Cho, Heecheol

Mathematical Finance(last update: December 19, 2008)

2nd Edition

Comments are Welcome(hccho@math.snu.ac.kr).

Abstract

MTL

Mathematical Finance

Hee Cheol Cho

Under Construction 수학의 입장에서 financial model을 공부하면서, 이미 나와있는 formula에 대한 이해를 제대로 하기 위해, 수학을 전공한 사람으로서 증명을 해보고 정리를 하려고 시작한 일이 계속되어 상당한 양의 Notes가 되었습니다.

이 Notes의 처음 목적은 제가 필요할 때 참고할 수 있는 내용을 정리하는 것이었습니다. 이렇게 정리를 해 놓고 보니, 오류는 항상 눈에 거슬리는 존재입니다만, 언제나 저를 방심할 수 없게 만드는 것이기도 합니다.

개인적으로는 상품의 pricing에 관심이 많습니다. Monte Carlo는 쉽게 구현할 수 있습니다만, hedging을 하기 위해 필요한 greeks계산이 unstable합니다. Monte Carlo의 대안으로 저는 Finite Difference Method(FDM)를 좋아합니다. 가격만 구한다면 Monte Carlo로 충분할 수 있지만, Monte Carlo로 Gamma까지 안정적으로 구할 수는 없지요. Single-Asset에 대한 FDM은 Multi-Asset에 비하면 상대적으로 쉬울 수도 있습니다. Multi-Asset에 대한 FDM에서 boundary condition은 항상 골치거리입니다.

Monte Carlo 방법은 조금만 공부해도 누구나 할 수 있는 방법입니다. 한 발 더 나아가는 길은 FDM 같이 편미분 방정식을 직접 다루는 방법에 익숙해지는 것이 아닐까 생각합니다.

그러나, underlying의 개수가 늘어나면, FDM은 속도의 한계를 드러냅니다(The Curse of Dimensionality). 그래서 Monte Carlo 방법을 미워할 수 만은 없지요. Malliavin Calculus 같이 Monte Carlo를 보완하는 방법이 있습니다만, 제대로 공부해 보지 않아 뭐라고 말할 수가 없네요...

이 Notes의 내용중 많은 부분은 직접 증명하고 유도하였으나, 일부분은 여러 대가들의 책과 논문을 인용하기도 했습니다.

금융수학을 공부하는 분들에게 도움이 되기를 바랍니다.

2004년 12월 정든 학교를 떠나며.

✎ 2nd Edition
Malliavin Calculus 부분을 추가하였습니다. Malliavin Calculus를 이용한 Barrier Option의 Greeks 계산에 관한 석사 논문을 쓴 서울대 수학과의 정두섭군은 Malliavin Calculus를 공부하고, 정리하는데 많은 도움을 주었습니다. 지면을 통해 감사의 마음을 전합니다.
✎ Acknowledgments
많은 오타를 수정해준 후배 김경범에게 고마움을 전합니다.

2005년 8월.

Typeset by LaTeX.

Contents

Abstract **i**

I Introduction to Mathematical Finance **1**

1 Basics **2**

1.1 Preliminary . . . 2

1.1.1 Basis risk . . . 2

1.1.2 Optimal Hedge Ratio . . . 2

1.1.3 Forward Price . . . 3

1.1.4 Value of Forward Contracts . . . 3

1.2 Put-Call Parity . . . 5

1.3 Portfolio Insurance . . . 6

1.4 Incomplete Market . . . 7

1.5 American & European Options . . . 8

1.5.1 Put Options with interest rate $r = 0$ . . . 8

1.5.2 Put Options with interest rate $r > 0$ . . . 8

1.5.3 Call Options with interest rate $r > 0$ . . . 9

1.6 Binomial Tree . . . 10

1.6.1 Matching Volatility with u and d . . . 11

1.6.2 General Tree Construction . . . 12

1.7 Binomial Tree & Geometric Brownian Motion . . . 15

1.7.1 Under Physical Measure . . . 15

1.7.2 Under Risk-Neutral Measure . . . 16

1.8 The Distribution of GBM . . . 18

1.8.1 Why Geometric Brownian Motion? . . . 18
1.8.2 Martingale & Dividend . . . 19
1.8.3 Discretization . . . 20
1.9 The Black-Scholes Model . . . 21
1.10 Black-Scholes Formula . . . 25
1.10.1 Method 1: Tedious algebra . . . 25
1.10.2 Method 2: Measure change(Using Girsanov Theorem) . . . 28
1.10.3 Method 3: Alternative Integration . . . 30
1.10.4 The Generalization of Black-Scholes Formula . . . 32
1.11 The Greek Letters . . . 34
1.12 Estimating Volatility vs. Drift . . . 41
1.13 The Market Price of Risk . . . 42
1.14 Market Price of Risk(λ), Measure Change and Numeraire . . . 43
1.14.1 Single Asset . . . 43
1.14.2 Multi-Assets . . . 44
1.15 Change of Numeraire . . . 46
1.16 Series Solution . . . 52
1.17 Static Hedging . . . 54
1.18 Financial Instruments . . . 55
1.19 Portfolio Fitting . . . 57

2 SDE & PDE **58**
2.1 Basics . . . 58
2.2 Some Examples . . . 58
2.3 Black-Scholes PDE . . . 65
2.3.1 Single Asset . . . 65
2.3.2 Two Assets with dividend . . . 66
2.3.3 Boundary Conditions . . . 67
2.4 BS PDE & Heat Equation . . . 69
2.4.1 Method 1: $u_s + \frac{1}{2}u_{yy} = 0$. . . . 69
2.4.2 Method 2(Time Dependent): $u_s = u_{yy}$. . . . 72
2.4.3 Method 3(Time Independent): . . . 74
2.4.4 Double Asset(Converting to the heat equation) . . . 77

2.5 A Generalization of The Black-Scholes PDE . . . 86
2.5.1 Derivation of Differential Equation with Multi-Asset . . . 86
2.5.2 Time-Dependent Parameters . . . 87
2.6 Markov Process & The Kolmogorov Equations . . . 90
2.7 Dividend Paying Assets . . . 91
2.7.1 Boundary Conditions . . . 91

II Pricing Financial Instruments **92**

3 American Options **93**
3.1 The Perpetual American Put . . . 93
3.2 Mathematical Formulation . . . 96
3.3 Buyer's & Seller's Price . . . 96
3.4 Optimal Stopping & American Options(discrete) . . . 98
3.4.1 The Snell Envelope . . . 98
3.4.2 American & European Option(discrete) . . . 100
3.4.3 American & European Option(continuous) . . . 102
3.5 Obstacle Problems . . . 103
3.6 Complementarity Problem For The American Put . . . 106
3.7 FDM for American Option . . . 107
3.7.1 Early Exercise and the Explicit Method . . . 107
3.7.2 Early Exercise and Crank-Nicolson . . . 108
3.7.3 The Projected SOR . . . 108
3.7.4 Implementation Example(Bermudan Style Option) . . . 109
3.8 Analytic Approximation to American Option Prices . . . 110
3.8.1 Barone, Adesi and Whaley's Approach . . . 110

4 Volatility **113**
4.1 Time Dependent Volatility . . . 113
4.2 Volatility Surface . . . 116
4.2.1 Local volatility with Black-Scholes Formula . . . 117
4.3 Stochastic Volatility . . . 121
4.4 The GARCH Models . . . 123

4.5 The Heston Model 126
4.6 The Fundamental Transform 127
4.7 Estimating Volatilities 128

III Exotic Options 129

5 Asian Options 130
5.1 Continuous Sampling(Geometric Average): The Pricing Equation 131
5.1.1 Reduction to a One-Dimensional Equation 134
5.2 Discrete Sampling(Geometric Average): The Pricing Equation 136
5.3 Arithmetic Average: Turnbull and Wakeman Approximation 138
5.4 Arithmetic Average: Levy's Approximation 141

6 Path Dependent Options 142
6.1 MIN-MAX Distribution of a Brownian Motion with Drift 142
6.2 Double Barrier Hitting Time Distribution of a BM with Drift 148
6.3 Double Barrier Hitting Time Distribution of a GBM 153
6.3.1 Four Basic integrals 154
6.4 Distribution of a Brownian Motion with Double Barrier 158

7 Lookback Options 165
7.1 Lookback Put 165
7.2 Lookback Call 169
7.3 Forward Lookback Call 172
7.4 Forward Lookback Put 175

8 Barrier Options 178
8.1 Reflection Principle 179
8.2 Single Barrier Options 187
8.3 Alternative Approach To Barrier Options 190
8.3.1 General In-Out Contracts 193
8.3.2 In-Out Bond 195
8.3.3 In-Out Calls 197
8.3.4 In-Out Puts 201

8.3.5 In-Out Asset . . . 204
8.3.6 In-Out Cash-Or-Nothing-Call . . . 206
8.3.7 In-Out Cash-Or-Nothing-Put . . . 208
8.3.8 In-Out Asset-Or-Nothing-Call . . . 210
8.3.9 In-Out Asset-Or-Nothing-Put . . . 212
8.4 Double Barrier Options . . . 214
8.4.1 Double Barrier Knock Out Call and Put . . . 214
8.4.2 Double Knock Out Binary . . . 218
8.4.3 American Binary Knock Out Option . . . 220

9 Miscellaneous Exotic Options **224**
9.1 Compound Option . . . 224
9.2 Chooser Option . . . 229
9.3 Binary(Digital) Option . . . 230
9.4 Asset or Nothing Option . . . 230
9.5 Forward Start Options . . . 230
9.6 Shout Option . . . 231
9.7 One-Touch Option . . . 231
9.8 Relative Digital Option . . . 237
9.9 Relative Outperformance Options . . . 238
9.10 Digital Options on Best or Worst of Two Assets . . . 239
9.11 Options on Best or Worst of Two Assets . . . 241

IV Interest Models **245**

10 Interest Rate Models **246**
10.1 Basics . . . 246
10.2 Constant Interest Rate . . . 247
10.3 Principal Components of Interest Rates . . . 253
10.4 Interest Rate Derivatives . . . 255
10.4.1 Black's Model . . . 255
10.4.2 Bond Options . . . 256
10.4.3 Caps And Floors . . . 257

10.4.4 Swap Options . . . 258
10.5 Convexity Adjustments . . . 258
10.6 Non-Constant Interest . . . 260
10.6.1 Generalities . . . 260
10.6.2 Swaps . . . 263
10.7 Instantaneous Short Rate . . . 269
10.7.1 Measure Change . . . 273
10.7.2 Replicating Strategies . . . 274
10.8 Interest Rate Products . . . 276
10.8.1 Forward Condtract . . . 276
10.8.2 Multiple Payment Contracts . . . 276
10.8.3 Floating Rate Notes . . . 277
10.8.4 Swaps . . . 278
10.8.5 Forward Swaps . . . 278
10.9 Short Rate Models . . . 280
10.10 Parameter Estimations(= Inverting The Yield Curve = Yield Curve Fitting) . . . 282
10.11 Affine Term Structures . . . 283
10.11.1 Vasicek Models(Affine Term Structure) . . . 286
10.11.2 Cox-Ingersoll-Ross Model(Affine Term Structure) . . . 290
10.11.3 Ho and Lee Model(Affine Term Structure) . . . 294
10.11.4 (Simplified)Hull-White Model(Affine Term Structure, Extended Vasicek) . . . 297
10.12 Some Comments about One Factor Short Rate Models . . . 303

11 Heath-Jarrow-Morton Model **304**
11.1 Single-factor HJM . . . 304
11.1.1 HJM in terms of the short rate . . . 307
11.1.2 HJM from Bond Price Dynamics . . . 308
11.1.3 Ho and Lee Model in HJM . . . 310
11.1.4 Hull-White Model in HJM(Extended Vasicek) . . . 313
11.2 Forward Measure . . . 315
11.2.1 Extended Black-Scholes Formula . . . 318
11.2.2 Valuing European Options on Zero-Coupon Bonds . . . 320
11.2.3 Valuing European Options on Coupon Bearing Bonds . . . 326

11.2.4 Contingent Claim on $r(T)$ under Hull-White Model . . . 327
11.2.5 Caps and Floors . . . 328
11.2.6 Swaption . . . 330

12 Brace-Gatarek-Musiela Model **331**
12.1 Musiela Parametrization . . . 332
12.2 Single-Factor Brace-Gatarek-Musiela Model . . . 334
12.3 Multi-Factor Brace-Gatarek-Musiela Model . . . 339
12.3.1 Pricing Derivatives . . . 345
12.3.2 Discretization For Simulation . . . 347
12.3.3 Volatility Structure and Calibration . . . 350

13 Interest Rate Trees **357**
13.1 Hull-White Two-Stage Procedure . . . 357
13.1.1 First Stage . . . 357
13.1.2 Second Stage . . . 359

V Implements **361**

14 Tree Approaches **362**
14.1 Path-Dependent Derivatives(Hull-White) . . . 362
14.2 Lookback Options(Reiner) . . . 362
14.3 Barrier Options . . . 362

15 Finite Difference Methods **363**
15.1 The Explicit FDM & Trinomial Tree . . . 364
15.1.1 The Transformation of Variables . . . 366
15.1.2 Stability of the Explicit FDM . . . 367
15.2 The Crank-Nicolson Scheme . . . 368
15.2.1 Implementation(Crank-Nicolson) . . . 369
15.2.2 Implementation(θ-Method) . . . 372
15.3 Alternative Approach to Explicit FDM . . . 374
15.4 Fully Implicit Scheme . . . 377
15.5 Alternative Approach to Crank-Nicolson . . . 378

15.6 θ-Method . . . 382
15.7 Greeks with Transformation . . . 388
15.8 Alternating Direction Implicit(ADI) . . . 389
15.8.1 Peaceman-Rachford Scheme . . . 389
15.8.2 D'Yakonov scheme . . . 390
15.8.3 Douglas-Rachford Scheme . . . 391
15.8.4 ADI with Mixed Derivatives . . . 392
15.9 Operator Splitting Method . . . 394
15.10 How To Handle Jumps in FDM . . . 398
15.10.1 Discrete Cashflow(a jump in the value) . . . 398
15.10.2 Discrete Dividend(a jump in the underlying asset) . . . 399
15.11 Summing Up . . . 400
15.11.1 Heat Equation Version . . . 400
15.12 Boundary Conditions . . . 401
15.12.1 Pricing Barrier Options In The PDE Framework . . . 401

16 Monte Carlo Simulations **403**
16.1 Introduction to Monte Carlo Method . . . 403
16.1.1 What is Monte Carlo Simulation? . . . 403
16.1.2 Pros And Cons . . . 404
16.2 Implementing Monte Carlo Simulation . . . 405
16.2.1 Example: Black-Scholes Call Option Model . . . 405
16.3 Examples . . . 407
16.3.1 Delta Hedging Simulation . . . 407
16.3.2 Path-Dependent Asian Option . . . 408
16.3.3 Interest Futures . . . 408
16.3.4 Hedging Sensitivities . . . 409
16.4 Efficiency, Speeding Up & Variance Reduction Methods . . . 410
16.4.1 Antithetic Variates . . . 411
16.4.2 Control Variates Technique . . . 412
16.4.3 Moments Matching Methods . . . 413
16.5 Multivariate Monte Carlo . . . 414
16.5.1 Cholesky Decomposition . . . 415

16.5.2 Example: Value at Risk . . . 416
16.6 Generating Distributions . . . 417
16.6.1 Inverse Transform Method . . . 417
16.6.2 Acceptance-Rejection Method . . . 418
16.6.3 Generating Normal Distribution . . . 419
16.7 Quasi Monte Carlo Simulation . . . 422
16.7.1 Halton Sequence . . . 422
16.7.2 Faure Sequence . . . 424
16.7.3 Sobol's Sequence . . . 426
16.8 Brownian Bridge . . . 431
16.9 Example: Discrete Delta Hedging . . . 433

17 Finite Element Methods **434**
17.1 The Principle of Weighted Residuals . . . 434
17.2 The Galerkin method . . . 437

18 The Malliavin Calculus **442**
18.1 Introduction to the Malliavin Calculus . . . 442
18.2 Simple Examples . . . 443
18.3 Likelihood Ratio Method . . . 448
18.4 Mathematical Preliminary . . . 450
18.5 Malliavin Calculus . . . 456
18.6 The European-Style Options . . . 460
18.7 Discrete Path Dependent Options . . . 473
18.8 Applications of Malliavin Calculus to Barrier Options . . . 483

VI Appendex **490**

Appendix **491**

A The Multivariate Normal Distribution **491**

B The Brownian Motion **494**

C The Girsanov Theorem **496**

D Transforming Variables **500**

E The Linear Stochastic Differential Equations **502**

F The Infinitesimal Operator **504**

F.1 Kolmogorov Backward Equations 506

F.2 Kolmogorov Forward Equations 506

G Tridiagonal Matrix **507**

G.1 An LU solver for tridiagonal systems 508

H Principal Components Analysis **510**

I C++ Program Codes **516**

I.1 Solving Tridiagonal Matrix 516

I.2 Explicit FDM 518

I.3 Projected SOR 521

I.4 Callable Equity Swap 523

J Normal Distribution Table **526**

K Notations **527**

Index **531**

List of Tables

2.1 Change of variables with $D_0 = 0$. 75
2.2 Change of variables with dividend 76
2.3 Relevant rate 84
2.4 Parameters 84

3.1 Bermudan Option 109

4.1 Implied volatilities of options. 115

5.1 The average rate of Asian options 130
5.2 Asian options 131

10.1 Comparision between Compoundings 251
10.2 Interest data 253
10.3 Daily differences of interest rates 253
10.4 Principal Components of difference of interest rate 253
10.5 Eigenvalues of the Correlation Matrix 254
10.6 Stock forward vs bond forward 260

11.1 Vasicek vs Hull-White Model 315
11.2 Option on Coupon-Bearing Bond 326

12.1 3-month forward LIBOR rates from 2001-07-31 to 2007-1-23. 352
12.2 Annualized log return for 3-month forward LIBOR: $\sqrt{250}\log\frac{L_n(0)^{i+1}}{L_n(0)^i}$ 352
12.3 Simple Statistics. 352
12.4 Covariance Matrix 353
12.5 The eigenvectors scaled to have the square of length equal to their eigenvalues 353

12.6 Eigenvalues 353

13.1 Probabilities for each branching pattern. 359

15.1 Convergence rate and stable condition of finite difference methods 368
15.2 1-dimensional FDM schemes 400
15.3 2-dimensional ADI(Peaceman-Rachford) schemes 400

16.1 Gray codes 429

18.1 Summary 488

H.1 미국 50개 주의 범죄 자료 512
H.2 표준화된 미국 50개 주의 범죄 자료 514
H.3 공분산 행렬 514
H.4 상관계수 행렬 514
H.5 Principal components(상관계수) 514
H.6 Eigenvalues of the Correlation Matrix 515

List of Figures

1.1 Portfolio Insurance 6
1.2 Two step tree 10
1.3 Two equivalent methods for building constant-volatility binomial tree. 12
1.4 Trinomial tree 13
1.5 Three equivalent methods for building constant-volatility trinomial tree. 14
1.6 An Extension of Black-Scholes Theory 23
1.7 Value of a European call option with $r = 0.1, \sigma = 0.2, K = 40, T = 1$ 27
1.8 Delta of call & put option 35
1.9 Delta w.r.t. underlying price 35
1.10 Delta w.r.t. time to maturity 35
1.11 Delta w.r.t. volatility 36
1.12 Gamma of Call Options 36
1.13 Call Value & Theta 37
1.14 Put Value & Theta 37
1.15 Vega of Option 38
1.16 Option Value with Interest($\tau = 0.5$) 39
1.17 Convexity of call option 40

2.1 Computation regions 85
2.2 Solution for ELS; valuation date-9 January 2005) 85
2.3 Option with dividend. 91

3.1 The solution to the perpetual American put. 95
3.2 Example of an obstacle problem. 103
3.3 Transformed put option's payoff(obstacle) 107

3.4 American option pricing. 108
3.5 Bermudan option pricing. 109
3.6 The price of Bermudan option($r = 10\%, \sigma = 20\%$). 109

4.1 Volatility term structure 115
4.2 Fat tails 121

6.1 Integration region 144
6.2 The absorbed process 145

7.1 Lookback option 166

8.1 Integration region 181
8.2 Conditional Probability: $S_0 = 100, H = 120, L = 60, \sigma = 0.3, T = 0.5$ 186
8.3 $P_L(x; K)$, with $L < K$. 192
8.4 payoffs of barrier call option 197
8.5 $C^L(x; L)$ with $K < L$ 199
8.6 $P^L(x; K)$. 202
8.7 Up & out cash or nothing call 207
8.8 Down & out cash or nothing put 208
8.9 Up & out cash or nothing put 209
8.10 Down & out asset or nothing call 210
8.11 Up & out asset or nothing call 211
8.12 Down & out asset or nothing put 212
8.13 Up & out asset or nothing put 213
8.14 Up-Out-Down-In Call Option 218
8.15 Double Knock-Out Binary 219
8.16 American Binary Knock-Out Option 221

9.1 The payoff of binary option 230

10.1 Calculating accrued interest. 247
10.2 Principal component of interest(spot rate) 254
10.3 Floating Rate 1 263
10.4 Floating Rate 2 263

10.5 Fixed Rate 1 263
10.6 Fixed Rate 2 263
10.7 Swap 264
10.8 $L(T_i)$ 265
10.9 Time horizon 265
10.10 Discounting Factor 285

11.1 Volatility Structure in the Hull-White Model 315
11.2 Floating Rate 328

12.1 Forward LIBOR 334
12.2 $\eta(t)$ is right continuous. 339
12.3 numeraire associated with spot measure 341
12.4 Lognormal Distribution of $L_n(t)$ 344
12.5 Time Horizon 345
12.6 Volatility factors 351
12.7 The first three eigenvectors 354
12.8 $\|\sigma(t)\|^2$ with $\|\sigma(i)\| =$ standard deviation of time bucket 354
12.9 $\overline{\sigma}(t)$ 354
12.10 Volatility matrix 356
12.11 Volatility factors 356

13.1 Alternative branching methods in a trinomial tree. 358
13.2 Tree for R^* in Hull-White model($\sigma = 0.01, a = 0.1, \Delta t = 1$) 359

15.1 Explicit Finite-Difference Discretization. 364
15.2 Probabilistic Interpretation. 366
15.3 Extrapolation from $V_{k+1,1}$ and f 380
15.4 Numerical solutions of 2-dimension heat equation. 395
15.5 Numerical solutions for a 2-dimension ELS. 398
15.6 Alternating direction implicit method 400
15.7 Boundary condition of a down-and-in barrier option 401

16.1 Monte Carlo Simulation을 이용한 π 계산 403
16.2 A sample path with geometric Brownian motion: S(0)=1,drift=10%, $\sigma = 40\%$ 406

16.3 Graphical example of the acceptance-rejection method 419

16.4 A pseudo code for polar rejection method 420

16.5 200 pairs 2D-random numbers 422

16.6 1000 points of the Halton sequence 423

16.7 First 10000 points of leaped Halton sequence with bases 109 and 113 424

16.8 Pseudo code for generating matrix $\mathbf{C}^{(i)}$ with size $r \times r$ 425

16.9 Brownian bridge construction 431

16.10Discrete hedging errors 433

17.1 Hat function: non-equidistant grid 436

17.2 $n_i(x)$ 438

18.1 Delta for digital option: $S_0 = 40, K = 40, r = 10\%, \sigma = 20\%, Div = 1\%, T = 0.5$, N = number of simulations. 463

18.2 Gamma for digital option: $S_0 = 40, K = 40, r = 10\%, \sigma = 20\%, Div = 1\%, T = 0.5$, N = number of simulations. 463

18.3 Numerical results of Multi-Chance Early Redemption Note: redeemable period= 0.5, maturity= 3,$K = 100, \sigma = 30\%, r = 5\%$, coupon rate= $4.5\% \times 2$. 482

Part I

Introduction to Mathematical Finance

Chapter 1

Basics

1.1 Preliminary

1.1.1 Basis risk

basis:= spot price of asset to be hedged - futures price of contract used.

1. S_1, S_2: spot price at time t_1, t_2
2. F_1, F_2: futures price with maturity T at time t_1, t_2
3. b_1, b_2: basis at time t_1, t_2.
4. $S_1^*, S_2^*, F_1^*, F_2^*$: different asset

t_1(현재)시점에서 t_2시점에 물건을 팔 필요가 있을 경우, 만기 $T(> t_2)$인 선물을 가격 F_1으로 short position을 취한 후, t_2시점에 close out 하면, 이자율을 무시한 payoff는

$$S_2 + F_1 - F_2 = F_1 + b_2.$$

만약 다른 asset으로 hedge를 한다면, t_2시점의 payoff는

$$S_2 + F_1^* - F_2^* = F_1^* + (S_2^* - F_2^*) + (S_2 - S_2^*) = F_1^* + b_2^* + (S_2 - S_2^*).$$

1.1.2 Optimal Hedge Ratio

When the hedger is long the asset and short h futures(F) at $t = 0$, the payoff at t is

$$V_t = S_t - h(F_t - F_0).$$

The change in the value of the hedger's position is

$$\begin{aligned} \Delta V &= \Delta S - h\Delta F \\ \mathrm{Var}[V] &= \sigma_S^2 + h^2\sigma_F^2 - 2h\rho\sigma_S\sigma_F \end{aligned}$$

To get the minimum of Var[V], differentiate Var[V]:

$$\begin{aligned}\frac{\partial \mathrm{Var}[V]}{\partial h} &= 2h\sigma_F^2 - 2\rho\sigma_S\sigma_F = 0, \\ h &= \rho\frac{\sigma_S}{\sigma_F}.\end{aligned}$$

1.1.3 Forward Price

The forward price is given by

$$F_0 = S_0 e^{rT}, \quad F_t = S_t e^{r(T-t)}.$$

To derive this relation, consider the following cases:

1. $F_0 > S_0 e^{rT}$:
   (a) Borrow S_0 at an interest rate r for T years.
   (b) Buy one futures.
   (c) Short a forward contract.

   After T years, the investor makes a profit $F_0 - S_0 e^{rT}$.
2. $F_0 < S_0 e^{rT}$:
   (a) Short 1 stock with S_0.
   (b) Invest S_0 at the risk-free rate r.
   (c) Take a long position in a forward contract.

   After T years, the investor makes a profit $S_0 e^{rT} - F_0$.
3. What if short sales are not possible? In this case an investor who owns one stock can
   (a) Sell the stock for S_0
   (b) Invest the proceeds at interest rate r for T years.
   (c) Take a long position in a forward contract.

   After T years, the cash invested has grown to $S_0 e^{rT}$. The stock is repurchased for F_0, and the investor makes a profit of $S_0 e^{rT} - F_0$ relative to the position the investor would have been in if the stock had been kept.

1.1.4 Value of Forward Contracts

The Value of Forward Contract: Martingale Viewpoint

Consider the forward contract with strike K, maturity T. Let the value of this forward contract at t be f_t. Then f_t/B_t is a martingale. Hence

$$\begin{aligned}f_t &= B_t E[(S_T - K)/B_T|\mathcal{F}_t] \\ &= B_t\left(\frac{S_t}{B_t} - \frac{K}{B_T}\right), \quad \frac{S_t}{B_t}: \text{ martingale,}\end{aligned}$$

$$= \quad S_t - Ke^{-r(T-t)}.$$

Let F_t be the forward Price[1] for the contract with maturity T. To make $f_t = 0$,

$$\begin{aligned} S_t - F_t e^{-r(T-t)} &= 0, \\ F_t &= S_t e^{r(T-t)}. \end{aligned}$$

It follows from this that since $\frac{S_t}{B_t}$ is a martingale, $F_t = \frac{S_t}{B_t} B_T$ is also martingale.

$$\boxed{\frac{S_t}{B_t}, \quad \frac{f_t}{B_t}, \quad F_t \text{ are martingales.}}$$

Also we get that

$$f_t \quad = \quad F_t e^{-r(T-t)} - Ke^{-r(T-t)} = (F_t - K)e^{-r(T-t)}.$$

다른 관점에서 보면, forward에 대한 long position인 경우, t시점에서 반대 position을 취하면, T시점에서의 가치는 $F_t - K$ 이므로, $(F_t - K)$를 할인하면 된다. 즉,

$$f_t = e^{-r(T-t)}(F_t - K).$$

Since F_t is a martingale,

$$\begin{aligned} F_t &= E[F_T|\mathcal{F}_t] \\ &= E[S_T|\mathcal{F}_t] \\ &= E\left[\frac{S_T}{B_T} \times B_T|\mathcal{F}_t\right] \\ &= \frac{S_t}{B_t} \times B_T \\ &= S_t e^{r(T-t)}. \end{aligned}$$

Furthermore we have, by Ito lemma,

$$\begin{aligned} dF_t &= dS_t e^{r(T-t)} + F_t(-rdt) \\ &= F_t(rdt + \sigma dW_t) - rF_t dt \\ &= \sigma F_t dW_t. \end{aligned}$$

[1] The delivery price in a forward contract that causes the contract to be worth zero.

1.2 Put-Call Parity

✎ **Stock without cost of carry**
Consider a portfolio consisting of

1 stock + 1 put option - 1 call option

with same maturity T and same strike price K. This portfolio is worth K at options' maturity regardless of stock price. Hence it is easily seen that

$$\begin{aligned} S_T + P_T - C_T &= K \\ E[S_T + P_T - C_T \mid \mathscr{F}_t] &= E[K \mid \mathscr{F}_t], \\ e^{rT} E\left[\frac{S_T}{e^{rT}} + \frac{P_T}{e^{rT}} - \frac{C_T}{e^{rT}} \mid \mathscr{F}_t\right] &= K, \\ e^{rT}\left(\frac{S_t}{e^{rt}} + \frac{P_t}{e^{rt}} - \frac{C_t}{e^{rt}}\right) &= K, \\ S_t + P_t - C_t &= Ke^{-r(T-t)}. \end{aligned}$$

✎ **Stock with cost of carry(dividend)**
Suppose that the stock price process, S_t, follows

$$d\Big(S_t e^{qt}\Big) = S_t e^{qt}\Big(rdt + \sigma dW_t\Big).$$

By virtue of Ito formula, this SDE is equivalent to

$$dS_t = S_t\Big((r-q)dt + \sigma dW_t\Big).$$

Then we find that

$$\left\{\frac{S_t e^{qt}}{e^{rt}}\right\}_{t\geq 0}$$

is a martingale. In this case, the put-call parity is given as follows:

$$\begin{aligned} E[S_T + P_T - C_T|\mathscr{F}_t] &= E[K], \\ e^{rT} E\left[\frac{S_T e^{qT}}{e^{rT}}\, e^{-qT} + \frac{P_T}{e^{rT}} - \frac{C_T}{e^{rT}}\Big|\mathscr{F}_t\right] &= K, \\ e^{rT}\Big(\frac{S_t e^{qt}}{e^{rt}}\, e^{-qT} + \frac{P_t}{e^{rt}} - \frac{C_t}{e^{rt}}\Big) &= K, \\ S_t e^{-q(T-t)} + P_t - C_t &= Ke^{-r(T-t)}. \end{aligned}$$

1.3 Portfolio Insurance

Portfolio insurance is a strategy entering into trades to ensure that the value of portfolio will not fall down below a certain level.

At time t, put-call parity with maturity T is given by

$$P_t + S_t \quad = \quad C_t + Xe^{-r(T-t)}.$$

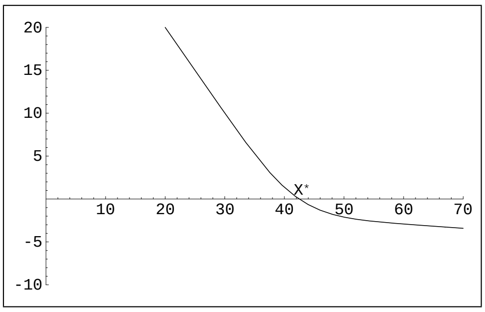
We cannot fine a put option such that

$$P(S_t, X^*, T-t, r, \sigma) + S_t \quad = \quad X^* e^{-r(T-t)},$$

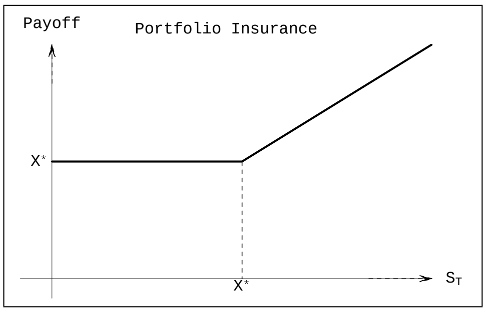
but, we can find a put option with exercise price X^* such that

$$P(S_t, X^*, T-t, r, \sigma) + S_t \quad = \quad X^*.$$

Note that the X^* always exists and $S_t \leq X^*$. Now, we can construct a portfolio with one stock and



(a) Finding Strike Price ($S_t = 40$)



(b) Final Payoff

Figure 1.1: Portfolio Insurance

one put option with exercise price X^*:

$$\text{one stock} + \text{one put option with exercise price } X^*$$

The initial portfolio value is X^* and the final payoff is given in Figure (1.1(b)).

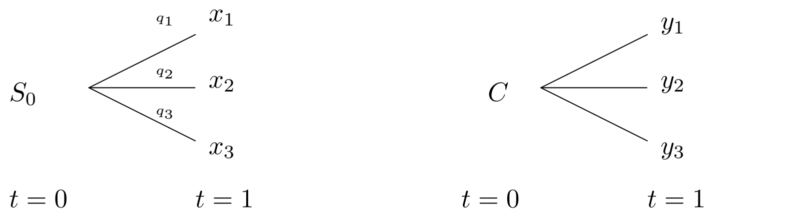
This portfolio is equivalent to the portfolio consist of call option and cash. In this equivalent portfolio, the call options value is $X^* - X^* e^{-r(T-t)}$ and cash amount is $X^* e^{-r(T-t)}$.

$$\boxed{\text{put option} + \text{Stock} \quad = \quad X^* \quad = \quad \text{call option}\Big(X^* - X^* e^{-r(T-t)}\Big) + \text{cash}\Big(X^* e^{-r(T-t)}\Big)}$$

1.4 Incomplete Market

Martingale measure의 존재성은 underlying asset의 현재-미래 값들간의 관계에서 주어진다.

✎ **Duality 1: seller's viewpoint**
$\mathbf{x} = (x_1, x_2, x_3)^t, \mathbf{y} = (y_1, y_2, y_3)^t$ are given[2].
$\mathbf{x}$ is the stock value distribution, $\mathbf{y}$ is contingent claim value distribution.



Stock price Value of the contingent claim

Let $\mathbf{q} = (q_1, q_2, q_3)^t$ be a martingale measure and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ be a replicating portfolio.

$$\begin{aligned}
\text{Seller's Price} &= \sup\{E_Q[X] : Q \text{ is a martingale measure.}\} \\
&= \inf\{\text{dominating portfolio with respect to } X\} \\
\text{Buyer's Price} &= \inf\{E_Q[X] : Q \text{ is a martingale measure.}\} \\
&= \sup\{\text{dominated portfolio by } X\}
\end{aligned}$$

Remark 1.4.1.

1. Martingale Measure의 존재성: $S_0 \in \text{Convex Hull}(S_1(\omega_1), \cdots, S_1(\omega_n))$.
2. Unique Martingale Measure(Complete Market): Every contingent claim is attainable.

---

[2]

$$\begin{array}{rl}
Max(Min)\ \mathbf{c}^t\mathbf{x} & Min(Max)\ \boldsymbol{\lambda}^t\mathbf{b} \\
A\mathbf{x} = \mathbf{b} & \boldsymbol{\lambda}^t A \geq \mathbf{c}^t\ (\leq) \\
\mathbf{x} \geq 0 & \boldsymbol{\lambda} :\ \text{free}
\end{array}$$

$$q_1 + q_2 + q_3 = 1$$
$$x_1 q_1 + x_2 q_2 + x_3 q_3 = S_0$$

$$\lambda_1 + \lambda_2 x_1 \geq y_1$$
$$\lambda_1 + \lambda_2 x_2 \geq y_2$$
$$\lambda_1 + \lambda_2 x_3 \geq y_3$$

$$\max \quad (y_1 q_1 + y_2 q_2 + y_3 q_3)$$
$$\begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} 1 \\ S_0 \end{bmatrix}$$
$$q_1, q_2, q_3 \geq 0.$$

$$\min \quad \lambda_1 + \lambda_2 S_0$$
$$\begin{bmatrix} \lambda_1 & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \end{bmatrix} \geq \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}$$

$$\max \quad \mathbf{y}^t \mathbf{q}$$
$$A\mathbf{q} = \mathbf{b}$$
$$\mathbf{q} \geq \mathbf{0}$$

$$\min \quad \boldsymbol{\lambda}^t \mathbf{b}$$
$$\boldsymbol{\lambda} A \geq \mathbf{y}^t$$
$$\boldsymbol{\lambda} : \text{ free}$$

1.5 American & European Options

1.5.1 Put Options with interest rate $r = 0$

European option value curve가 만기의 payoff 곡선보다 아래로 내려가는 경우가 없다. 따라서

$$\boxed{\text{American put} = \text{European put}}$$

1.5.2 Put Options with interest rate $r > 0$

American put option은 만기의 payoff보다 항상 위에 있지만, European put option은 discount하면 만기의 payoff보다 아래에 있다. European call option의 $S = 0$ 일때의 값은 $Xe^{-r(T-t)}$인 것으로 부터 알 수도 있다.

American put option은 exercise boundary 보다 낮은 주가에서는 즉시 행사해야 한다. 그 이유로는 즉시 행사하여, 실현된 이익으로 European call option을 사고, 그 차액을 얻는 것이 더 유리하기 때문이다.

$$\boxed{\text{American put} \geq \text{European put}}$$

1.5.3 Call Options with interest rate $r > 0$

American call = European call

1. t 시점에서 European call option 과 $Xe^{-r(T-t)}$의 deposit으로 구성된 portfolio와 주식 하나로 구성된 portfolio의 만기에서의 payoff를 보면

$$\max(S_T, X) \geq S_T.$$

   따라서,

$$c_t > S_t - Xe^{-r(T-t)}$$

   가 되어 European call option의 value curve는 $r > 0$인 경우, 만기의 payoff curve 보다 위에 놓이게 된다.

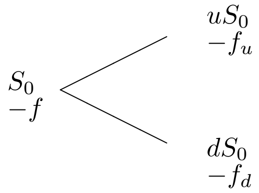
2. 만기(T)의 payoff와 $t(< T)$ 시점에서의 value curve가 convex increasing 이므로 discount 하면, discount 하기 전의 curve 보다 위에 놓인다.

3. American call option을 early exercise한다는 것은 만기의 payoff를 현재 취하는 것인데, 이렇게 되면 arbitrage가 발생한다. 따라서, American call option을 만기 전에 행사하면 Arbitrage Pricing Axiom에 위배된다.

4. American call option의 early exercise가 불합리한 또 다른 이유로, 행사가를 deposit하여 이자를 얻고, 만기에 행사하는 것이 더 유리한 것도 있고, 또한 주가가 떨어질 위험에 대비하는 측면도 있다.

1.6 Binomial Tree

Let S_t be the underlying stock and f be the price of contingent claim [3]. we can choose Δ such that it makes portfolio with Δ stock and -1 contingent claim risk free.



$$\Delta S_0 u - f_u = \Delta S_0 d - f_d = (\Delta S_0 - f)e^{rT}.$$

Hence we get Δ and f as follows:

$$\begin{aligned}
\Delta &= \frac{f_u - f_d}{S_0 u - S_0 d}, \\
f &= e^{-rT}\big(pf_u + (1-p)f_d\big), \qquad \text{where } p = \frac{e^{rT} - d}{u - d}.
\end{aligned}$$

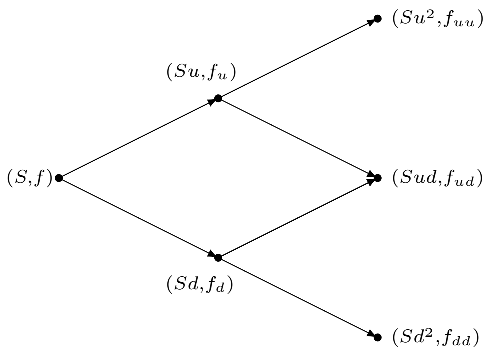


Figure 1.2: Two step tree

$$\begin{aligned}
p &= \frac{e^{r\Delta t} - d}{u - d} \\
f_u &= e^{-r\Delta t}\big(pf_{uu} + (1-p)f_{ud}\big), \\
f_d &= e^{-r\Delta t}\big(pf_{ud} + (1-p)f_{dd}\big), \\
f &= e^{-r\Delta t}\big(pf_u + (1-p)f_d\big) \\
&= e^{-2r\Delta t}\big(p^2 f_{uu} + 2p(1-p)f_{ud} + (1-p)^2 f_{dd}\big).
\end{aligned}$$

[3] Note that u, d, S_0, f_u and f_d are given, but Δ and f are unknown. We can find Δ and f from known parameters.

1.6.1 Matching Volatility with u and d

From $pu + (1-p)d = e^{r\Delta t}$, we find that

$$p = \frac{e^{r\Delta t} - d}{u - d}. \tag{1.6.1}$$

The variance of the stock price return is[4]

$$pu^2 + (1-p)d^2 - \big(pu + (1-p)d\big)^2 = \sigma^2 \Delta t. \tag{1.6.2}$$

The LHS can be written as

$$\begin{aligned}
pu^2 + (1-p)d^2 - \big(pu + (1-p)d\big)^2 &= pu^2 + (1-p)d^2 - p^2u^2 - (1-p)^2d^2 - 2p(1-p)ud \\
&= p(1-p)(u^2 + d^2) - 2p(1-p)ud \\
&= p(1-p)(u+d)^2 \\
&= \frac{e^{r\Delta t} - d}{u - d} \times \frac{u - e^{r\Delta t}}{u - d} \times (u+d)^2 \\
&= (e^{r\Delta t} - d)(u - e^{r\Delta t}).
\end{aligned}$$

To solve u, d, and p, we need at least three equations, but we only have two equations (1.6.1) and (1.6.2). Also thus assume that $d = 1/u$. If we solve the equations

$$\begin{cases} ud &= 1 \\ \sigma^2 \Delta t &= (e^{r\Delta t} - d)(u - e^{r\Delta t}) \end{cases}$$

then we get

$$\begin{aligned}
u &= \frac{1}{2} e^{-r\Delta t}\Big(1 + e^{2r\Delta t} + \Delta t \sigma^2 + \sqrt{-4e^{2r\Delta t} + (1 + e^{2r\Delta t} + \Delta t\sigma^2)^2}\Big) \\
&= 1 + \frac{\Delta t}{2}\sigma^2 + \sqrt{\Delta t}\sqrt{\sigma^2} + O(\Delta t^{3/2}) \\
&\approx e^{\sigma\sqrt{\Delta t}}, \\
d &= e^{-\sigma\sqrt{\Delta t}}.
\end{aligned}$$

This solution for u and d was proposed by Cox, Ross, and Rubinstein.

Remark 1.6.1. From Taylor series expansions, when terms of higher order than Δt are ignored, we have

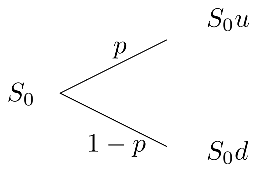
$$\begin{aligned}
u &= e^{\sigma\sqrt{\Delta t}} \quad \approx \quad 1 + \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2 \Delta t, \\
d &= e^{-\sigma\sqrt{\Delta t}} \quad \approx \quad 1 - \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2 \Delta t, \\
e^{r\Delta t} &\approx 1 + r\Delta t, \quad e^{2r\Delta t} \quad \approx \quad 1 + 2r\Delta t.
\end{aligned}$$

These approximations imply

$$\begin{aligned}
(e^{r\Delta t} - d)(u - e^{r\Delta t}) &= e^{r\Delta t}(u+d) - ud - e^{2\mathbf{r}\Delta t} \\
&= (1 + r\Delta t)(2 + \sigma^2 \Delta t) - 1 - (1 + 2r\Delta t) + o(\Delta t) \\
&= \sigma^2 \Delta t + o(\Delta t).
\end{aligned}$$

[4]The variance of the return $\frac{\Delta S}{S_0} = \frac{S_{\Delta t} - S_0}{S_0}$ is equal to the variance of $\frac{S_{\Delta t}}{S_0}$.
The volatility of stock price, σ, is defined so that $\sigma\sqrt{\Delta t}$ is the standard deviation of the return on the stock price in a short period of time of length Δt.

1.6.2 General Tree Construction



There are three parameters u, d, and p. Since we are only trying to match mean and variance we have a free choice for one of the parameters.

1. Cox, Ross and Rubinstein set the jump sizes to be equal, which leads to

$$\begin{aligned} u &= e^{\sigma\sqrt{\Delta t}}, \\ d &= e^{-\sigma\sqrt{\Delta t}}, \\ p &= \frac{1}{2} + \frac{r - \frac{1}{2}\sigma^2}{2\sigma}\sqrt{\Delta t}. \end{aligned}$$

2. Jarrow and Rudd chose the probabilities to be one-half which leads to

$$\begin{aligned} u &= e^{\left(r-\frac{1}{2}\sigma^2\right)\Delta t + \sigma\sqrt{\Delta t}}, \\ d &= e^{\left(r-\frac{1}{2}\sigma^2\right)\Delta t - \sigma\sqrt{\Delta t}}, \\ p &= \frac{1}{2}. \end{aligned}$$

There are in general an infinite number of equivalent binomial trees, all representing the same discrete constant volatility world. This is due to a freedom in the choice of overall growth of the price at tree nodes. The methods described here will all converge to the same theory, i.e. the constant volatility Black-Scholes theory, in the continuous limit.

$$\begin{aligned} u &= e^{\sigma\sqrt{\Delta t}}, \\ d &= e^{-\sigma\sqrt{\Delta t}}, \\ p &= \frac{1}{2} + \frac{r - \frac{1}{2}\sigma^2}{2\sigma}\sqrt{\Delta t}. \end{aligned}$$

(a) CRR binomial tree

$$\begin{aligned} u &= e^{\left(r-\frac{1}{2}\sigma^2\right)\Delta t + \sigma\sqrt{\Delta t}}, \\ d &= e^{\left(r-\frac{1}{2}\sigma^2\right)\Delta t - \sigma\sqrt{\Delta t}}, \\ p &= \frac{1}{2}. \end{aligned}$$

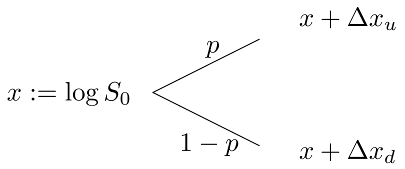
(b) JR binomial tree

Figure 1.3: Two equivalent methods for building constant-volatility binomial tree.

✎ **Log-Transformed Tree**

A problem with above two formulations is that the approximation is only good over a small time interval, Δt, we cannot freely choose arbitrarily large time steps. In order to solve this problem, we formulate the model in terms of the logarithms of the asset price. We now equate the mean and variance of the binomial process for x with mean $\left(r - \frac{1}{2}\sigma^2\right)\Delta t$ and variance $\sigma^2\Delta t$:

$$E[\Delta x] = p\Delta x_u + (1-p)\Delta x_d = \nu\Delta t,$$



$$E[\Delta x^2] \;=\; p\Delta x_u^2 + (1-p)\Delta x_d^2 \;=\; \sigma^2 \Delta t + \nu^2 \Delta t^2,$$

where $\nu = r - \frac{1}{2}\sigma^2$. As we have already mentioned, we have a free choice of one of the parameters. The two obvious choices are to set the probabilities to be equal to one-half or to set the jump size to be equal.

1. Equal probabilities of one-half leads to the following:

$$\begin{aligned} \Delta x_u &= \nu \Delta t + \sigma \sqrt{\Delta t}, \\ \Delta x_d &= \nu \Delta t - \sigma \sqrt{\Delta t}. \end{aligned}$$

2. Equal jump sizes, i.e. $\Delta x_d = -\Delta x_u$, lead to:

$$\begin{aligned} \Delta x &= \sqrt{\sigma^2 \Delta t + \nu^2 \Delta t^2} \\ p &= \frac{1}{2} + \frac{1}{2}\frac{\nu \Delta t}{\Delta x}. \end{aligned}$$

   This solution was proposed by Trigeorgis.

✎ **Trinomial Tree**
Δx cannot be chosen independently of Δt and a good choice is $\Delta x = \sigma\sqrt{3\Delta t}$.

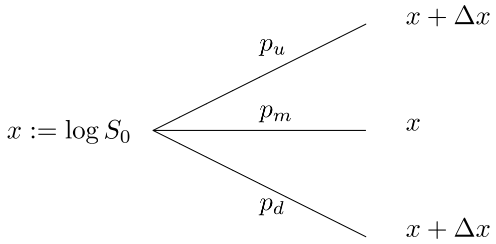


Figure 1.4: Trinomial tree

Matching first two moments leads to

$$\begin{aligned} E[\Delta x] &= p_u \Delta x + p_m \times 0 + p_d(-\Delta x) \;=\; \nu \Delta t, \\ E[\Delta x^2] &= p_u \Delta x^2 + p_m \times 0^2 + p_d(-\Delta x)^2 \;=\; \sigma^2 \Delta t + \nu^2 \Delta t^2, \\ 1 &= p_u + p_m + p_d. \end{aligned}$$

Solving equations gives

$$p_u \;=\; \frac{1}{2}\left(\frac{\sigma^2 \Delta t + \nu^2 \Delta t^2}{\Delta x^2} + \frac{\nu \Delta t}{\Delta x}\right),$$

$$
\begin{aligned}
p_m &= 1 - \frac{\sigma^2 \Delta t + \nu^2 \Delta t^2}{\Delta x^2}, \\
p_u &= \frac{1}{2}\left(\frac{\sigma^2 \Delta t + \nu^2 \Delta t^2}{\Delta x^2} - \frac{\nu \Delta t}{\Delta x}\right).
\end{aligned}
$$

$$
\begin{aligned}
u &= e^{\sigma\sqrt{2\Delta t}} \\
m &= 1 \\
d &= e^{-\sigma\sqrt{2\Delta t}} \\
p_u &= \left(\frac{e^{r\Delta t/2} - e^{\sigma\sqrt{\Delta t/2}}}{e^{\sigma\sqrt{2\Delta t}} - e^{-\sigma\sqrt{\Delta t/2}}}\right)^2 \\
p_m &= 1 - p_u - p_d \\
p_d &= \left(\frac{e^{\sigma\sqrt{\Delta t/2}} - e^{r\Delta t/2}}{e^{\sigma\sqrt{2\Delta t}} - e^{-\sigma\sqrt{\Delta t/2}}}\right)^2
\end{aligned}
$$

(a) CRR trinomial tree

$$
\begin{aligned}
u &= e^{(r-\sigma^2/2)\Delta t + \sigma\sqrt{2\Delta t}} \\
m &= e^{(r-\sigma^2/2)\Delta t} \\
d &= e^{(r-\sigma^2/2)\Delta t - \sigma\sqrt{2\Delta t}} \\
p_u &= 1/4 \\
p_m &= 1/2 \\
p_d &= 1/4
\end{aligned}
$$

(b) JR trinomial tree

$$
\begin{aligned}
u &= e^{(r-\sigma^2/2)\Delta t + \sigma\sqrt{3\Delta t}} \\
m &= e^{(r-\sigma^2/2)\Delta t} \\
d &= e^{(r-\sigma^2/2)\Delta t - \sigma\sqrt{3\Delta t}} \\
p_u &= 1/3 \\
p_m &= 1/3 \\
p_d &= 1/3
\end{aligned}
$$

(c) Equal probability tree

Figure 1.5: Three equivalent methods for building constant-volatility trinomial tree.

Remark 1.6.2. General tree construction:

1. General binomial tree: In a recombining constant volatility binomial tree u and d have the general form

$$
u = e^{\pi\Delta t + \sigma\sqrt{\Delta t}}, \quad d = e^{\pi\Delta t - \sigma\sqrt{\Delta t}},
$$

for any reasonable π.

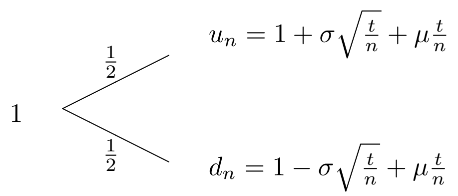
2. General trinomial tree: In a recombining constant volatility binomial tree u, m, and d have the general form

$$
u = e^{\pi\Delta t + \phi\sigma\sqrt{\Delta t}}, \quad m = e^{\pi\Delta t}, \quad d = e^{\pi\Delta t - \phi\sigma\sqrt{\Delta t}},
$$

for $\phi > 1$ and any reasonable π.

1.7 Binomial Tree & Geometric Brownian Motion

1.7.1 Under Physical Measure



$$X_i = \begin{cases} \ \ 1 \\ -1 \end{cases}$$

$$\begin{aligned} U_n &= \#\text{ of up's}, \\ D_n &= \#\text{ of down's.} \quad \text{Then } n = U_n + D_n. \\ W_n &= X_1 + X_2 + \cdots + X_n. \end{aligned}$$

$$\begin{aligned} U_n &= \frac{1}{2}(n + W_n), \\ D_n &= \frac{1}{2}(n - W_n), \\ S_n(t) &= S_0 u_n^{U_n} d_n^{D_n}. \end{aligned}$$

Then we get that

$$\begin{aligned} \log\frac{S_n(t)}{S_0} &= U_n \log u_n + D_n \log d_n \\ &= \frac{1}{2}(n + W_n)\log\Big(1 + \sigma\sqrt{\frac{t}{n}} + \mu\frac{t}{n}\Big) + \frac{1}{2}(n - W_n)\log\Big(1 - \sigma\sqrt{\frac{t}{n}} + \mu\frac{t}{n}\Big) \\ &= \frac{n}{2}\left[\log\Big(1 + \sigma\sqrt{\frac{t}{n}} + \mu\frac{t}{n}\Big) + \log\Big(1 - \sigma\sqrt{\frac{t}{n}} + \mu\frac{t}{n}\Big)\right] \\ &\quad + \frac{W_n}{2}\left[\log\Big(1 + \sigma\sqrt{\frac{t}{n}} + \mu\frac{t}{n}\Big) - \log\Big(1 - \sigma\sqrt{\frac{t}{n}} + \mu\frac{t}{n}\Big)\right] \\ &= I + II. \end{aligned}$$

Note that $\log(1 + x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots$.

$$\begin{aligned} \log\Big(1 + \sigma\sqrt{\frac{t}{n}} + \mu\frac{t}{n}\Big) &= \sigma\sqrt{\frac{t}{n}} + \mu\frac{t}{n} - \frac{1}{2}\sigma^2\frac{t}{n} + O(n^{-\frac{3}{2}}), \\ \log\Big(1 - \sigma\sqrt{\frac{t}{n}} + \mu\frac{t}{n}\Big) &= -\sigma\sqrt{\frac{t}{n}} + \mu\frac{t}{n} - \frac{1}{2}\sigma^2\frac{t}{n} + O(n^{-\frac{3}{2}}). \end{aligned}$$

Hence

$$\begin{aligned} I &= (\mu - \frac{1}{2}\sigma^2)t + O(n^{-\frac{1}{2}}), \\ II &= \frac{W_n}{2}\Big[2\sigma\sqrt{\frac{t}{n}} + O(n^{-\frac{3}{2}})\Big] \end{aligned}$$

$$= \quad W_n \sigma \sqrt{\frac{t}{n}} + W_n O(n^{-\frac{3}{2}}).$$

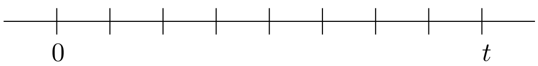
By the central limit theorem,

$$\frac{W_n}{\sqrt{n}} \quad \longrightarrow \quad Z,$$

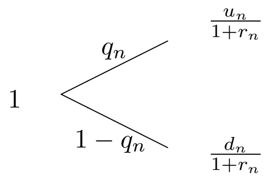
where Z is standard normal distribution. This gives us that

$$\begin{aligned}
\log \frac{S_n(t)}{S_0} \quad &= \quad \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma \sqrt{\frac{t}{n}} W_n + W_n O(n^{-\frac{3}{2}}) \\
&= \quad \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma \sqrt{t} \frac{W_n}{\sqrt{n}} + \frac{W_n}{\sqrt{n}} O(n^{-\frac{1}{2}}) \\
&\longrightarrow \quad \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma \sqrt{t} Z. \\
\therefore \ S_n(t) \quad &= \quad S_0 \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma \sqrt{t} Z\right].
\end{aligned}$$

1.7.2 Under Risk-Neutral Measure



Let $r_n = \frac{t}{n} r$ be interest rate per each interval.



$$\frac{u_n}{1+r_n} q_n + \frac{d_n}{1+r_n}(1 - q_n) = 1$$
$$q_n = \frac{1+r_n-d_n}{u_n-d_n}$$

Put $q_n, 1 - q_n$ be the risk-neutral measure. Let us check $0 < q_n < 1$ for large n.

$$\begin{aligned}
1 + r_n - d_n \quad &= \quad 1 + r\frac{t}{n} - \left(1 - \sigma\sqrt{\frac{t}{n}} + \mu\frac{t}{n}\right) \\
&= \quad \sigma\sqrt{\frac{t}{n}} + (r - \mu)\frac{t}{n} \quad > \quad 0. \\
u_n - (1 + r_n) \quad &= \quad \left(1 + \sigma\sqrt{\frac{t}{n}} + \mu\frac{t}{n}\right) - (1 + r\frac{t}{n}) \\
&= \quad \sigma\sqrt{\frac{t}{n}} + (\mu - r)\frac{t}{n} \quad > \quad 0.
\end{aligned}$$

$$q_n \quad = \quad \frac{1 + r_n - d_n}{u_n - d_n}$$

$$
\begin{aligned}
&= \frac{1 + r\frac{t}{n} - \left(1 - \sigma\sqrt{\frac{t}{n}} + \mu\frac{t}{n}\right)}{2\sigma\sqrt{\frac{t}{n}}} \\
&= \frac{1}{2}\left(1 + \frac{r-\mu}{\sigma}\sqrt{\frac{t}{n}}\right)
\end{aligned}
$$

$$
\begin{aligned}
E_{Q_n}\left[\frac{W_n}{\sqrt{n}}\right] &= \frac{n}{\sqrt{n}} E_{Q_n}[X_1] \\
&= \sqrt{n}\big(q_n + (1-q_n)(-1)\big) \\
&= \sqrt{n}(2q_n - 1) \\
&= \frac{r-\mu}{\sigma}\sqrt{t}. \\
\mathrm{Var}\left[\frac{W_n}{\sqrt{n}}\right] &= n \times \frac{1}{n}\mathrm{Var}[X_1] \\
&= (1^2 q_n + (-1)^2(1-q_n)) - E[X_1]^2 \\
&= 1 - \frac{1}{n}\left(\frac{r-\mu}{\sigma}\sqrt{t}\right)^2 \\
&\to 1, \quad \text{as } n \to \infty.
\end{aligned}
$$

By the central limit theorem,

$$
\frac{W_n}{\sqrt{n}} - \frac{r-\mu}{\sigma}\sqrt{t} \quad \longrightarrow \quad Z.
$$

$$
\begin{aligned}
\log\frac{S_n(t)}{S_0} &= \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma\sqrt{\frac{t}{n}}W_n + W_n O(n^{-\frac{3}{2}}) \\
&= \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma\sqrt{t}\frac{W_n}{\sqrt{n}} + \frac{W_n}{\sqrt{n}}O(n^{-\frac{1}{2}}) \\
&\longrightarrow \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma\sqrt{t}\left(\frac{r-\mu}{\sigma}\sqrt{t} + Z\right) \\
&= \left(r - \frac{1}{2}\sigma^2\right)t + \sigma\sqrt{t}Z.
\end{aligned}
$$

Remark 1.7.1. Another choices for u_n and d_n can be given as follows:

1.
$$
\begin{cases}
u_n = 1 + \sigma\sqrt{\frac{t}{n}} \\
d_n = 1 - \sigma\sqrt{\frac{t}{n}}
\end{cases}
$$

2.
$$
\begin{cases}
u_n = e^{\sigma\sqrt{\frac{t}{n}}} \\
d_n = e^{-\sigma\sqrt{\frac{t}{n}}}
\end{cases}
$$

1.8 The Distribution of GBM

If S_t satisfies the following dynamics

$$dS_t = S_t(\mu dt + \sigma dW_t),$$

then

$$S_t = S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right).$$

It follows from this that $\dfrac{S_t}{S_0}$ is log-normally distributed, and

$$\begin{aligned} E\left[\frac{S_t}{S_0}\right] &= e^{\mu t}, \\ \mathrm{Var}\left[\frac{S_t}{S_0}\right] &= e^{2\mu t}\left(e^{\sigma^2 t} - 1\right). \end{aligned}$$

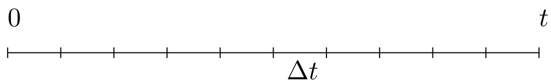
✎ **The probability density function**

$$\begin{aligned} P(S_T \le x) &= \int_0^x \frac{1}{y\sqrt{2\pi(T-t)}\,\sigma} e^{-\frac{\left(\log y - \log S_t - (\mu - \frac{1}{2}\sigma^2)(T-t)\right)^2}{2\sigma^2(T-t)}} dy \\ &= \int_0^x \frac{1}{y\sqrt{2\pi(T-t)}\,\sigma} e^{-\frac{\left(\log \frac{y}{S_t} - (\mu - \frac{1}{2}\sigma^2)(T-t)\right)^2}{2\sigma^2(T-t)}} dy. \end{aligned}$$

By (1.10.1), this probability can be written as

$$P(S_T \le x) = \int_{-\infty}^{\frac{\log \frac{x}{S_t} - (\mu - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz.$$

1.8.1 Why Geometric Brownian Motion?



Let $\Delta t = \frac{t}{n}$. Assume that the return, r, of asset, S, is given by

$$r_n(k) := \log \frac{S_{(k+1)\Delta t}}{S_{k\Delta t}} = \mu\Delta t + \sigma\sqrt{\Delta t}\ \epsilon_{k+1}, \quad \epsilon_k \sim \mathcal{N}(0,1)$$

where $\{\epsilon_i\}_{i\in\mathbb{N}}$ are i.i.d. We employ the central limit theorem and obtain that

$$\sum_{k=0}^{n-1} r_n(k) = \sum_{k=0}^{n-1} \log \frac{S_{(k+1)\Delta t}}{S_{k\Delta t}}$$

$$
\begin{aligned}
&= n\mu\Delta t + \sigma\sqrt{\Delta t}\,(\epsilon_1 + \cdots + \epsilon_n) \\
&= n\mu\Delta t + \sigma\sqrt{n\Delta t}\,\frac{\epsilon_1 + \cdots + \epsilon_n}{\sqrt{n}} \\
&= \mu t + \sigma\sqrt{t}\,\frac{\epsilon_1 + \cdots + \epsilon_n}{\sqrt{n}} \\
&\overset{d}{\to} \mu t + \sigma W_t, \quad \text{as } n \to \infty.
\end{aligned}
$$

Thus we have

$$
\log\frac{S_t}{S_0} = \mu t + \sigma W_t. \tag{1.8.1}
$$

Under the risk-neutral measure we want to be

$$
E\left[\frac{S_t}{S_0}\right] = e^{rt}.
$$

From (1.8.1), we have

$$
E\left[\frac{S_t}{S_0}\right] = e^{\mu t + \frac{1}{2}\sigma^2 t}.
$$

Thus we get

$$
\begin{aligned}
\mu &= r - \frac{1}{2}\sigma^2, \\
\log\frac{S_t}{S_0} &= \left(r - \frac{1}{2}\sigma^2\right)t + \sigma W_t, \\
dS_t &= S_t\,(rdt + \sigma dW_t)\,.
\end{aligned}
$$

1.8.2 Martingale & Dividend

Assume that

$$
d\Big(S_t e^{qt}\Big) = S_t e^{qt}\Big(rdt + \sigma dW_t\Big).
$$

The Ito lemma implies that

$$
\begin{aligned}
dS_t e^{qt} + qS_t e^{qt}dt &= S_t e^{qt}\Big(rdt + \sigma dW_t\Big), \\
dS_t &= S_t\Big((r-q)dt + \sigma dW_t\Big).
\end{aligned}
$$

It is now easily seen that

$$
\left\{\frac{S_t e^{qt}}{e^{rt}}\right\}_{t\geq 0}
$$

is a martingale.

1.8.3 Discretization

The discrete-time version of geometric Brownian motion is

$$\frac{\Delta S}{S} = \mu \Delta t + \sigma \sqrt{\Delta t}\, \epsilon, \quad \epsilon \sim N(0,1).$$

This implies that ΔS have a normal distribution. Thus we have

$$\begin{aligned} E[\Delta S] &= S\mu\Delta t, \\ E[\Delta S^2] &= S^2(\sigma^2\Delta t - \mu^2\,\Delta t^2) \quad \approx \quad S^2\sigma^2\Delta t. \end{aligned}$$

1.9 The Black-Scholes Model

In the early 1970s, Fischer Black and Myron Scholes made a major breakthrough by deriving the PDE and the price for options[5]. The Black-Scholes formalism relies on the following central assumptions:

- Frictionless Markets:
  - The interest rate for borrowing and lending money are equal.
  - All parties have immediate access to any information.
  - All securities and credits are available at any time and in any size.
  - Individual trading will not influence the price.
  - There are no arbitrage opportunities.
- No transaction costs.
- Constant and known volatility(σ) and interest(r) rate.

Let B_t be a risk-less bank account i.e

$$dB_t = rB_t dt, \quad B_0 = 1.$$

Assume that the stock process S_t satisfies

$$dS_t = S_t(\mu dt + \sigma dW_t)$$

where W_t is Brownian motion under the physical measure Q. We can prove that there exists a probability measure $\widetilde{Q}$ which is equivalent to Q, under which

$$\widetilde{W}_t := W_t + \frac{\mu - (r-\delta)}{\sigma} t$$

is a standard Brownian motion[6]. Also, we can deduce that $Z_t := \dfrac{S_t e^{\delta t}}{B_t}$ is a martingale with respect to $\widetilde{Q}$(check!!) and

$$\begin{aligned} dS_t &= S_t(\mu dt + \sigma dW_t) = S_t\Big((r-\delta)dt + \sigma d\widetilde{W}_t\Big), \\ dZ_t &= Z_t \sigma d\widetilde{W}_t. \end{aligned}$$

The measure $\widetilde{Q}$ is referred to as the risk neutral measure or the martingale measure.

Definition 1.9.1.

1. A trading strategy $\Psi = (\#\text{ of bond}, \#\text{ of stock}) = (\psi_t, \phi_t)$ is admissible if it is self-financing and if the discounted value $\widetilde{V}_t(\Psi) = \psi_t + e^{-\delta t}\phi_t Z_t$ of the corresponding portfolio is for all t, non-negative and such that $\sup_{t\in[0,T]} \widetilde{V}_t$ is square integrable under $\widetilde{Q}$.
2. A contingent claim is said to be *replicable* if its payoff at maturity is equal to the final value of an admissible strategy.

□

[5] See Black-Scholes(1973).
[6] Why the risk-neutral measure?

✎ The Black-Scholes Model

For a given contingent claim, $X \in \mathscr{F}_T$, with maturity T, define E_t as follows.

$$E_t \quad := \quad E_{\widetilde{Q}}\left[\frac{X}{B_T}\middle|\mathscr{F}_t\right], \quad 0 \le t \le T.$$

1. By definition E_t is a martingale[7] under $\widetilde{Q}$ and by the martingale representation theorem[8] there exists a predictable process $\widetilde{\phi}_t$ such that

$$dE_t \quad = \quad \widetilde{\phi}_t dZ_t.$$

   Note that since Z_t is not a martingale w.r.t. the physical measure Q(but martingale w.r.t $\widetilde{Q}$) we cannot apply martingale representation theorem.

2. Let $\phi_t := e^{\delta t}\widetilde{\phi}_t$. We want to choose ψ_t(the units of B_t) so that the portfolio (# of B_t,# of S_t) = (ψ_t, ϕ_t) replicates the (undiscounted) contingent claim $V_t := B_t E_t$. If we set

$$\psi_t := E_t - \widetilde{\phi}_t Z_t,$$

   we have

$$V_t \quad = \quad B_t E_t \quad = \quad B_t(\widetilde{\phi}_t Z_t + \psi_t) \quad = \quad \phi_t S_t + \psi_t B_t. \tag{1.9.1}$$

   This means that V_t is the value of replicating portfolio at time t.

3. This trading strategy (ψ_t, ϕ_t) is self-financing because

$$\begin{aligned} dV_t &= d(B_t E_t) \\ &= E_t dB_t + B_t dE_t \quad \text{(by Ito formula)} \\ &= E_t dB_t + B_t \widetilde{\phi}_t dZ_t \\ &= E_t dB_t + B_t \widetilde{\phi}_t \left(\frac{e^{\delta t}}{B_t} dS_t - \frac{e^{\delta t} S_t}{B_t^2} dB_t + \delta \frac{e^{\delta t} S_t}{B_t} dt\right) \\ &= (E_t - \widetilde{\phi}_t Z_t) dB_t + \big(\phi_t dS_t + \delta \phi_t S_t dt\big) \\ &= \psi_t dB_t + \big(\phi_t dS_t + \delta \phi_t S_t dt\big). \end{aligned} \tag{1.9.2}$$

4. Since $V_T (= B_T E_T) = X$, by the no arbitrage condition the value of X at time t(denote X_t) must be equal to V_t. i.e.

$$\begin{aligned} X_t &= V_t \\ &= B_t E_t \\ &= B_t E_{\widetilde{Q}}\left[\frac{X}{B_T}\middle|\mathscr{F}_t\right] \\ &= e^{-r(T-t)} E_{\widetilde{Q}}\left[X \mid \mathscr{F}_t\right]. \end{aligned}$$

   This means that the value of contingent clamin X at time t is the discounted expection with respect to the martingale measure $\widetilde{Q}$.

[7] For $s < t$ we have

$$E_{\widetilde{Q}}\left[E_t|\mathscr{F}_s\right] \quad = \quad E_{\widetilde{Q}}\left[E_{\widetilde{Q}}\left[\frac{X}{B_T}\middle|\mathscr{F}_t\right]\middle|\mathscr{F}_s\right] \quad = \quad E_{\widetilde{Q}}\left[\frac{X}{B_T}\middle|\mathscr{F}_s\right] \quad = \quad E_s.$$

[8] See Oksendal(2003).

✎ **Dynamic Hedging**

In practice, the existence of replicating portfolio (ψ_t, ϕ_t) is not satisfactory and it is essential to be able to build a real replicating portfolio to hedge an option.

There exists a function $f(t, x)$ such that

$$V_t = f(t, S_t).$$

Since

$$\begin{aligned} dV_t &= \phi_t dS_t + \phi_t \delta S_t dt + \psi_t dB_t, \\ df(t, S_t) &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dS_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dS_t)^2 \\ &= \left(\frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 f}{\partial x^2}\right) dt + \frac{\partial f}{\partial x} dS_t, \end{aligned}$$

we have

$$\begin{aligned} \phi_t &= \frac{\partial f}{\partial x}(t, S_t), \\ \frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 f}{\partial x^2} &= \phi_t \delta S_t + r\psi_t B_t \\ &= \phi_t \delta S_t + r(V_t - \phi_t S_t) \quad \text{by (1.9.1)} \\ &= -(r-\delta) S_t \phi_t + r V_t \\ &= -(r-\delta) S_t \frac{\partial f}{\partial x} + r f. \end{aligned}$$

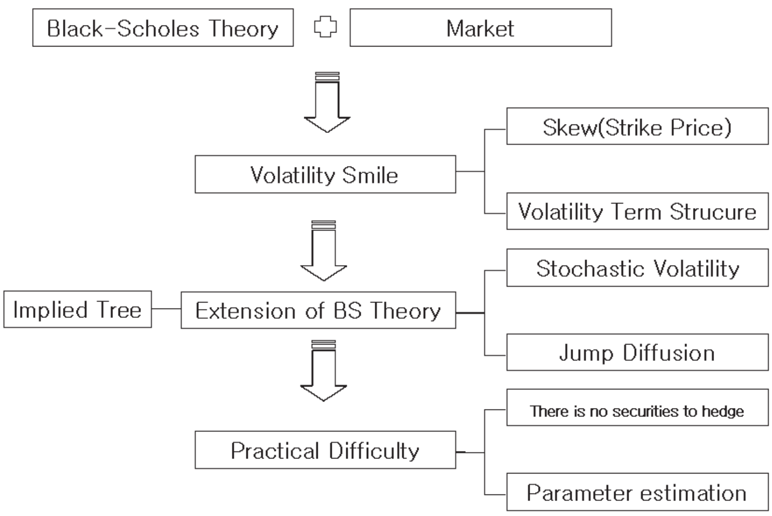


Figure 1.6: An Extension of Black-Scholes Theory

Remark 1.9.2.

1. The main feature of the Black-Scholes model is the dynamic replication of the portfolio and the economic consequences of this argument, rather than as is frequently asserted in the literature, the option pricing formula *per se*. In other words, the essence of the model is risk-neutral replication of securities in a market that is said to be complete.

2. The logic of replication is that a security whose payoff can be replicated purely by the continuous trading of a portfolio of underlying securities is redundant.

$\square$

1.10 Black-Scholes Formula

✎ **Measure change from physical to martingale(risk neutral)**
Assume the pricing dynamics are governed by a geometric Brownian motion.

$$\begin{aligned} dS_t &= S_t(\mu dt + \sigma dW_t), \\ B_t &= e^{rt}, \\ Z_t &= \frac{S_t}{B_t} = e^{-rt} S_t. \end{aligned}$$

By the Ito formula, we have

$$\begin{aligned} dZ_t &= e^{-rt} dS_t - r e^{-rt} S_t dt \\ &= Z_t(\mu dt + \sigma dW_t) - r Z_t dt \\ &= Z_t\big((\mu - r)dt + \sigma dW_t\big) \end{aligned}$$

We want to find a new measure Q under which Z_t is a martingale. Define

$$dQ = \exp\left(\gamma W_t - \frac{1}{2}\gamma^2 t\right) dP,$$

where γ is to be determined later so that Z_t is a martingale. By the Girsanov theorem,

$$d\widetilde{W}_t = dW_t - \gamma dt.$$

is Q-Brownian motion. Now let us determine γ.

$$\begin{aligned} dZ_t &= Z_t\big((\mu - r)dt + \sigma(d\widetilde{W}_t + \gamma dt)\big) \\ &= Z_t\big((\mu - r + \sigma\gamma)dt + \sigma d\widetilde{W}_t\big). \end{aligned}$$

If we set

$$\gamma := \frac{r - \mu}{\sigma},$$

since

$$dZ_t = \sigma Z_t d\widetilde{W}_t,$$

Z_t is a Q-martingale and

$$\begin{aligned} dS_t &= S_t\big(\mu dt + \sigma(d\widetilde{W}_t + \gamma dt)\big) \\ &= S_t\big(r dt + \sigma d\widetilde{W}_t\big), \\ S_t &= S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right) t + \sigma \widetilde{W}_t\right). \end{aligned}$$

1.10.1 Method 1: Tedious algebra

The call option's value at time 0 is given by

$$C(0, S_0) := e^{-rT} \int (S_T - K)^+ dQ$$

$$\begin{aligned}
&= e^{-rT}\int_{\{S_T\geq K\}}(S_T-K)dQ \\
&= e^{-rT}\int_{\{S_T\geq K\}}S_T dQ - e^{-rT}\int_{\{S_T\geq K\}}K dQ \\
&= I - II
\end{aligned}$$

where

$$\begin{aligned}
I &= e^{-rT}\int_{\{S_T\geq K\}}S_T dQ, \\
II &= e^{-rT}\int_{\{S_T\geq K\}}K dQ.
\end{aligned}$$

Straightforward calculations yield

$$\begin{aligned}
II &= e^{-rT}KQ(S_T\geq K) \\
&= e^{-rT}KQ\left(S_0\exp\left((r-\frac{1}{2}\sigma^2)T+\sigma\widetilde{W}_T\right)>K\right) \\
&= e^{-rT}KQ\left((r-\frac{1}{2}\sigma^2)T+\sigma\widetilde{W}_T>\log\frac{K}{S_0}\right) \\
&= e^{-rT}KQ\left(\sigma\widetilde{W}_T>\log\frac{K}{S_0}-(r-\frac{1}{2}\sigma^2)T\right) \\
&= e^{-rT}KQ\left(\sigma\sqrt{T}\ \widetilde{W}_1>\log\frac{K}{S_0}-(r-\frac{1}{2}\sigma^2)T\right) \\
&= e^{-rT}KQ\left(\widetilde{W}_1>\frac{\log\frac{K}{S_0}-(r-\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) \\
&= e^{-rT}KN\left(\frac{\log\frac{S_0}{K}+(r-\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right).
\end{aligned}$$

To evaluate I, define first d_1 and d_2 as follows:

$$\begin{aligned}
d_1 &= \frac{\log\frac{S_0}{K}+\left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}, \\
d_2 &= d_1-\sigma\sqrt{T} \;=\; \frac{\log\frac{S_0}{K}+\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}.
\end{aligned}$$

Let $f(x)$ be the probability density function for the standard normal distribution Z. Note that

$$\begin{aligned}
& S_T\geq K \\
\Longleftrightarrow\ & S_0\exp\left((r-\tfrac{1}{2}\sigma^2)T+\sigma\sqrt{T}Z\right)\geq K \\
\Longleftrightarrow\ & (r-\tfrac{1}{2}\sigma^2)T+\sigma\sqrt{T}Z\geq\log\tfrac{K}{S_0} \qquad (1.10.1)\\
\Longleftrightarrow\ & Z\geq\tfrac{\log\frac{K}{S_0}-(r-\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \\
\Longleftrightarrow\ & Z\geq -d_2.
\end{aligned}$$

Now let us calculate I.

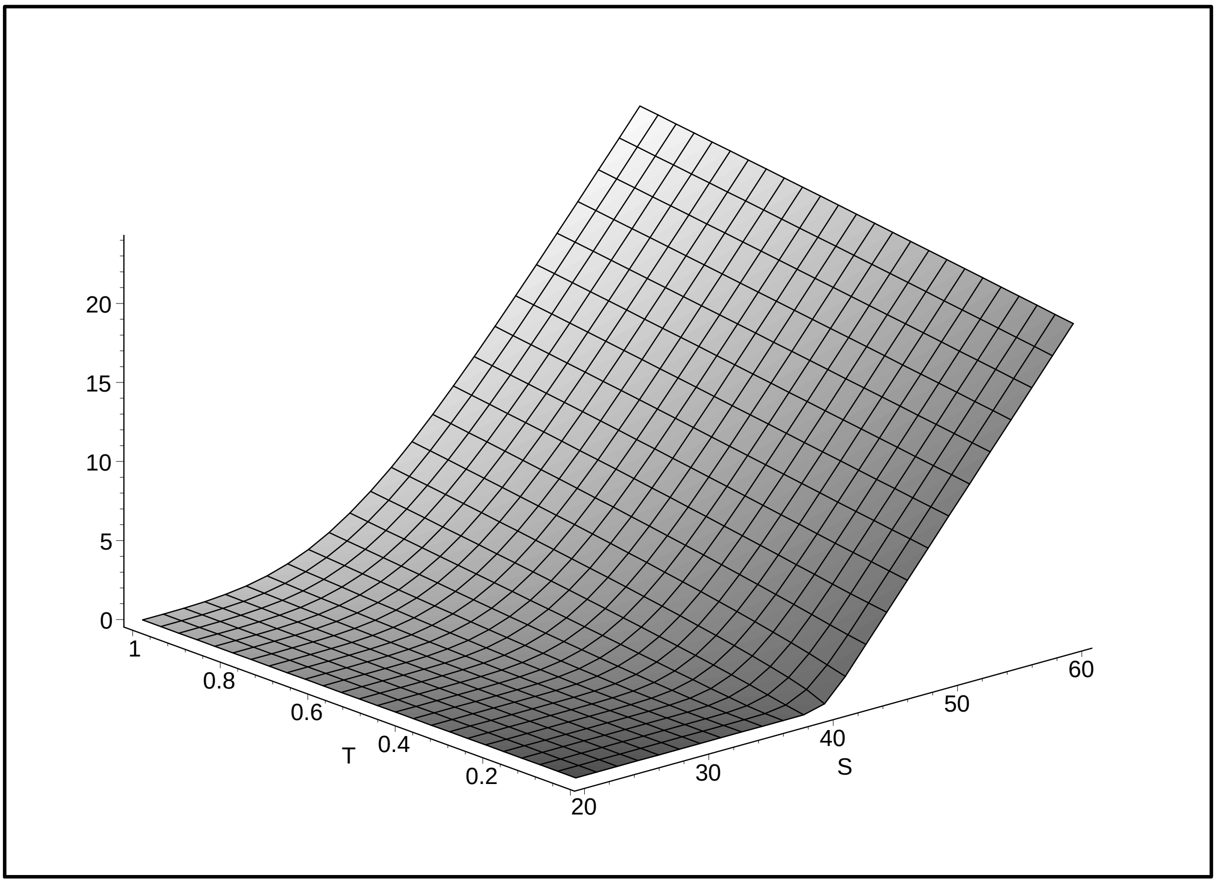


Figure 1.7: Value of a European call option with $r = 0.1, \sigma = 0.2, K = 40, T = 1$

$$\begin{aligned}
I &= e^{-rT} \int_{\{S_T \geq K\}} S_T dQ \\
&= e^{-rT} \int_{\{Z \geq -d_2\}} S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z \right) dQ \\
&= e^{-rT} S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2 \right) T \right) \int_{-d_2}^{\infty} \exp(\sigma\sqrt{T}z) f(z) dz \\
&= \frac{1}{\sqrt{2\pi}} S_0 \exp\left(-\frac{1}{2}\sigma^2 T \right) \int_{-d_2}^{\infty} \exp(\sigma\sqrt{T}z) \exp\left(-\frac{1}{2}z^2 \right) dz \\
&= \frac{1}{\sqrt{2\pi}} S_0 \int_{-d_2}^{\infty} \exp\left(-\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}z - \frac{1}{2}z^2 \right) dz \\
&= \frac{1}{\sqrt{2\pi}} S_0 \int_{-d_2}^{\infty} \exp\left(-\frac{1}{2}(z - \sigma\sqrt{T})^2 \right) dz \\
&= \frac{1}{\sqrt{2\pi}} S_0 \int_{-d_2 - \sigma\sqrt{T}}^{\infty} \exp\left(-\frac{1}{2}w^2 \right) dw \\
&= S_0 \left\{ 1 - N\left(-d_2 - \sigma\sqrt{T} \right) \right\}
\end{aligned}$$

$$
\begin{aligned}
&= S_0\{1 - N(-d_1)\} \\
&= S_0 N(d_1) \\
&= S_0 N\left(\frac{\log\frac{S_0}{K} + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right).
\end{aligned}
$$

Hence we get the Black-Scholes formula as

$$
C(0, S_0) \;=\; S_0 N\left(\frac{\log\frac{S_0}{K} + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) - e^{-rT} K N\left(\frac{\log\frac{S_0}{K} + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right).
$$

Remark 1.10.1. Let $h(x)$ be the PDF of $X := \log(\frac{S_T}{S_0})$. To get the PDF of S_T, let

$$
g(x) = S_0 e^x.
$$

Then $S_T = g(X)$. Hence the PDF of S_T is given by

$$
\begin{aligned}
h(g^{-1}(x))(g^{-1}(x))' &= h\big(\log(\frac{x}{S_0})\big) \cdot \frac{1}{x} \\
&= \frac{1}{\sqrt{2\pi T}\sigma x} \exp\left(-\frac{\left(\log\frac{x}{S_0} - \left(r - \delta - \frac{1}{2}\sigma^2\right)T\right)^2}{2T\sigma^2}\right).
\end{aligned}
$$

1.10.2 Method 2: Measure change(Using Girsanov Theorem)

Recall that

$$
\begin{aligned}
I &= e^{-rT}\int_{\{S_T \geq K\}} S_T dQ \\
&= \int_{\{S_T \geq K\}} Z_T dQ
\end{aligned}
$$

Note that $Z_0 \neq 1$. Define a new measure $\widetilde{Q}$ by

$$
d\widetilde{Q} \;=\; \frac{Z_t}{Z_0} dQ.
$$

Since

$$
Z_t \;=\; Z_0 \exp\left(-\frac{1}{2}\sigma^2 t + \sigma\widetilde{W}_t\right),
$$

by virtue of the Girsanov theorem,

$$
d\widetilde{\widetilde{W}}_t \;=\; d\widetilde{W}_t - \sigma t
$$

is a Brownian motion under $\widetilde{Q}$. Hence we have

$$
\begin{aligned}
I &= Z_0 \widetilde{Q}(S_T > K) \\
&= S_0 \widetilde{Q}(Z_T > K e^{-rT})
\end{aligned}
$$

$$
\begin{aligned}
&= S_0\widetilde{Q}\left(Z_0 \exp\left(-\frac{1}{2}\sigma^2 T + \sigma\widetilde{W}_T\right) > Ke^{-rT}\right) \\
&= S_0\widetilde{Q}\left(\log\frac{S_0}{K} + \left(-\frac{1}{2}\sigma^2 T + \sigma\widetilde{W}_T\right) > -rT\right) \\
&= S_0\widetilde{Q}\left(\log\frac{S_0}{K} + \left(\frac{1}{2}\sigma^2 T + \sigma\widetilde{\widetilde{W}}_T\right) > -rT\right) \\
&= S_0\widetilde{Q}\left(\sigma\widetilde{\widetilde{W}}_T > -\log\frac{S_0}{K} - \frac{1}{2}\sigma^2 T - rT\right) \\
&= S_0\widetilde{Q}\left(\widetilde{\widetilde{W}}_1 > -\frac{\log\frac{S_0}{K} + \frac{1}{2}\sigma^2 T + rT}{\sigma\sqrt{T}}\right) \\
&= S_0 N\left(\frac{\log\frac{S_0}{K} + (\frac{1}{2}\sigma^2 + r)T}{\sigma\sqrt{T}}\right).
\end{aligned}
$$

✎ **Dividend Paying Stock**

When the dividend rate $q \neq 0$, we have the call option price as follows:

$$C(S_0, K, \tau, r, q, \sigma) \quad := \quad S_0 e^{-q\tau} N(d_1) - Ke^{-r\tau} N(d_2).$$

By the put call parity, the put option's value is given by

$$
\begin{aligned}
P(S_0, K, \tau, r, q, \sigma) &= C(S_0, K, \tau, r, q, \sigma) + Ke^{-r\tau} - S_0 e^{-q\tau} \\
&= S_0 e^{-q\tau} N(d_1) - Ke^{-r\tau} N(d_2) + Ke^{-r\tau} - S_0 e^{-q\tau} \\
&= Ke^{-r\tau}\Big(1 - N(d_2)\Big) - S_0 e^{-q\tau}\Big(1 - N(d_1)\Big) \\
&= Ke^{-r\tau} N(-d_2) - S_0 e^{-q\tau} N(-d_1).
\end{aligned}
$$

Also notice that

$$C(S_0, K, \tau, r, q, \sigma) \quad = \quad P(K, S_0, \tau, q, r, \sigma).$$

For example, when $r = q$,

$$
\begin{aligned}
C(S, K) &= P(K, S) \\
&= \frac{K}{S}\, P\left(S, \frac{S^2}{K}\right).
\end{aligned}
$$

This means that the value of call option with a strike price K equals that of $\frac{K}{S}$ put options with a strike $\frac{S^2}{K}$.

1.10.3 Method 3: Alternative Integration

Theorem 1.10.2. The call option price is given by

$$\begin{aligned} C(0,S_0) &= e^{-rT}\int_K^\infty Q\Big(S_T \ge y\Big)\,dy \\ &= e^{-rT}\int_K^\infty N\Big(d_2(S_0,y)\Big)\,dy \\ &= S_0 e^{-qT} N\Big(d_1(S_0,K)\Big) - Ke^{-rT}N\Big(d_2(S_0,K)\Big) \end{aligned}$$

where Q is the standard risk-neutral measure and

$$\begin{aligned} d_1(u,v) &:= \frac{\log\frac{u}{v} + (r-q+\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}, \\ d_2(u,v) &:= \frac{\log\frac{u}{v} + (r-q-\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}. \end{aligned}$$

Proof. Let $f(x)$ be the probability density function for S_T. We have

$$\begin{aligned} & e^{-rT}\int_K^\infty Q\Big(S_T \ge y\Big)\,dy \\ =\ & e^{-rT}\int_K^\infty \int_y^\infty f(x)\ dx\,dy \\ =\ & e^{-rT}\int_K^\infty \int_K^x f(x)\ dy\,dx \\ =\ & e^{-rT}\int_K^\infty (x-K)f(x)dx \\ =\ & e^{-rT}E_Q\Big[(S_T-K)^+\Big] \\ =\ & C(0,S_0). \end{aligned}$$

Now let us calculate the integral:

$$\begin{aligned} & \int_K^\infty Q\Big(S_T \ge y\Big)\,dy \\ =\ & \int_K^\infty N\Big(d_2(S_0,y)\Big)\,dy, \quad Q\Big(S_T\ge y\Big) = Q\Big(W_1 \le d_2(S_0,y)\Big), \\ =\ & \int_K^\infty \int_{-\infty}^{d_2(S_0,y)} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx\,dy \\ =\ & \int_{-\infty}^{d_2(S_0,K)} \int_K^{S_0\exp\left(-\sigma\sqrt{T}x+(r-q-\frac{1}{2}\sigma^2)T\right)} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dy\,dx \\ =\ & \int_{-\infty}^{d_2(S_0,K)} \left\{S_0\exp\left(-\sigma\sqrt{T}x+(r-q-\frac{1}{2}\sigma^2)T\right)-K\right\}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\,dx \\ =\ & \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_2(S_0,K)} S_0\exp\left(-\sigma\sqrt{T}x+(r-q-\frac{1}{2}\sigma^2)T-\frac{1}{2}x^2\right)\,dx - KN\Big(d_2(S_0,K)\Big) \end{aligned}$$

$$
\begin{aligned}
&= \frac{S_0 e^{(r-q)T}}{\sqrt{2\pi}} \int_{-\infty}^{d_2(S_0,K)} \exp\Big(-\sigma\sqrt{T}x - \frac{1}{2}\sigma^2 T - \frac{1}{2}x^2\Big)\, dx - KN\Big(d_2(S_0,K)\Big) \\
&= \frac{S_0 e^{(r-q)T}}{\sqrt{2\pi}} \int_{-\infty}^{d_2(S_0,K)} \exp\Big(-\frac{1}{2}(x+\sigma\sqrt{T})^2\Big)\, dx - KN\Big(d_2(S_0,K)\Big) \\
&= \frac{S_0 e^{(r-q)T}}{\sqrt{2\pi}} \int_{-\infty}^{d_1(S_0,K)} \exp\Big(-\frac{x^2}{2}\Big)\, dx - KN\Big(d_2(S_0,K)\Big) \\
&= S_0 e^{(r-q)T} N\Big(d_1(S_0,K)\Big) - KN\Big(d_2(S_0,K)\Big).
\end{aligned}
$$

■

Theorem 1.10.3. The put option price is given by

$$
\begin{aligned}
P(0,S_0) &= e^{-rT} \int_0^K Q\Big(S_T \le y\Big)\, dy \\
&= e^{-rT} \int_0^K N\Big(-d_2(S_0,y)\Big)\, dy \\
&= Ke^{-rT} N\Big(-d_2(S_0,K)\Big) - S_0 e^{-qT} N\Big(-d_1(S_0,K)\Big).
\end{aligned}
$$

Proof. We follow the same notations in Theorem (1.10.2). By the Fubini theorem, we have

$$
\begin{aligned}
\int_0^K Q\Big(S_T \le y\Big)\, dy &= \int_0^K \int_0^y f(x)\, dxdy \\
&= \int_0^K \int_x^K f(x)\, dydx \\
&= \int_0^K (K-x) f(x) dx \\
&= E\Big[(K-S_T)^+\Big].
\end{aligned}
$$

■

1.10.4 The Generalization of Black-Scholes Formula

Theorem 1.10.4. If V is log-normally distributed and the standard deviation of $\log V$ is s then

$$\begin{aligned} E[(V-X)^+] &= E[V]N(d_1) - XN(d_2), \\ E[(X-V)^+] &= XN(-d_2) - E[V]N(-d_1), \end{aligned}$$

where

$$d_1 = \frac{\log\left(\frac{E[V]}{X}\right) + \frac{s^2}{2}}{s}, \quad d_2 = d_1 - s.$$

Proof. Let m be the mean of $\log V$, where

$$m = \log E[V] - \frac{s^2}{2}.$$

Define a new variable $Z = \frac{\log V - m}{s}$. This variable is normally distributed with a mean zero and a standard deviation of 1.0. Denote the density function of standard normal distribution by $h(z)$

$$\begin{aligned} E[(V-X)^+] &= E[(e^{sZ+m} - X)^+] \\ &= \int_{\frac{\log X - m}{s}}^{\infty} (e^{sz+m} - X)h(z)dz \\ &= \int_{\frac{\log X - m}{s}}^{\infty} e^{sz+m}h(z)dz - X\int_{\frac{\log X - m}{s}}^{\infty} h(z)dz. \end{aligned}$$

Now

$$\begin{aligned} e^{sz+m}h(z) &= \frac{1}{\sqrt{2\pi}} e^{(-z^2+2sz+2m)/2} \\ &= \frac{1}{\sqrt{2\pi}} e^{[-(z-s)^2+2m+s^2]/2} \\ &= \frac{e^{m+s^2/2}}{\sqrt{2\pi}} e^{-(z-s)^2/2} \\ &= e^{m+s^2/2}h(z-s). \end{aligned}$$

If we denote $N(x)$ as the probability that a variable with a mean of zero and a standard deviation 1.0 is less than x, we have that

$$\begin{aligned} E[(V-X)^+] &= e^{m+s^2/2}\left(1 - N\left(\frac{\log X - m}{s} - s\right)\right) - X\left(1 - N\left(\frac{\log X - m}{s}\right)\right) \\ &= e^{m+s^2/2}N\left(-\frac{\log X - m}{s} + s\right) - XN\left(-\frac{\log X - m}{s}\right) \\ &= e^{m+s^2/2}N\left(\frac{\log\left(\frac{E[V]}{X}\right) + \frac{s^2}{2}}{s}\right) - XN\left(\frac{\log\left(\frac{E[V]}{X}\right) - \frac{s^2}{2}}{s}\right) \\ &= E[V]N(d_1) - XN(d_2). \end{aligned}$$

■

Corollary 1.10.5. Black - Scholes call option price is given by

$$\begin{aligned} c &= e^{-rT}E[S_T]N(d_1) - e^{-rT}XN(d_2) \\ &= e^{-rT}S_0e^{rT}N(d_1) - e^{-rT}XN(d_2), \quad E[S_T] = S_0e^{rT} \\ &= S_0N(d_1) - e^{-rT}XN(d_2), \end{aligned}$$

where

$$\begin{aligned} d_1 &= \frac{\log\left(\dfrac{S_0e^{rT}}{X}\right) + \dfrac{\sigma^2 T}{2}}{\sigma\sqrt{T}} \\ &= \frac{\log\left(\dfrac{S_0}{X}\right) + (r+\sigma^2/2)T}{\sigma\sqrt{T}}. \end{aligned}$$

Corollary 1.10.6 (Futures Options). Suppose that we are given the futures call option whose underlying is futures on S_t, exercise price X and options maturity T. The option's price c at time 0 is given by

$$c = e^{-rT}\Big[F_0N(d_1) - XN(d_2)\Big],$$

where F_0 is the futures price at time 0 and

$$d_1 = \frac{\log\left(\dfrac{F_0}{X}\right) + \dfrac{\sigma^2}{2}T}{\sigma\sqrt{T}}, \quad d_2 = d_1 - \sigma\sqrt{T}.$$

Proof. Let F_t be the futures price at time 0. Since F_t is a martingale under risk-neutral measure,

$$\begin{aligned} E[F_T] &= E[F_T|\mathscr{F}_0] = F_0, \\ c &= e^{-rT}\Big(E[F_T]N(d_1) - XN(d_2)\Big) \\ &= e^{-rT}\Big(F_0N(d_1) - XN(d_2)\Big). \end{aligned}$$

■

Example 1.10.7. Replacing S_0 by S_0e^{-qT} in the Black-Scholes formulas, we have the Black-Scholes formulas with continuous dividend rate q.

1.11 The Greek Letters

The Black-Scholes formula says that

$$\begin{aligned} C_t &= S_t e^{-q(T-t)} N(d_1) - X e^{-r(T-t)} N(d_2), \quad P_t = X e^{-r(T-t)} N(-d_2) - S_t e^{-q(T-t)} N(-d_1), \\ \text{where} \quad d_1 &= \frac{\log\left(\frac{S}{X}\right) + (r - q + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 = d_1 - \sigma\sqrt{T-t}, \\ N(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt, \quad N'(x) = n(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \end{aligned}$$

In the following computations, we use the useful equation:

$$n(d_2) = n(d_1)\frac{S}{X} e^{(r-q)\tau}. \tag{1.11.1}$$

Let τ be time to maturity, i.e. $T-t$.

1. Delta:

$$\begin{aligned} \Delta_c &= \frac{\partial C}{\partial S} \\ &= e^{-q(T-t)} N(d_1) + S e^{-q(T-t)} N'(d_1) \frac{\partial d_1}{\partial S} - X e^{-r\tau} N'(d_2) \frac{\partial d_2}{\partial S} \\ &= e^{-q(T-t)} N(d_1) + S e^{-q(T-t)} n(d_1) \frac{1}{\sigma\sqrt{\tau} S} - X e^{-r\tau} n(d_2) \frac{1}{\sigma\sqrt{\tau} S} \\ &= e^{-q(T-t)} N(d_1) + \frac{1}{\sigma\sqrt{\tau}} \left[e^{-q(T-t)} n(d_1) - \frac{X}{S} e^{-r\tau} n(d_2) \right] \\ &= e^{-q(T-t)} N(d_1) + \frac{1}{\sigma\sqrt{\tau}} \left[e^{-q(T-t)} n(d_1) - \frac{X}{S} e^{-r\tau} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} d_1^2 + \sigma\sqrt{\tau} d_1 - \frac{1}{2}\sigma^2 \tau} \right] \\ &= e^{-q(T-t)} N(d_1) + \frac{1}{\sigma\sqrt{\tau}} \left[e^{-q(T-t)} n(d_1) - \frac{X}{S} e^{-r\tau} n(d_1) e^{\sigma\sqrt{\tau} d_1 - \frac{1}{2}\sigma^2 \tau} \right] \\ &= e^{-q(T-t)} N(d_1) + \frac{e^{-q(T-t)}}{\sigma\sqrt{\tau}} \left[n(d_1) - n(d_1) \right] \\ &= e^{-q(T-t)} N(d_1). \end{aligned}$$

By the put-call parity, we find that

$$\begin{aligned} S e^{-q(T-t)} + P &= C + X e^{-r(T-t)}, \\ e^{-q(T-t)} + \frac{\partial P}{\partial S} &= \frac{\partial C}{\partial S}, \end{aligned} \tag{1.11.2}$$

and we have

$$\Delta_p = \frac{\partial P}{\partial S} = e^{-q(T-t)} N(d_1) - e^{-q(T-t)} = -e^{-q(T-t)} N(-d_1).$$

Example 1.11.1. When $q = 0$, the delta of at the money call option is greater than 0.5 because

$$d_1 = \frac{(r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} > 0.$$

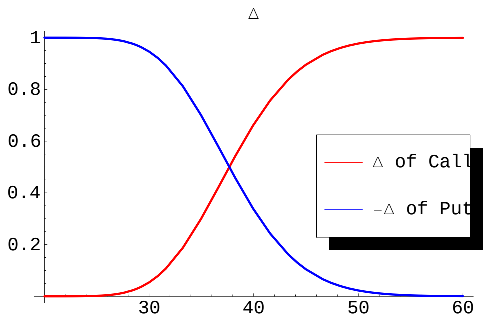


Figure 1.8: Delta of call & put option

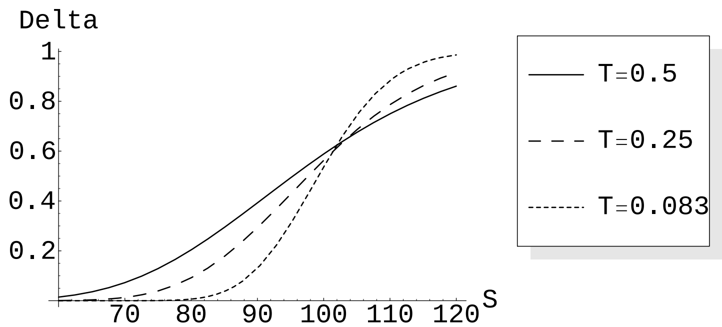


Figure 1.9: Delta w.r.t. underlying price

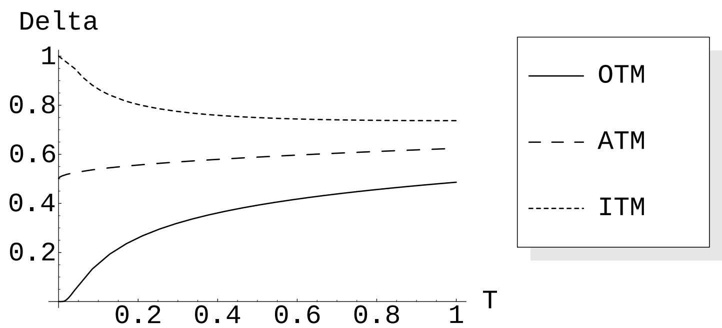


Figure 1.10: Delta w.r.t. time to maturity

2. Gamma: From (1.11.2) we have

$$\begin{aligned}\Gamma_c \;=\; \Gamma_p &= \frac{\partial^2 C}{\partial S^2} \\ &= e^{-q(T-t)} n(d_1)\, \frac{1}{\sigma\sqrt{\tau} S} \;>\; 0.\end{aligned}$$

The gamma is a measure of how much one might have to rehedge, and gives a measure of the amount of transaction costs from delta hedging.

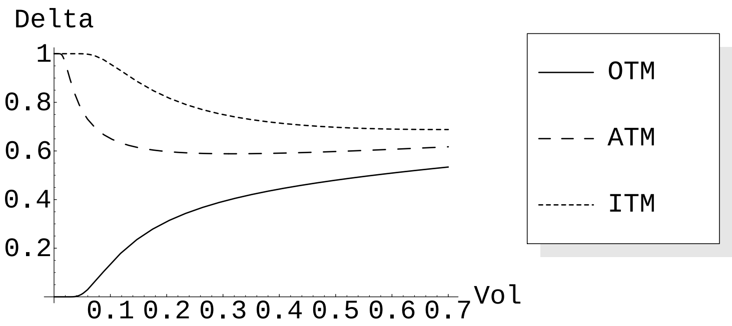
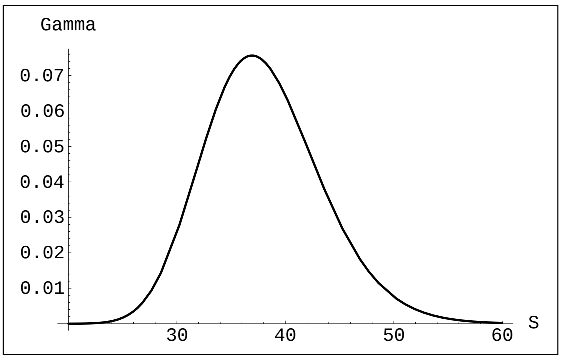
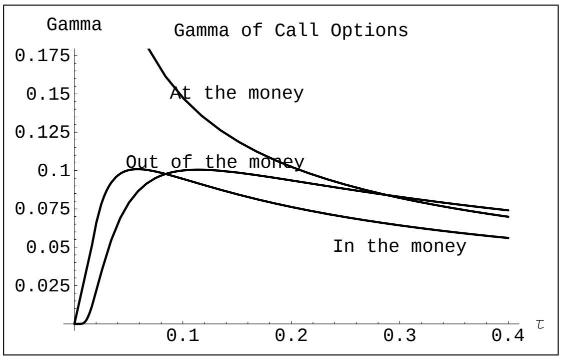


Figure 1.11: Delta w.r.t. volatility



(a) Gamma with stock price ($X = 40, \sigma = 0.2, \tau = 0.5, r = 0.1$)



(b) Gamma with time to maturity

Figure 1.12: Gamma of Call Options

3. Theta:

$$\begin{aligned}
\Theta_c &= \frac{\partial C}{\partial \tau} \\
&= -qSe^{-q\tau}N(d_1) + Se^{-q\tau}n(d_1)\frac{\partial d_1}{\partial \tau} + Xre^{-r\tau}N(d_2) - Xe^{-r\tau}n(d_2)\frac{\partial d_2}{\partial \tau} \\
&= -qSe^{-q\tau}N(d_1) + Se^{-q\tau}n(d_1)\frac{\partial d_1}{\partial \tau} + Xre^{-r\tau}N(d_2) - Se^{-q\tau}n(d_1)\frac{\partial d_2}{\partial \tau}, \quad \text{by (1.11.1)} \\
&= -qSe^{-q\tau}N(d_1) + Se^{-q\tau}n(d_1)\left[\frac{\partial d_1}{\partial \tau} - \frac{\partial d_2}{\partial \tau}\right] + Xre^{-r\tau}N(d_2) \\
&= -qSe^{-q\tau}N(d_1) + \frac{S\sigma}{2\sqrt{\tau}}\,n(d_1) + Xre^{-r\tau}N(d_2).
\end{aligned}$$

Thus we have

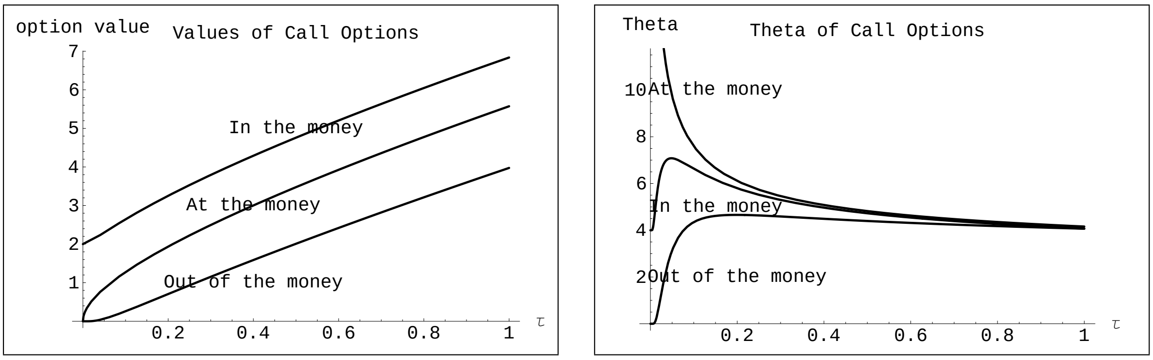
$$\begin{aligned}
\frac{\partial C}{\partial t} &= -\frac{\partial C}{\partial \tau} \\
&= qSe^{-q\tau}N(d_1) - \frac{S\sigma}{2\sqrt{\tau}}\,n(d_1) - Xre^{-r\tau}N(d_2).
\end{aligned}$$

For the put option, by the put-call parity, we have

$$\begin{aligned}
\Theta_p &= \frac{\partial P}{\partial \tau} \\
&= \Theta_c + qSe^{-q\tau} - rXe^{-r\tau} \\
&= -qSe^{-q\tau}\Big(N(d_1) - 1\Big) + \frac{S\sigma}{2\sqrt{\tau}}\, n(d_1) + Xre^{-r\tau}\Big(N(d_2) - 1\Big).
\end{aligned}$$

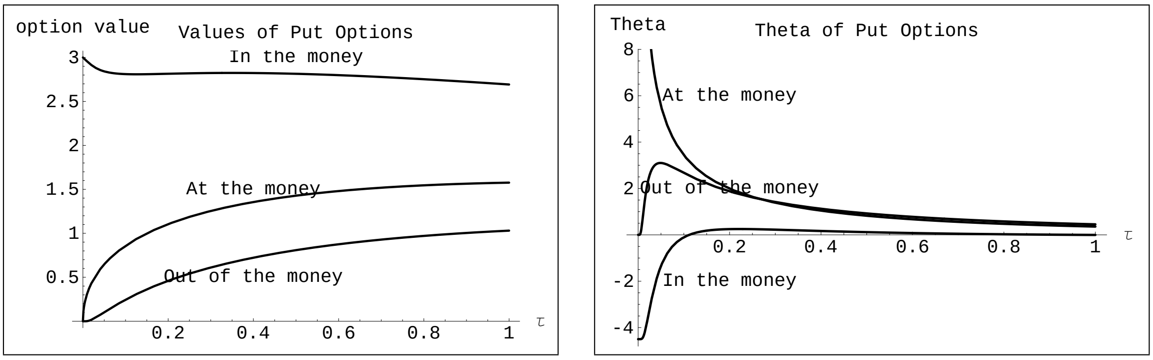
This gives us

$$\begin{aligned}
\frac{\partial P}{\partial t} &= -\frac{\partial P}{\partial \tau} \\
&= -qSe^{-q\tau} N(-d_1) - \frac{S\sigma}{2\sqrt{\tau}}\, n(d_1) + Xre^{-r\tau} N(-d_2).
\end{aligned}$$



(a) Call Option values with Time to Maturity (b) Theta of Call Options with Time

Figure 1.13: Call Value & Theta



(a) Put Option values with Time to Maturity (b) Theta of Put Options with Time

Figure 1.14: Put Value & Theta

4. Vega:

$$\mathscr{V}_c = \frac{\partial C}{\partial \sigma}$$

$$
\begin{aligned}
&= Se^{-q\tau}n(d_1)\frac{\partial d_1}{\partial \sigma} - Xe^{-r\tau}n(d_2)\frac{\partial d_2}{\partial \sigma} \\
&= Se^{-q\tau}n(d_1)\left[\frac{\partial d_1}{\partial \sigma} - \frac{\partial d_2}{\partial \sigma}\right] \quad \text{by (1.11.1)} \\
&= Se^{-q\tau}n(d_1)\sqrt{\tau}.
\end{aligned}
$$

By the put-call parity

$$
\mathscr{V}_c = \mathscr{V}_p = Se^{-q\tau}n(d_1)\sqrt{\tau}.
$$

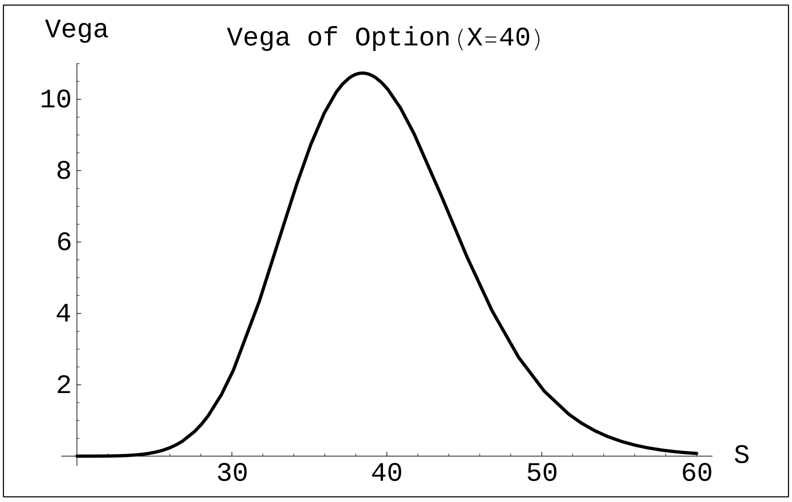


Figure 1.15: Vega of Option

5. Rho:

$$
\begin{aligned}
\rho_c &= \frac{\partial C}{\partial r} \\
&= Se^{-q\tau}n(d_1)\frac{\partial d_1}{\partial r} - Xe^{-r\tau}n(d_2)\frac{\partial d_2}{\partial r} + X\tau e^{-r\tau}N(d_2) \\
&= X\tau e^{-r\tau}N(d_2).
\end{aligned}
$$

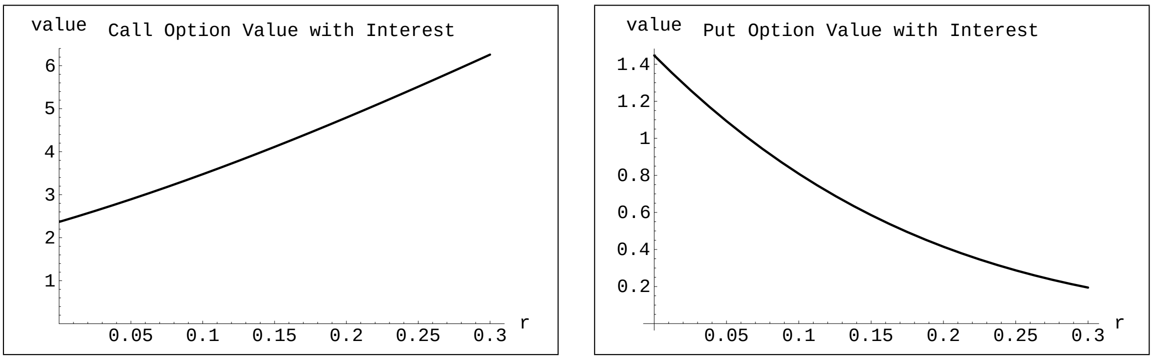
By the put-call parity

$$
\begin{aligned}
\rho_p &= \frac{\partial C}{\partial r} - X\tau e^{-r\tau} \\
&= X\tau e^{-r\tau}\Big(N(d_2) - 1\Big).
\end{aligned}
$$

✎ **Relationship Among Delta, Theta, and Gamma**

The Black-Scholes PDE that must be satisfied by the price, f, of any derivative on a non-dividend-paying stock is

$$
\frac{\partial f}{\partial t} + (r-q)S\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf.
$$



(a) Call Option values with Interest

(b) Put Option values with Interest

Figure 1.16: Option Value with Interest($\tau = 0.5$)

Since

$$\Theta = \frac{\partial f}{\partial t}, \quad \Delta = \frac{\partial f}{\partial S}, \quad \Gamma = \frac{\partial^2 f}{\partial S^2}$$

it follows that

$$\Theta + (r-q)S\Delta + \frac{1}{2}\sigma^2 S^2 \Gamma \quad = \quad rf.$$

Example 1.11.2. The call option value C_t satisfies the following SDE: By Ito's formula,

$$\begin{aligned}
dC_t &= \frac{\partial C_t}{\partial t}dt + \frac{\partial C_t}{\partial S}dS + \frac{1}{2}\frac{\partial^2 C_t}{\partial S^2}(dS)^2 \\
&= \frac{\partial C_t}{\partial t}dt + \frac{\partial C_t}{\partial S}S(rdt + \sigma dW_t) + \frac{1}{2}\frac{\partial^2 C_t}{\partial S^2}S^2\sigma^2 dt \\
&= \Big(\frac{\partial C_t}{\partial t} + rS\frac{\partial C_t}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C_t}{\partial S^2}\Big)dt + \frac{\partial C_t}{\partial S}S\sigma dW_t \\
&= rC_t dt + \Delta S_t \sigma dW_t, \\
&= C_t\left(rdt + \Delta\frac{S_t}{C_t}\sigma dW_t\right).
\end{aligned}$$

Note that since C_t is convex(see Figure(1.17)),

$$\frac{C_t}{S_t} \le \Delta, \quad \text{i.e. } 1 \le \Delta\frac{S_t}{C_t}.$$

Also we can verify that the market prices of risk of S_t and C_t are same. i.e.

$$\begin{aligned}
dS_t &= S_t\Big(\mu_s dt + \sigma_s dW_t\Big), \\
dC_t &= C_t\Big(\mu_c dt + \Delta\frac{S_t}{C_t}\sigma_s dW_t\Big), \\
\Longrightarrow \quad \frac{\mu_s - r}{\sigma_s} &= \frac{\mu_c - r}{\Delta\frac{S_t}{C_t}\sigma_s}.
\end{aligned}$$

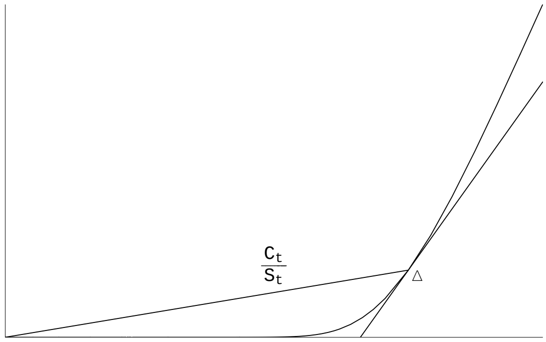


Figure 1.17: Convexity of call option

Thus we have

$$\mu_c - r = \Delta \frac{S_t}{C_t}(\mu_s - r). \tag{1.11.3}$$

In fact, $\mu_s = \mu_c = r$, (1.11.3) is trivial.

1.12 Estimating Volatility vs. Drift

To estimate the volatility of a stock price empirically, the stock price(S_i) is usually observed at fixed intervals of time(Δt). Let $X_1, X_2, \cdots, X_n$ denote a random sample of size $n \geq 2$ from a distribution

$$N\left(\left(\mu - \frac{1}{2}\sigma^2\right)\Delta t, \Delta t\sigma^2\right).$$

Then X_i is the continuously compounded return process(not annualized) in the ith interval. Define

$$\begin{aligned}
\overline{X} &:= \frac{1}{n}\sum_{i=1}^{n} X_i, \\
\overline{\sigma}^2 &:= \frac{1}{\Delta t}\,\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2, \\
\overline{\mu} &:= \frac{1}{\Delta t}\,\overline{X} + \frac{1}{2}\overline{\sigma}^2.
\end{aligned}$$

Note that $\overline{X}$ and $\overline{\sigma}^2$ are independent. The means and variances of the estimators, $\overline{X}, \overline{\sigma}$, and $\overline{\mu}$, are given by

$$\begin{aligned}
E[\overline{X}] &= \left(\mu - \frac{1}{2}\sigma^2\right)\Delta t, \qquad \mathrm{Var}[\overline{X}] = \frac{\Delta t\sigma^2}{n}, \\
E[\overline{\sigma}^2] &= \sigma^2, \qquad \mathrm{Var}[\overline{\sigma}^2] = \frac{2\sigma^4}{n-1}, \\
E[\overline{\mu}] &= \mu, \qquad \mathrm{Var}[\overline{\mu}] = \frac{\sigma^2}{n\Delta t} + \frac{\sigma^4}{2(n-1)}.
\end{aligned} \tag{1.12.1}$$

To obtain (1.12.1), note that[9]

$$\frac{(n-1)}{\Delta t\sigma^2}\,\Delta t\overline{\sigma}^2 \sim \chi(n-1), \quad \text{i.e. } \overline{\sigma}^2 \sim \frac{\sigma^2}{(n-1)}\,\chi(n-1)$$

From this we can see that $\overline{\sigma}^2$ and $\overline{\mu}$ are unbiased estimators of σ^2 and μ, respectively.

Example 1.12.1. Suppose that we are given

$$\mu = 0.1, \quad \sigma = 0.2, \quad n = 1000, \quad \Delta t = \frac{1}{365}.$$

Then

$$\begin{aligned}
\mathrm{std}[\overline{\sigma}^2] &= 0.00179\cdots, \\
\mathrm{std}[\overline{\mu}] &= 0.12083\cdots.
\end{aligned}$$

The standard error of $\overline{\mu}$ is too large!!! □

[9] See pp 214 in Hogg and Craig(1995).

1.13 The Market Price of Risk

Suppose that f_1 and f_2 are the prices of two derivatives. These could be options or other instruments. Suppose that the processes followed by f_1 and f_2 are

$$\begin{aligned} \frac{df_1}{f_1} &= \mu_1 dt + \sigma_1 dW_t, \quad \text{and} \\ \frac{df_2}{f_2} &= \mu_2 dt + \sigma_2 dW_t. \end{aligned}$$

If Π is the value of the portfolio,

$$\Pi = (\sigma_2 f_2) f_1 - (\sigma_1 f_1) f_2.$$

From the self-financing condition, we have that

$$\begin{aligned} d\Pi &= (\sigma_2 f_2) df_1 - (\sigma_1 f_1) df_2 \\ &= f_1 f_2 (\sigma_2 \mu_1 - \sigma_1 \mu_2) dt. \end{aligned} \tag{1.13.1}$$

Because the portfolio is instantaneously riskless, it must earn the risk-free rate. Hence

$$\begin{aligned} d\Pi &= r\Pi dt \\ &= r f_1 f_2 (\sigma_2 - \sigma_1) dt. \end{aligned} \tag{1.13.2}$$

Equations (1.13.1) and (1.13.2) give

$$\sigma_2 \mu_1 - \sigma_1 \mu_2 = r(\sigma_2 - \sigma_1)$$

or

$$\frac{\mu_1 - r}{\sigma_1} = \frac{\mu_2 - r}{\sigma_2}. \tag{1.13.3}$$

Define λ as the value of each side in equation (1.13.3), so that

$$\frac{\mu_1 - r}{\sigma_1} = \frac{\mu_2 - r}{\sigma_2} = \lambda.$$

The parameter λ is referred to as the **market price of risk**. In traditional risk-neutral world, the market price of risk for all tradable assets is zero.

1.14 Market Price of Risk(λ), Measure Change and Numeraire

1.14.1 Single Asset

Suppose we are given the following three assets: money account, security f and g:

$$\begin{cases} \dfrac{dB_t}{B_t} &= rdt, \\ \dfrac{df_t}{f_t} &= \mu_f\, dt + \sigma_f\, dW_t = rdt + \sigma_f\, dW_t^B, \\ \dfrac{dg_t}{g_t} &= \mu_g\, dt + \sigma_g\, dW_t = rdt + \sigma_g\, dW_t^B. \end{cases}$$

1. If we choose the money account, B_t, as numeraire, the volatility of B_t is market price of risk, i.e. $\lambda = 0$.

$$\begin{cases} \dfrac{dB_t}{B_t} &= rdt, \\ \dfrac{df_t}{f_t} &= (r + \lambda\sigma_f)\, dt + \sigma_f\, dW_t^B = r\, dt + \sigma_f\, dW_t^B, \\ \dfrac{dg_t}{g_t} &= (r + \lambda\sigma_g)\, dt + \sigma_g\, dW_t^B = r\, dt + \sigma_g\, dW_t^B. \end{cases}$$

2. If we choose f_t as numeraire, the volatility of f_t is market price of risk, i.e. $\lambda = \sigma_f$.

$$\begin{cases} \dfrac{dB_t}{B_t} &= rdt, \\ \dfrac{df_t}{f_t} &= (r + \lambda\sigma_f)\, dt + \sigma_f\, dW_t^f = (r + \sigma_f^2)\, dt + \sigma_f\, dW_t^f, \\ \dfrac{dg_t}{g_t} &= (r + \lambda\sigma_g)\, dt + \sigma_g\, dW_t^f = (r + \sigma_f\sigma_g)\, dt + \sigma_g\, dW_t^f. \end{cases}$$

3. If we choose g_t as numeraire, the volatility of g_t is market price of risk, i.e. $\lambda = \sigma_g$.

$$\begin{cases} \dfrac{dB_t}{B_t} &= rdt, \\ \dfrac{df_t}{f_t} &= (r + \lambda\sigma_f)\, dt + \sigma_f\, dW_t^g = (r + \sigma_g\sigma_f)\, dt + \sigma_f\, dW_t^g, \\ \dfrac{dg_t}{g_t} &= (r + \lambda\sigma_g)\, dt + \sigma_g\, dW_t^g = (r + \sigma_g^2)\, dt + \sigma_g\, dW_t^g. \end{cases}$$

Choosing a particular market price of risk is referred to as defining the probability measure.

1.14.2 Multi-Assets

Suppose that in the traditional risk-neutral world(i.e. riskless bond is the numeraire), we have

$$\begin{bmatrix} \frac{df_1}{f_1} \\ \frac{df_2}{f_2} \\ \vdots \\ \frac{df_m}{f_m} \end{bmatrix} = \begin{bmatrix} r \\ r \\ \vdots \\ r \end{bmatrix} dt + \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_{mn} \end{bmatrix} \begin{bmatrix} dW_1 \\ dW_2 \\ \vdots \\ dW_n \end{bmatrix}$$

where $(W_1, W_2, \cdots, W_n)$ are independent Wiener processes.

If we choose f_i as a new numeraire, then the new Brownian motion is given by

$$\begin{bmatrix} W_1^i \\ W_2^i \\ \vdots \\ W_n^i \end{bmatrix} := \begin{bmatrix} W_1 - \sigma_{i1}t \\ W_2 - \sigma_{i2}t \\ \vdots \\ W_n - \sigma_{in}t \end{bmatrix}$$

and market price of risk is given by $(\lambda_1^i, \lambda_2^i, \cdots, \lambda_n^i) = (\sigma_{i1}, \sigma_{i2}, \cdots, \sigma_{in})$. When f_i is the numeraire, the securities follow processes of the form

$$\frac{df_j}{f_j} = \mu_j dt + \sum_{k=1}^{n} \sigma_{jk} dW_k^i, \quad j = 1, \cdots, m, \tag{1.14.1}$$

and

$$\mu_j - r = \sum_{k=1}^{n} \lambda_k^i \sigma_{jk} \left(= \sum_{k=1}^{n} \sigma_{ik}\sigma_{jk} \right).$$

Hence Eq(1.14.1) can be rewritten as

$$\frac{df_j}{f_j} = \left(r + \sum_{k=1}^{n} \lambda_k^i \sigma_{jk} \right) dt + \sum_{k=1}^{n} \sigma_{jk} dW_k^i, \quad j = 1, \cdots, m. \tag{1.14.2}$$

Also for any other tradable security g dependent on the n stochastic variables,

$$\frac{dg}{g} = \mu dt + \sum_{k=1}^{n} \sigma_k dW_k^i,$$

we have

$$\mu - r = \sum_{k=1}^{n} \lambda_k^i \sigma_k \left(= \sum_{k=1}^{n} \sigma_{ik}\sigma_k \right).$$

Example 1.14.1. Suppose that in the risk-neutral world we are given

$$\frac{df}{f} = r\,dt + \sigma_f\,dZ_1$$

$$\frac{dg}{g} = r\,dt + \sigma_g\,dZ_2,$$

where Z_1 and Z_2 are Wiener processes with correlation ρ.

We can find independent Wiener processes W_1 and W_2 such that

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \end{bmatrix}. \tag{1.14.3}$$

Then we have

$$\begin{aligned} \frac{df}{f} &= r\,dt + \sigma_f\,dW_1 \\ \frac{dg}{g} &= r\,dt + \rho\sigma_g\,dW_1 + \sqrt{1-\rho^2}\,\sigma_g dW_2. \end{aligned}$$

If we choose f as a new numeraire, then the new Brownian motion, (W_1^f, W_2^f), is given by

$$\begin{bmatrix} W_1^f \\ W_2^f \end{bmatrix} = \begin{bmatrix} W_1 - \sigma_f t \\ W_2 \end{bmatrix}. \tag{1.14.4}$$

This gives us

$$\begin{aligned} \frac{df}{f} &= (r+\sigma_f^2)\,dt + \sigma_f\,dW_1^f &&= (r+\sigma_f^2)\,dt + \sigma_f\,d\widetilde{Z}_1, \\ \frac{dg}{g} &= (r+\rho\sigma_f\sigma_g)\,dt + \rho\sigma_g\,dW_1^f + \sqrt{1-\rho^2}\,\sigma_g dW_2^f &&= (r+\rho\sigma_f\sigma_g)\,dt + \sigma_g d\widetilde{Z}_2 \end{aligned}$$

where

$$\begin{bmatrix} \widetilde{Z}_1 \\ \widetilde{Z}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} W_1^f \\ W_2^f \end{bmatrix}. \tag{1.14.5}$$

We can summarize this result as follows: Substituting from equations (1.14.3) and (1.14.4) into equation (1.14.5), we obtain

$$\begin{aligned} \begin{bmatrix} \widetilde{Z}_1 \\ \widetilde{Z}_2 \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} W_1 - \sigma_f t \\ W_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix} \left\{ \begin{bmatrix} -\sigma_f t \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix}^{-1} \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \right\} \\ &= \begin{bmatrix} -\sigma_f\,t + Z_1 \\ -\rho\sigma_f\,t + Z_2 \end{bmatrix}. \end{aligned}$$

1.15 Change of Numeraire

Let $(\Omega, \mathscr{F})$ be a probability space, $\mathscr{F}_t$ be a filtration and $\widetilde{Q}, Q$ be probability measure defined on $\mathscr{F}$. Assume $\widetilde{Q} \ll Q$. Define

$$\zeta_t = E_Q\left[\frac{d\widetilde{Q}}{dQ} \;\middle|\; \mathscr{F}_t\right].$$

Lemma 1.15.1. For any s, $d\widetilde{Q} = \zeta_s dQ$ on $\mathscr{F}_s$.

Proof. For any $B \in \mathscr{F}_s$,

$$\begin{aligned}\int_B \zeta_s dQ &= \int_B E_Q\left[\frac{d\widetilde{Q}}{dQ} \;\middle|\; \mathscr{F}_s\right] dQ \\ &= \int_B \frac{d\widetilde{Q}}{dQ} dQ \\ &= \int_B d\widetilde{Q}.\end{aligned}$$

■

Theorem 1.15.2.

$$\zeta_s E_{\widetilde{Q}}\left[X_t \;\middle|\; \mathscr{F}_s\right] = E_Q\left[\zeta_t X_t \;\middle|\; \mathscr{F}_s\right], \quad (s < t).$$

Proof. For any $B \in \mathscr{F}_s$,

$$\begin{aligned}\int_B \zeta_s E_{\widetilde{Q}}\left[X_t \;\middle|\; \mathscr{F}_s\right] dQ &= \int_B E_{\widetilde{Q}}\left[X_t \;\middle|\; \mathscr{F}_s\right] d\widetilde{Q}, \quad \text{by Lemma (1.15.1)} \\ &= \int_B X_t d\widetilde{Q} \\ &= \int_B X_t \zeta_t dQ, \quad \text{by Lemma (1.15.1) and } \mathscr{F}_s \subset \mathscr{F}_t \\ &= \int_B E_Q\left[\zeta_t X_t \;\middle|\; \mathscr{F}_s\right] dQ.\end{aligned}$$

Thus we have

$$E_Q\left[\zeta_t X_t \;\middle|\; \mathscr{F}_s\right] = \zeta_s E_{\widetilde{Q}}\left[X_t \;\middle|\; \mathscr{F}_s\right].$$

■

Corollary 1.15.3. If X_t is a $\widetilde{Q}$-martingale, then $\zeta_t X_t$ is Q-martingale.

Proof.

i. $\zeta_t X_t \in \mathscr{F}_t$.

ii.

$$\int |\zeta_t X_t| dQ = \int |X_t| \zeta_t dQ = \int |X_t| d\widetilde{Q} < \infty.$$

iii. By Theorem (1.15.2), for $s < t$

$$\begin{aligned} E_Q\Big[\zeta_t X_t|\mathscr{F}_s\Big] &= \zeta_s E_{\widetilde{Q}}\Big[X_t|\mathscr{F}_s\Big] \\ &= \zeta_s X_s. \end{aligned}$$

■

The inverse of Corollary (1.15.3) holds.

Corollary 1.15.4. If X_t is a Q-martingale, then $\frac{X_t}{\zeta_t}$ is $\widetilde{Q}$-martingale.

Proof.

i. $\frac{X_t}{\zeta_t} \in \mathscr{F}_t$.

ii.

$$\int \left|\frac{X_t}{\zeta_t}\right| d\widetilde{Q} = \int |X_t|\frac{1}{\zeta_t}d\widetilde{Q} = \int |X_t|dQ < \infty.$$

iii. By Theorem (1.15.2), for $s < t$

$$\begin{aligned} E_{\widetilde{Q}}\left[\frac{X_t}{\zeta_t} \;\Big|\; \mathscr{F}_s\right] &= \frac{1}{\zeta_s}\,\zeta_s\, E_{\widetilde{Q}}\left[\frac{X_t}{\zeta_t} \;\Big|\; \mathscr{F}_s\right] \\ &= \frac{1}{\zeta_s}\,E_Q\left[\zeta_t\,\frac{X_t}{\zeta_t} \;\Big|\; \mathscr{F}_s\right] \\ &= \frac{X_s}{\zeta_s}. \end{aligned}$$

■

Definition 1.15.5 (Numeraire). Numeraire is a basic security relative to which the value of other securities can be judged. □

Definition 1.15.6 (Market Price of Risk). A measure of the trade-offs investors make between risk and return. □

Suppose we have a number of securities including some stocks $S_t^1, \cdots, S_t^n$ and two others B_t and C_t either of which might be a numeraire. If we choose B_t to be our numeraire, we need to find a measure Q under which

$$B_t^{-1}S_t^i(i = 1, \cdots, n) \quad \text{and} \quad B_t^{-1}C_t$$

are Q-martingales. Then the value at time t of a derivative payoff X at time T is

$$V_t = B_t E_Q\Big[B_T^{-1}X|\mathscr{F}_t\Big].$$

Suppose however that we choose C_t to be our numeraire instead. Then we would have a different measure $\widetilde{Q}$ under which

$$C_t^{-1}S_t^i(i = 1, \cdots, n) \quad \text{and} \quad B_t^{-1}C_t$$

are $\widetilde{Q}$-martingales. If we apply Corollary (1.15.3), $\zeta_t = \frac{C_t}{B_t}$. Thus the Radon-Nikodym derivative of $\widetilde{Q}$ with respect to Q is the ratio of the numeraire C to the numeraire B,

$$\begin{aligned} \frac{d\widetilde{Q}}{dQ} &= E_Q\left[\frac{d\widetilde{Q}}{dQ}|\mathscr{F}_T\right] = \frac{C_T}{B_T} = \zeta_T, \\ \zeta_t &= \frac{C_t}{B_t} = E_Q\left[\frac{d\widetilde{Q}}{dQ}|\mathscr{F}_t\right] \end{aligned}$$

The price of a payoff X maturing at T under the $\widetilde{Q}$ measure is

$$\begin{aligned} V_t^C &= C_t E_{\widetilde{Q}}\Big[C_T^{-1} X|\mathscr{F}_t\Big] \\ &= C_t \zeta_t^{-1} \zeta_t E_{\widetilde{Q}}\Big[C_T^{-1} X|\mathscr{F}_t\Big] \\ &= C_t \zeta_t^{-1} E_Q\Big[\zeta_T C_T^{-1} X|\mathscr{F}_t\Big] \\ &= B_t E_Q\Big[B_T^{-1} X|\mathscr{F}_t\Big] \\ &= V_t. \end{aligned}$$

Remark 1.15.7 (Alternative). Let X be a T contract. The price of X at time $t = 0$ is given by

$$X_0 = E_Q\left[\frac{X_T}{B_T}\right]. \tag{1.15.1}$$

Assume, on the other hand, that the measure $\widetilde{Q}$ actually exists, with a Random-Nikodym derivative process

$$L_t = \frac{d\widetilde{Q}}{dQ}, \quad \text{on } \mathscr{F}_t.$$

To make $\frac{X_t}{C_t}$ martingale, we have

$$\frac{X_0}{C_0} = E_{\widetilde{Q}}\left[\frac{X_T}{C_T}\right] = E_Q\left[L_T\frac{X_T}{C_T}\right], \quad X_0 = E_Q\left[C_0 L_T \frac{X_T}{C_T}\right]. \tag{1.15.2}$$

Comparing (1.15.1) with (1.15.2), we see that a natural candidate as likelihood process for the intended change of measure is given by

$$L_t = \frac{d\widetilde{Q}}{dQ} = \frac{C_t}{C_0 \cdot B_t} = \frac{C_t/C_0}{B_t/B_0}.$$

□

✎ **Girsanov Transformation**

Suppose we are given

$$dS_0 = \Big(r(t) - q(t)\Big) S_0(t)dt + \sigma(t) S_0(t) dW_t,$$

i.e.

$$d(S_0 e^{q(t)t}) = (S_0 e^{q(t)t})\Big(r(t)dt + \sigma(t)dW_t\Big),$$

We can see that from $L_t := \frac{S_0(t)e^{q(t)t}}{S_0(0)B_t}$ and $L_0 = 1$,

$$\begin{aligned} dL_t &= \frac{1}{S_0(0)B_t} d(S_0 e^{q(t)t}) + \frac{S_0(t)e^{q(t)t}}{S_0(0)} \left(-\frac{1}{B_t^2} \right) dB_t \\ &= \frac{S_0(t)e^{q(t)t}}{S_0(0)B_t} \Big(r(t)dt + \sigma(t)dW_t \Big) + \frac{S_0(t)e^{q(t)t}}{S_0(0)} \Big(-\frac{1}{B_t^2} \Big) r(t)B_t dt \\ &= L_t \sigma(t) dW_t. \end{aligned}$$

Theorem 1.15.8 (Girsanov Transformation). If W_t and is the Brownian motion under old risk neutral measure Q(with B as numeraire), then the new Brownian motion $\widetilde{W}_t$ under $\widetilde{Q}$(with S_0 as numeraire) is given by

$$d\widetilde{W}_t = dW_t - \sigma(t)dt.$$

In fact $\sigma(t)$ is the volatility of the new numeraire S_0.

Proof.

1. Method 1: Since L_t is the Radon-Nikodym derivative and

$$L_t = \exp\Big(\int_0^t \sigma(s)dW_s - \frac{1}{2}\int_0^t \sigma(s)^2 ds \Big),$$

   we can see that by Girsanov Theorem,

$$d\widetilde{W}_t = dW_t - \sigma(t)dt.$$

2. Method 2: Since

$$\begin{aligned} d\left(\frac{B_t}{S_0(t)e^{q(t)t}} \right) &= \frac{r(t)B_t}{S_0(t)e^{q(t)t}} dt - \frac{B_t}{S_0(t)^2 e^{2q(t)t}} d\Big(S_0(t)e^{q(t)t} \Big) + \frac{1}{2} \frac{2B_t}{S_0(t)^3 e^{3q(t)t}} \Big\{ d\Big(S_0(t)e^{2q(t)t} \Big) \Big\}^2 \\ &= -\frac{B_t}{S_0(t)e^{q(t)t}} \sigma(t) \Big(dW_t - \sigma(t)dt \Big), \end{aligned}$$

   to make $\frac{B_t}{S_0(t)e^{q(t)t}}$ be martingale, it should be $d\widetilde{W}_t = dW_t - \sigma(t)dt$.

■

Example 1.15.9 (Black-Scholes Formula). The price at time t of a call option that matures at time T on a stock with strike K is

$$\begin{aligned} C_t &= B_t E\left[\frac{S_T}{B_T} \mathbf{1}_{\{S_T \geq K\}} | \mathscr{F}_t \right] - B_t E\left[\frac{K}{B_T} \mathbf{1}_{\{S_T \geq K\}} | \mathscr{F}_t \right] \\ &= S_t E^S \left[\frac{S_T}{S_T} \mathbf{1}_{\{S_T \geq K\}} | \mathscr{F}_t \right] - B_t \frac{K}{B_T} E\left[\mathbf{1}_{\{S_T \geq K\}} | \mathscr{F}_t \right] \\ &= S_t \mathrm{Prob}^S \Big(S_T \geq K | \mathscr{F}_t \Big) - K e^{-r(T-t)} \mathrm{Prob}\Big(S_T \geq K | \mathscr{F}_t \Big). \end{aligned}$$

Since $dW_t^S = dW_t - \sigma dt$,

$$S_T = S_t \exp\left(\left(r - \frac{1}{2}\sigma^2 \right)(T-t) + \sigma W_{T-t} \right)$$

$$= \quad S_t \exp\left(\left(r + \frac{1}{2}\sigma^2\right)(T-t) + \sigma W^S_{T-t}\right)$$

and

$$\begin{aligned}
\mathrm{Prob}^S\Big(S_T \geq K\Big) &\iff \mathrm{Prob}^S\left(\left(r + \frac{1}{2}\sigma^2\right)(T-t) + \sigma W^S_{T-t} \geq \log\frac{K}{S_t}\right) \\
&\iff \mathrm{Prob}^S\left(W^S_{T-t} \geq \frac{\log\frac{K}{S_t} - (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma}\right) \\
&\iff \mathrm{Prob}^S\left(W^S_1 \geq \frac{\log\frac{K}{S_t} - (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right) \\
&\iff \mathrm{Prob}^S\left(W^S_1 \leq -\frac{\log\frac{K}{S_t} - (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right)
\end{aligned}$$

we have that

$$\begin{aligned}
C_t &= S_t N\left(-\frac{\log\frac{K}{S_t} - \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right) - Ke^{-r(T-t)}N\left(-\frac{\log\frac{K}{S_t} - \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right) \\
&= S_t N\left(\frac{\log\frac{S_t}{K} + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right) - Ke^{-r(T-t)}N\left(\frac{\log\frac{S_t}{K} + \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right).
\end{aligned}$$

□

Example 1.15.10 (Exchange Option). The T-claim to be priced is an exchange option, which gives the holder the right, but not obligation, to exchange on S_0 share for one S_1 share at time T.[10] Formally this means that the claim is given by $X_T = \max[S_1(T) - S_0(T), 0]$, and we note that we have a linearly homogeneous contract function. Assume that

$$\begin{aligned}
dS_0 &= \alpha_0 S_0 dt + \sigma_0 S_0 dW_0, \\
dS_1 &= \alpha_1 S_1 dt + \sigma_1 S_1 dW_1.
\end{aligned}$$

Here W_0 and W_1 are assumed to be Q-Wiener processes with correlation ρ. Under the standard risk neutral measure $\widetilde{Q}$(with B as numeraire) the price dynamics will be given by

$$\begin{aligned}
dS_0 &= (r - q_0)S_0 dt + \sigma_0 S_0 d\widetilde{W}_0, \\
dS_1 &= (r - q_1)S_1 dt + \sigma_1 S_1 d\widetilde{W}_1,
\end{aligned}$$

where $d\widetilde{W}_0 \cdot d\widetilde{W}_1 = \rho dt$. By Example (1.14.1), we can put

$$d\widetilde{\widetilde{W}}_0 = d\widetilde{W}_0 - \sigma_0 dt, \tag{1.15.3}$$

$$d\widetilde{\widetilde{W}}_1 = d\widetilde{W}_1 - \rho\sigma_0 dt. \tag{1.15.4}$$

Then we have $d\widetilde{\widetilde{W}}_0 \cdot d\widetilde{\widetilde{W}}_1 = \rho dt$. By Theorem (1.15.8) these dynamics are converted to

$$dS_0 = \left(r - q_0 + \sigma_0^2\right)S_0 dt + \sigma_0 S_0 d\widetilde{\widetilde{W}}_0,$$

[10] See Margrabe(1978).

$$dS_1 = \big(r - q_1 + \rho\sigma_0\sigma_1\big)S_1 dt + \sigma_1 S_1 d\widetilde{\widetilde{W}}_1,$$

where $\widetilde{\widetilde{W}}_0$ and $\widetilde{\widetilde{W}}_1$ are Brownian motions with correlation ρ under risk neutral measure $\widetilde{\widetilde{Q}}$(with S_0 as numeraire). Now, let $Z_t = \frac{S_1}{S_0} = \frac{S_1(t)}{S_0(t)}$. Then we have

$$\begin{aligned}
dZ_t &= \frac{dS_1}{S_0} - \frac{S_1}{S_0^2}dS_0 - \frac{1}{S_0^2}dS_0 \cdot dS_1 + \frac{1}{2}\,\frac{2S_1}{S_0^3}\,dS_0 \cdot dS_0 \\
&= \frac{S_1}{S_0}\Big((r - q_1 + \rho\sigma_0\sigma_1)dt + \sigma_1 d\widetilde{\widetilde{W}}_1\Big) - \frac{S_1}{S_0}\Big((r - q_0 + \sigma_0^2)dt + \sigma_0 d\widetilde{\widetilde{W}}_0\Big) - \frac{S_1}{S_0}\,\rho\sigma_0\sigma_1 dt + \frac{S_1}{S_0}\sigma_0^2 dt \\
&= \frac{S_1}{S_0}\Big(- (q_1 - q_0)dt + \sigma_1 d\widetilde{\widetilde{W}}_1 - \sigma_0 d\widetilde{\widetilde{W}}_0\Big) \\
&= Z_t\left(-(q_1 - q_0)dt + \sqrt{\sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1}\; dW^*\right), \\
&= Z_t\left(-q dt + \sigma\; dW^*\right)
\end{aligned}$$

where $q = q_1 - q_0, \sigma = \sqrt{\sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1}$ and W^* is a standard $\widetilde{\widetilde{Q}}$-Brownian motion. Since $\left\{\frac{X_t}{S_0(t)e^{q_0 t}}\right\}_t$ is a martingale under measure $\widetilde{\widetilde{Q}}$, we have

$$\begin{aligned}
X_t &= e^{rt} E_{\widetilde{Q}}\left[\frac{X_T}{e^{rT}} \;\Big|\; \mathscr{F}_t\right] \\
&= S_0(t)e^{q_0 t} E_{\widetilde{\widetilde{Q}}}\left[\frac{X_T}{S_0(T)e^{q_0 T}} \;\Big|\; \mathscr{F}_t\right] \\
&= S_0(t)e^{-q_0(T-t)} E_{\widetilde{\widetilde{Q}}}\left[\big(Z_T - 1\big)^+ \;\Big|\; \mathscr{F}_t\right].
\end{aligned}$$

We are thus in fact valuing a European call option on Z_T, with strike price $K = 1$, short rate of interest $r = 0$, dividend rate $q = q_1 - q_0$ and a stock volatility $\sigma = \sqrt{\sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1}$. Therefore

$$\begin{aligned}
X_t &= S_0(t)e^{-q_0(T-t)}\left[Z_t e^{-q(T-t)} N\left(\frac{\log Z_t + \big(-q + \frac{1}{2}\sigma^2\big)(T-t)}{\sigma\sqrt{T-t}}\right) - N\left(\frac{\log Z_t + \big(-q - \frac{1}{2}\sigma^2\big)(T-t)}{\sigma\sqrt{T-t}}\right)\right] \\
&= S_0(t)e^{-q_0(T-t)}\left[Z_t e^{-(q_1-q_0)(T-t)} N\left(\frac{\log Z_t + \big(-(q_1 - q_0) + \frac{1}{2}(\sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1)\big)(T-t)}{\sqrt{(\sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1)(T-t)}}\right)\right. \\
&\qquad\left. - N\left(\frac{\log Z_t \big(-(q_1 - q_0) - \frac{1}{2}(\sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1)\big)(T-t)}{\sqrt{(\sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1)(T-t)}}\right)\right] \\
&= S_1(t)e^{-q_1(T-t)} N\left(\frac{\log\frac{S_1(t)}{S_0(t)} + \big(-(q_1 - q_0) + \frac{1}{2}(\sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1)\big)(T-t)}{\sqrt{(\sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1)(T-t)}}\right) \\
&\qquad - S_0(t)e^{-q_0(T-t)} N\left(\frac{\log\frac{S_1(t)}{S_0(t)} + \big(-(q_1 - q_0) - \frac{1}{2}(\sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1)\big)(T-t)}{\sqrt{(\sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1)(T-t)}}\right).
\end{aligned}$$

Note the price of exchange option is independent of the interest rate r. □

1.16 Series Solution

Sometimes we have boundary conditions at two finite values of S, L and $H(L < S < H)$, for example double barrier option.

Suppose that $V(S,t)$ satisfies

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r-\delta) S \frac{\partial V}{\partial S} - rV = 0.$$

If we write

$$V(S,t) \quad = \quad e^{\alpha x + \beta\tau} U(x,\tau),$$

where

$$\alpha \quad = \quad -\frac{1}{2}\left(\frac{2(r-\delta)}{\sigma^2} - 1\right), \quad \beta \quad = \quad -\frac{1}{4}\left(\frac{2(r-\delta)}{\sigma^2} - 1\right)^2 - \frac{2r}{\sigma^2}, \quad x \quad = \quad \log S, \quad \text{and } \tau = \frac{\sigma^2}{2}(T-t),$$

then $U(x,\tau)$ satisfies the heat equation

$$\frac{\partial U}{\partial \tau} \quad = \quad \frac{\partial^2 U}{\partial x^2}.$$

For the detail, see page 74.

Now write the solution $V(S,t) = U(x,\tau)$ as

$$e^{\alpha x + \beta\tau} \sum_{n=0}^{\infty} \Big[a_n(\tau)\sin(\omega n x) + b_n(\tau)\cos(\omega n x) \Big], \tag{1.16.1}$$

for some ω and some function a and b to be found.

If $a_n(\tau)\sin(\omega n x)$ and $b_n(\tau)\cos(\omega n x)$ are to satisfy the heat equation respectively then we must have

$$\frac{da_n}{d\tau} = -n^2\omega^2 a_n(\tau), \quad \text{and } \frac{db_n}{d\tau} = -n^2\omega^2 b_n(\tau).$$

The solutions are

$$a_n(\tau) = A_n e^{-n^2\omega^2\tau}, \quad \text{and } b_n(\tau) = B_n e^{-n^2\omega^2\tau}.$$

The solution of the Black-Scholes equation is therefore

$$e^{\alpha x + \beta\tau} \sum_{n=0}^{\infty} e^{-n^2\omega^2\tau} \Big[A_n \sin(\omega n x) + B_n \cos(\omega n x) \Big]. \tag{1.16.2}$$

We've got the general solution now. All that we need to do now is to satisfy boundary and initial conditions.

Consider the example where the payoff at time $\tau = 0$ is $f(x)$ but the contract become worthless if ever $x = x_d$ or $x = x_u$. Rewrite (1.16.2) as

$$e^{\alpha x + \beta\tau} \sum_{n=0}^{\infty} e^{-n^2\omega^2\tau} \left[C_n \sin\left(\widetilde{\omega} n \frac{x - x_d}{x_u - x_d}\right) + D_n \cos\left(\widetilde{\omega} n \frac{x - x_d}{x_u - x_d}\right) \right].$$

Choose $D_n = 0$ and $\widetilde{\omega} = \pi$, the boundary conditions are satisfied.

All that remains is to choose the C_n to satisfy the final condition:

$$e^{\alpha x} \sum_{n=0}^{\infty} C_n \sin\left(\pi n \frac{x - x_d}{x_u - x_d}\right) = f(x).$$

Multiplying both sides by

$$\sin\left(\pi m \frac{x - x_d}{x_u - x_d}\right),$$

and integrating between x_d and x_u we find that

$$C_m = \frac{2}{x_u - x_d} \int_{x_d}^{x_u} f(x) e^{-\alpha x} \sin\left(\pi m \frac{x - x_d}{x_u - x_d}\right) dx.$$

Note that

$$\begin{aligned}
&\int_{x_d}^{x_u} \sin\left(\pi m \frac{x - x_d}{x_u - x_d}\right) \sin\left(\pi n \frac{x - x_d}{x_u - x_d}\right) dx \\
&= \frac{x_u - x_d}{\pi} \int_0^{\pi} \sin(my)\sin(ny) dy \\
&= \begin{cases} \frac{x_u - x_d}{2}, & n = m, \\ 0, & n \neq m. \end{cases}
\end{aligned}$$

1.17 Static Hedging

Some exotic options are easier to hedging than the corresponding regular options; others are more difficult. In general, Asian options are easier to hedge because the payoff becomes progressively more certain as we approach maturity. Barrier options can be more difficult to hedge because delta is liable to be discontinuous at the barrier. One approach to hedging an exotic option, known as static options replication [11], is to find a portfolio of regular options whose value matches the value of the exotic option on some boundary, i.e. if two portfolio are worth the same on a certain boundary, they are also worth the same at all interior points of the boundary. The exotic option is hedged by shorting this portfolio. For details, see Hull(2000)

Example 1.17.1. Consider a European down-and-out-call with strike K, lower barrier H. We assume in this particular example that H and K are equal. There are two classes of scenarios for the stock price path: scenario 1 in which the barrier is avoided and the option finishes in-the-money; and scenario 2 in which the barrier is hit before expiration and the option expires worthless.

In scenario 1, the call pays out $S_T - K$, where S_T is the stock price at expiration. This is the same as the payoff of a forward contract with delivery price K. This forward has a theoretical value

$$\begin{aligned} F_t &= e^{-r(T-t)} E[S_T - K \mid \mathscr{F}_t] \\ &= e^{-r(T-t)} \left(S_t e^{(r-q)(T-t)} - K \right) \\ &= S_t e^{-q(T-t)} - K e^{-r(T-t)}, \end{aligned}$$

where q is the continuously paid dividend yield of stock.

For paths in scenario 3, where the stock price hits the barrier at any time t', the call immediately expires with zero value. In that case, the above forward $F_{t'}$ that replicates the barrier-avoiding scenario of type 1 is worth $Ke^{-q(T-t')} - Ke^{-r(T-t')}$.

If the riskless interest rate r equals the dividend yield q, a forward with delivery price K will exactly replicate a down-and-out call with barrier and strike at the same level K, no matter whether the barrier is struck or avoided. □

[11] See Derman, Ergener and Kani(1995).

1.18 Financial Instruments

1. **callable bond**: This is a bond that contains provision allowing the issuing firm to buy back the bond at a predetermined price at certain times in the future. The holder of such a bond has sold a call option to the issuer.

bond + short position of call option

2. **puttable bond**:

bond + put option

3. **swaption(swap option)**: This gives the holder the right to enter into a certain interest rate swap at a certain time in the future. If a swaption gives the holder the right to pay fixed and receive floating, it is a put option on the fixed-rate bond with strike price equal to the principal. If a swaption gives the holder the right to pay floating and receive fixed, ti is a call option on the fixed-rate bond with a strike price equal to the principal.

4. **Range forward, Risk Reversal**: The combination of a long call and short put or the combination of a short call and long put.

   (a) A short-range forward contract consists of a long position in a put with a low strike price, K_1, and a short position in a call with a high strike price, K_2.

   (b) A long-range forward contract consists of a short position in a put with the low strike price, K_1, and a long position in a call with a high strike price K_2.

   The risk reversal is a combination of a long call, with strike above the current spot, and a short put with a strike below the current spot, or vice versa. Both have the same expiry. Also known as a cylinder, a break forward or range forward.

5. **Forward start option**: Forward start option is designed so that it will be at-the-money at some time in the future.

6. **Compound option**: Compound options are options on options.

7. **Corridor option**: An Option to pay 1 if the underlying at maturity is between $S_{\min}$ and $S_{\max}$.

8. **Cliquet option**: A Cliquet Option settles periodically and resets the strike at the then spot level. It is therefore a series of at-the-money options, but where the total premium is determined in advance. A Cliquet can be thought of as a series of "pre-purchased" at-the-money options. The payout on each option can either be paid at the final maturity, or at the end of each reset period.

9. **Parisian option**: Parisian options are essentially a crossover between barrier options and Asian options. They have predominant barrier option features in that they can be knocked in or out depending on hitting a barrier from under or above; they differ from standard barrier options in that extreme outlier asset movements will not trigger the Parisian, and for the trigger to be activated or extinguished, the asset must lie outside or inside the barrier for a predetermined time period t. The effect of this more rigorous triggering criterion is to smooth the option value (and delta and gamma) near the barrier to make hedging somewhat easier.

10. **Reverse Convertible**: Reverse convertibles are bonds which carry a fixed rate of interest but do not guarantee the full redemption of the original investment on maturity. Indeed, if the price of the underlying asset is lower than a given strike price fixed when the bonds are issued, investors will receive a predetermined number of shares in place of the principal amount they subscribed for.

11. **Constant Maturity Swap**: A swap where a swap rate is exchanged for either a fixed rate or a floating rate on each payment date.

12. **Constant Maturity Treasury Swap** A swap where the yield on a Treasury bond is exchanged for either a fixed rate or a floating rate on each payment date.

1.19 Portfolio Fitting

Suppose we are given piecewise-linear function $f(S)$ with broken point $K_i(i = 1, \cdots, N)$.

$$\begin{aligned} \text{Min} \quad & \sum_{i=1}^{N} |a_i| + |b_i| + |c| \\ \text{s.t.} \quad & \sum_{i=1}^{N} \Big[a_i C(s_k, K_i) + b_i P(s_k, K_i) \Big] + c s_k = f(s_k), \quad \text{for all } s_k \in \widehat{S} \end{aligned}$$

where $C(S, K_i)$ and $P(S, K_i)$ are payoff functions of call and put options, respectively and $\widehat{S}$ contains all K_i's and middle points of K_i's.

Chapter 2

SDE & PDE

2.1 Basics

Theorem 2.1.1 (Ito's Formula). Under some conditions, for given processes X_t, Y_t and functions $f(t,x), g(t,x,y)$ we have

$$\begin{aligned} df(t,X_t) &= \frac{\partial f}{\partial t}\,dt + \frac{\partial f}{\partial x}\,dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\,dX_t\cdot dX_t, \\ dg(t,X_t,Y_t) &= \frac{\partial g}{\partial t}\,dt + \frac{\partial g}{\partial x}\,dX_t + \frac{\partial g}{\partial y}\,dY_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}\,dX_t\cdot dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial y^2}\,dY_t\cdot dY_t + \frac{\partial^2 g}{\partial x\partial y}\,dX_t\cdot dY_t. \end{aligned}$$

This Ito formula is sometimes called **Ito's lemma**.

Example 2.1.2. Let X_t, Y_t be Ito processes in $\mathbb{R}$. We have

$$d(X_tY_t) = dX_t\,Y_t + X_t\,dY_t + dX_t\cdot dY_t.$$

□

2.2 Some Examples

Example 2.2.1. Find

$$\int_0^T W_t dW_t.$$

It is quite easy to see

$$\begin{aligned} dW_t^2 &= 2W_tdW_t + dt, \\ \int_0^T W_tdW_t &= \frac{1}{2}W_t^2 - \frac{1}{2}T. \end{aligned}$$

Example 2.2.2 (Geometric Brownian Motion).

$$dS_t = \mu S_t dt + \sigma S_t dW_t.$$

$$
\begin{aligned}
d(\log S_t) &= \frac{1}{S_t}dS_t + \frac{1}{2}\left(-\frac{1}{S_t^2}\right)(dS_t)^2 \\
&= \frac{1}{S_t}dS_t + \frac{1}{2}\left(-\frac{1}{S_t^2}\right)(\sigma S_t dW_t)^2 \\
&= \mu dt + \sigma dW_t - \frac{1}{2}\sigma^2 dt \\
&= \left(\mu - \frac{1}{2}\sigma^2\right)dt + \sigma dW_t, \\
\therefore \quad \log S_t - \log S_0 &= \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t, \\
S_t &= S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right).
\end{aligned}
$$

□

Example 2.2.3. Let $f(x)$ be the PDF of $\log(\frac{S_t}{S_0})$. Then

$$
f(x) = \frac{1}{\sqrt{2\pi t}\sigma}\exp\Big(-\frac{\big(x - \big(\mu - \frac{1}{2}\sigma^2\big)t\big)^2}{2t\sigma^2}\Big).
$$

Also, the PDF of S_t is $g(x) = f\big(\log(\frac{x}{S_0})\big)\cdot\frac{1}{x}$. i.e.

$$
g(x) = \frac{1}{x}\frac{1}{\sqrt{2\pi t}\sigma}\exp\Big(-\frac{\big(\log(\frac{x}{S_0}) - \big(\mu - \frac{1}{2}\sigma^2\big)t\big)^2}{2t\sigma^2}\Big)
$$

□

Example 2.2.4. $f(t,x)$에 대하여, $df(t,X_t)$를 알 때, $dg(f(t,X_t))$=?

$$
\begin{aligned}
dg(f(t,X_t)) &= g'(f(t,X_t))\frac{\partial f}{\partial t}dt + g'(f(t,X_t))\frac{\partial f}{\partial x}dX_t \\
&\quad + \frac{1}{2}\left[g''(f(t,X_t))\left(\frac{\partial f}{\partial x}\right)^2 + g'(f(t,X_t))\frac{\partial^2 f}{\partial x^2}\right](dX_t)^2 \\
&= g'(f(t,X_t))\Big[\frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX_t)^2\Big] \\
&\quad + \frac{1}{2}g''(f(t,X_t))\left(\frac{\partial f}{\partial x}\right)^2(dX_t)^2 \\
&= g'(f(t,X_t))(df) + \frac{1}{2}g''(f(t,X_t))(df)^2
\end{aligned}
$$

□

Example 2.2.5. $f(t,s,x)$에 대하여, $df(t,s,X_t)$를 알 때, $g(t,X_t) := f(t,t,X_t)$에 대한 dg=?

$$
\begin{aligned}
dg(t,X_t) &= \frac{\partial g}{\partial t}dt + \frac{\partial g}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(dX_t)^2 \\
&= \left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial s}\right)dt + \frac{\partial g}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(dX_t)^2
\end{aligned}
$$

$$\begin{aligned} &= \frac{\partial f}{\partial s}dt + \frac{\partial f}{\partial t}dt + \frac{\partial g}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(dX_t)^2 \\ &= \frac{\partial f}{\partial s}dt + df(t,t,X_t). \end{aligned}$$

□

Example 2.2.6 (Ornstein-Uhlenbeck Process). For the $X(t)$, consider a ODE

$$\frac{dX}{dt} = b(a - X), \quad b > 0.$$

The solution of this ODE is given by

$$X(t) = a + c\,e^{-bt}, \qquad c = \text{ constant}.$$

Note that

$$\lim_{t\to\infty} X(t) = a.$$

The Ornstein-Uhlenbeck Process is given by

$$dX_t = \beta(\alpha - X_t)dt + \sigma dW_t.$$

Check that

$$\begin{aligned} d\Big(e^{\beta t}(X_t - \alpha)\Big) &= e^{\beta t}dX_t + \beta e^{\beta t}(X_t - \alpha)dt \\ &= e^{\beta t}\Big(\beta(\alpha - X_t)dt + \sigma dW_t\Big) + \beta e^{\beta t}(X_t - \alpha)dt \\ &= \sigma e^{\beta t}dW_t. \end{aligned}$$

After integration, we have that

$$\begin{aligned} e^{\beta t}(X_t - \alpha) &= (X_0 - \alpha) + \sigma\int_0^t e^{\beta s}dW_s, \\ X_t &= (1 - e^{-\beta t})\alpha + e^{-\beta t}X_0 + \sigma e^{-\beta t}\int_0^t e^{\beta s}dW_s. \end{aligned}$$

Now how to handle $\int_0^t e^{\beta s}dW_s$?

Try 1. Note that $\int_0^t X_s dY_s = X_tY_t - X_0Y_0 - \int_0^t Y_s dX_s - \int_0^t dX_s \cdot dY_s$.

$$\begin{aligned} \int_0^t e^{\beta s}dW_s &= e^{-\beta t}W_t - \int_0^t W_s\beta e^{\beta s}ds \\ &= e^{-\beta t}W_t - \beta\int_0^t e^{\beta s}W_s ds. \end{aligned}$$

Try 2. Let $M_t = \int_0^t e^{\beta s}dW_s$. Then M_t is a martingale. i.e. $dM_t = e^{\beta t}dW_t$.

$$\begin{aligned} d\langle M\rangle_t &= dM_t \cdot dM_t \\ &= e^{2\beta t}dW_t \cdot dW_t \end{aligned}$$

$$
\begin{aligned}
&= e^{2\beta t}dt. \\
\langle M\rangle_t &= \int_0^t e^{2\beta s}ds \\
&= \frac{e^{2\beta t}-1}{2\beta}
\end{aligned}
$$

Therefore

$$M_t \stackrel{d}{=} W_{\langle M\rangle_t} = W_{\frac{e^{2\beta t}-1}{2\beta}} = W\left(\frac{e^{2\beta t}-1}{2\beta}\right).$$

Now we have the following result:

$$
\begin{aligned}
X_t &\stackrel{d}{=} (1-e^{-\beta t})\alpha + e^{-\beta t}X_0 + \sigma e^{-\beta t}\int_0^t e^{\beta s}dW_s \\
&= (1-e^{-\beta t})\alpha + e^{-\beta t}X_0 + \sigma e^{-\beta t}W\left(\frac{e^{2\beta t}-1}{2\beta}\right) \\
&= (1-e^{-\beta t})\alpha + e^{-\beta t}X_0 + W\left(\sigma^2 e^{-2\beta t}\ \frac{e^{2\beta t}-1}{2\beta}\right) \\
&= (1-e^{-\beta t})\alpha + e^{-\beta t}X_0 + W\left(\frac{1-e^{-2\beta t}}{2\beta}\ \sigma^2\right).
\end{aligned}
$$

If $\beta > 0$, as $t \to \infty$

$$X_t \longrightarrow N\left(\alpha, \frac{\sigma^2}{2\beta}\right).$$

If $\beta < 0$,

$$e^{\beta t}X_t = (e^{\beta t}-1)\alpha + X_0 + \sigma W\left(\frac{e^{2\beta t}-1}{2\beta}\right).$$

Hence as $t \to \infty$,

$$e^{\beta t}X_t \longrightarrow -\alpha + X_0 + \sigma W\left(\frac{-1}{2\beta}\right) \quad \text{a.s.}$$

□

Theorem 2.2.7 (The linear SDE). Consider the n-dimensional linear SDE

$$
\begin{cases}
d\mathbf{X}_t = \underbrace{\mathbf{A}}_{n\times n}\mathbf{X}_t + \underbrace{\mathbf{b}_t}_{n\times 1}dt + \underbrace{\boldsymbol{\sigma}_t}_{n\times m}d\underbrace{\mathbf{W}_t}_{m\times 1} \\
\mathbf{X}_0 = \mathbf{x}_0
\end{cases}
$$

The solution of this equation is given by

$$\mathbf{X}_t = e^{\mathbf{A}t}\mathbf{x}_0 + \int_0^t e^{\mathbf{A}(t-s)}\mathbf{b}_s ds + \int_0^t e^{\mathbf{A}(t-s)}\boldsymbol{\sigma}_s d\mathbf{W}_s.$$

Proof. See Arnold(1974) chapter 8. ■

Example 2.2.8. Compute $E[e^{\alpha W_t}]$, where α is a constant.

Try 1. Let $Z_t = e^{\alpha W_t}$. the Ito formula gives us

$$dZ_t = \alpha Z_t dW_t + \frac{1}{2}\alpha^2 Z_t dt, \quad Z_0 = 1.$$

In integral form this reads

$$Z_t = 1 + \frac{1}{2}\alpha^2 \int_0^t Z_s ds + \alpha \int_0^t Z_s dW_s.$$

Let $m(t) = E[Z_t]$. We obtain the equation

$$m(t) = 1 + \frac{1}{2}\alpha^2 \int_0^t m(s) ds.$$

From this we obtain the ODE

$$\begin{cases} m'(t) &= \frac{1}{2}\alpha^2 m(t), \\ m(0) &= 1. \end{cases}$$

Solving this equation gives us the answer

$$E[e^{\alpha W_t}] = E[Z_t] = m(t) = \exp\left(\frac{1}{2}\alpha^2 t\right).$$

Try 2. (Exponential Martingale)

$$\begin{aligned} \exp(\alpha W_t) &= \exp\left(\int_0^t \alpha dW_s\right) \\ &= \exp\left(\int_0^t \alpha dW_s - \frac{1}{2}\int_0^t \alpha^2 ds + \frac{1}{2}\int_0^t \alpha^2 ds\right) \\ &= \exp\left(\int_0^t \alpha dW_s - \frac{1}{2}\int_0^t \alpha^2 ds\right) \exp\left(\frac{1}{2}\int_0^t \alpha^2 ds\right) \\ &= \exp\left(\frac{1}{2}\alpha^2 t\right) \exp\left(\int_0^t \alpha dW_s - \frac{1}{2}\int_0^t \alpha^2 ds\right). \end{aligned}$$

From the exponential martingales,

$$E[e^{\alpha W_t}] = \exp\left(\frac{1}{2}\alpha^2 t\right).$$

□

Lemma 2.2.9. Let $\sigma(t)$ be a given deterministic function of time and define the process X by

$$X_t = \int_0^t \sigma(s) dW_s.$$

Then X_t has a normal distribution with zero mean, and variance given by

$$\mathrm{Var}[X_t] = \int_0^t \sigma^2(s) ds.$$

Proof. We can use the technique described in Example (2.2.8). Let us calculate the characteristic function.

$$\begin{aligned}
E\left[\exp\left(iuX_t\right)\right] &= E\left[\exp\left(\int_0^t iu\sigma(s)dW_s\right)\right] \\
&= E\left[\exp\left(\int_0^t iu\sigma(s)dW_s + \frac{1}{2}\int_0^t u^2\sigma^2(s)ds - \frac{1}{2}\int_0^t u^2\sigma^2(s)ds\right)\right] \\
&= \exp\left(-\frac{1}{2}u^2\int_0^t \sigma^2(s)ds\right)\ E\left[\exp\left(\int_0^t iu\sigma(s)dW_s + \frac{1}{2}\int_0^t u^2\sigma^2(s)ds\right)\right] \\
&= \exp\left(-\frac{1}{2}u^2\int_0^t \sigma^2(s)ds\right).
\end{aligned}$$

This shows that X_t is normally distributed with zero mean and a variance given by

$$\mathrm{Var}[X_t] = \int_0^t \sigma^2(s)ds.$$

■

Alternatively we can obtain that by the Ito isometry

$$\begin{aligned}
\mathrm{Var}[X_t] &= E[X_t^2] \\
&= E\left[\left(\int_0^t \sigma(s)dW_s\right)^2\right] \\
&= E\left[\int_0^t \sigma^2(s)ds\right] \\
&= \int_0^t \sigma^2(s)ds.
\end{aligned}$$

Example 2.2.10 (Exponential Martingale). Show that

$$\exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right)$$

is a martingale.

Proof. For $s < t$,

$$\begin{aligned}
& E\left[\exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right)\ \Big|\ \mathscr{F}_s\right] \\
=\ & E\left[\exp\left(\sigma W_s - \frac{1}{2}\sigma^2 s + \sigma W_t - \sigma W_s - \frac{1}{2}\sigma^2(t-s)\right)\ \Big|\ \mathscr{F}_s\right] \\
=\ & \exp\left(\sigma W_s - \frac{1}{2}\sigma^2 s\right) E\left[\exp\left(\sigma(W_t - W_s) - \frac{1}{2}\sigma^2(t-s)\right)\ \Big|\ \mathscr{F}_s\right].
\end{aligned}$$

Since $W_t - W_s$ is independent of $\mathscr{F}_s$, we have

$$E\left[\exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right)\ \Big|\ \mathscr{F}_s\right]$$

$$
\begin{aligned}
&= \exp\left(\sigma W_s - \frac{1}{2}\sigma^2 s\right) E\left[\exp\left(\sigma(W_t - W_s) - \frac{1}{2}\sigma^2(t-s)\right)\right] \\
&= \exp\left(\sigma W_s - \frac{1}{2}\sigma^2 s\right) \exp\left(\frac{1}{2}\sigma^2(t-s) - \frac{1}{2}\sigma^2(t-s)\right) \\
&= \exp\left(\sigma W_s - \frac{1}{2}\sigma^2 s\right).
\end{aligned}
$$

■

2.3 Black-Scholes PDE

2.3.1 Single Asset

Consider Portfolio $\Pi = f - \Delta S$ (Δ will be determined later). Since Π must have a risk-free return by arbitrage pricing theory,

$$d\Pi - r\Pi dt = 0.$$

By the Ito formula

$$\begin{aligned}
d\Pi - r\Pi dt &= d\Big(f - \Delta S\Big) - r\Big(f - \Delta S\Big)dt \\
&= df - \Delta dS - \delta S\Delta dt - r(f - \Delta S)dt \qquad \text{(by self-financing condition)} \\
&= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial S}dS + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}(dS)^2 - \Delta dS - \delta S\Delta dt - r(f - \Delta S)dt.
\end{aligned}$$

If we take $\Delta = \frac{\partial f}{\partial S}$, then we obtain

$$0 = d\Pi - r\Pi dt = \left(\frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 f}{\partial S^2} + (r-\delta)S\frac{\partial f}{\partial S} - rf\right)dt.$$

Hence we derived Black-Scholes partial differential equation(PDE)as following:

$$\frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 f}{\partial S^2} + (r-\delta)S\frac{\partial f}{\partial S} - rf = 0.$$

✎ **Alternative Approach 1:**
Consider a replicating portfolio (ϕ_t, ψ_t).

$$\begin{aligned}
f &= \phi_t S_t + \psi_t B_t, \\
df &= d(\phi_t S_t + \psi_t B_t) \\
&= \phi_t dS_t + \phi_t\delta S_t dt + \psi_t dB_t, \qquad \text{by self-financing condition,} \\
&= \phi_t\Big((r-\delta)S_t dt + \sigma S_t dW_t\Big) + \phi_t\delta S_t dt + \psi_t r B_t dt \\
&= r(\phi_t S_t + \psi_t B_t)dt + \phi_t\sigma S_t dW_t \\
&= rf dt + \phi_t\sigma S_t dW_t, \\
df &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial S}dS + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}(dS)^2 \\
&= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial S}\Big((r-\delta)S_t dt + \sigma S_t dW_t\Big) + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S_t^2 dt \\
&= \Big(\frac{\partial f}{\partial t} + (r-\delta)S_t\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 f}{\partial S^2}\Big)dt + \sigma S_t\frac{\partial f}{\partial S}dW_t.
\end{aligned}$$

From this we get

$$rf = \frac{\partial f}{\partial t} + (r-\delta)S_t\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 f}{\partial S^2}, \quad \phi_t = \frac{\partial f}{\partial S}.$$

✎ **Alternative Approach 2:**
We will assume that the process followed by S is

$$\frac{dS}{S} = \mu\, dt + \sigma\, dW_t.$$

Suppose that f is the price of derivative dependent only on S and t and suppose that the process followed by f are

$$\frac{df}{f} = m\, dt + s\, dW_t. \tag{2.3.1}$$

Ito's lemma gives us that

$$df = \left(\frac{\partial f}{\partial t} + \mu S\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2}\right) dt + \sigma S \frac{\partial f}{\partial S} dW_t \tag{2.3.2}$$

From (2.3.1) and (2.3.2), we have

$$\begin{aligned} mf &= \frac{\partial f}{\partial t} + \mu S\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2}, \\ sf &= \sigma S \frac{\partial f}{\partial S} \end{aligned}$$

If λ is the market price risk,

$$m - \lambda s = r.$$

This equation gives us the Black-Scholes PDE:

$$\frac{\partial f}{\partial t} + (\mu - \lambda\sigma) S\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf$$

✎ **Alternative Approach 3(Martingale):**
Suppose $e^{-rt}f(S,t)$ is a martingale. Ito's lemma gives

$$\begin{aligned} d\Big(e^{-rt}f(S,t)\Big) &= \frac{\partial f}{\partial t}e^{-rt}dt - rfe^{-rt}dt + e^{-rt}\frac{\partial f}{\partial S}dS + \frac{1}{2}e^{-rt}\frac{\partial^2 f}{\partial S^2}(dS)^2 \\ &= e^{-rt}\left(\frac{\partial f}{\partial t} - rf + (r-\delta)S\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2}\right) dt + (\cdots)dW_t. \end{aligned}$$

Since $e^{-rt}f(S,t)$ is a martingale, dt term should be zero, i.e.

$$\frac{\partial f}{\partial t} - rf + (r-\delta)S\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = 0.$$

2.3.2 Two Assets with dividend

Suppose we have that

$$\begin{aligned} dS_1 &= (r-\delta_1)S_1 dt + \sigma_1 S_1 dz_1, \\ dS_2 &= (r-\delta_2)S_1 dt + \sigma_2 S_2 dz_2 \end{aligned}$$

where the two assets have correlation ρ i.e. $dz_1 \cdot dz_2 = \rho dt$. Let $\Pi = f - \Delta_1 S_1 - \Delta_2 S_2$. Then $d\Pi - r\Pi dt = 0$.

$$
\begin{aligned}
0 &= df - \Delta_1 dS_1 - \Delta_2 dS_2 - \delta_1 \Delta_1 S_1 dt - \delta_2 \Delta_2 S_2 dt - r\Pi dt \\
&= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial S_1}dS_1 + \frac{\partial f}{\partial S_2}dS_2 + \frac{1}{2}\frac{\partial^2 f}{\partial {S_1}^2}dS_1^2 + \frac{1}{2}\frac{\partial^2 f}{\partial {S_2}^2}dS_2^2 + \frac{\partial^2 f}{\partial S_1 \partial S_2}dS_1 \cdot dS_2 \\
&\quad -\Delta_1 dS_1 - \Delta_2 dS_2 - \delta_1 \Delta_1 S_1 dt - \delta_2 \Delta_2 S_2 dt - r\Pi dt \\
&= \frac{\partial f}{\partial t}dt + \frac{1}{2}\frac{\partial^2 f}{\partial {S_1}^2}dS_1^2 + \frac{1}{2}\frac{\partial^2 f}{\partial {S_2}^2}dS_2^2 + \frac{\partial^2 f}{\partial S_1 \partial S_2}dS_1 \cdot dS_2 - \delta_1 \Delta_1 S_1 dt - \delta_2 \Delta_2 S_2 dt - r\Pi dt \\
&= \frac{\partial f}{\partial t}dt + \frac{1}{2}\sigma_1^2 S_1^2 \frac{\partial^2 f}{\partial {S_1}^2}dt + \frac{1}{2}\sigma_2^2 S_2^2 \frac{\partial^2 f}{\partial {S_2}^2}dt + \rho\sigma_1\sigma_2 S_1 S_2 \frac{\partial^2 f}{\partial S_1 \partial S_2}dt \\
&\quad -\delta_1 S_1 \frac{\partial f}{\partial S_1}dt - \delta_2 S_2 \frac{\partial f}{\partial S_2}dt - r(f - \frac{\partial f}{\partial S_1}S_1 - \frac{\partial f}{\partial S_2}S_2)dt \\
&= \frac{\partial f}{\partial t}dt + \frac{1}{2}\sigma_1^2 S_1^2 \frac{\partial^2 f}{\partial {S_1}^2}dt + \frac{1}{2}\sigma_2^2 S_2^2 \frac{\partial^2 f}{\partial {S_2}^2}dt + \rho\sigma_1\sigma_2 S_1 S_2 \frac{\partial^2 f}{\partial S_1 \partial S_2}dt \\
&\quad +(r-\delta_1)S_1 \frac{\partial f}{\partial S_1}dt + (r-\delta_2)S_2 \frac{\partial f}{\partial S_2}dt - rf dt
\end{aligned}
$$

Hence we get

$$
\frac{\partial f}{\partial t} + \frac{1}{2}\sigma_1^2 S_1^2 \frac{\partial^2 f}{\partial {S_1}^2} + \frac{1}{2}\sigma_2^2 S_2^2 \frac{\partial^2 f}{\partial {S_2}^2} + \rho\sigma_1\sigma_2 S_1 S_2 \frac{\partial^2 f}{\partial S_1 \partial S_2} + (r-\delta_1)S_1 \frac{\partial f}{\partial S_1} + (r-\delta_2)S_2 \frac{\partial f}{\partial S_2} - rf = 0.
$$

2.3.3 Boundary Conditions

To uniquely specify a problem we must prescribe boundary conditions an initial or final condition. Boundary conditions tell us how the solution must behave for all times at certain values of the asset. Since the Black-Scholes equation is a backward equation, we therefore have to impose a final condition. This is usually the payoff function at expiry.

The Black-Scholes equation is linear. Linear diffusion equations have some very nice properties.

1. Even if we start out with a discontinuity in the final data, due to a discontinuity in the payoff, this immediately gets smoothed out, due to the diffusion nature of the equation.
2. Another nice property is the uniqueness of the solution. provided the solution is not allowed to grow too fast as S tends to infinity the solution will be unique. The precise definition of 'too fast' need not worry us, we will not have to worry about uniqueness for any problems we encounter.

For details, refer to Wilmott(2000).

Example 2.3.1. For a call option, the boundary conditions are given by

$$
\begin{aligned}
f(0,t) &= 0, \\
f(S_t,t) &\sim S_t \quad \text{as } S \to \infty, \quad \text{more accurate expression: } S_t - Ke^{-r(T-t)}
\end{aligned}
$$

This relation can be obtained from the put-call parity. i.e.

$$
S_t + P_t = C_t + Ke^{-r(T-t)},
$$

and for a large S_t, we have $P_t \approx 0$

$$S_t = C_t + Ke^{-r(T-t)}.$$

$\square$

2.4 BS PDE & Heat Equation

Heat Equation:

$$\begin{aligned} & \frac{\partial u}{\partial t} - k\frac{\partial^2 u}{\partial x^2} = 0 && \frac{\partial v}{\partial t} + k\frac{\partial^2 v}{\partial x^2} = 0, \quad k > 0 \\ \text{i.e.} \quad & (\partial_t - k\Delta)u = 0 && (\partial_t + k\Delta)v = 0, \\ & u(0,x) = f(x) && v(T,x) = g(x). \end{aligned}$$

Let us define heat kernel:

$$K_t(x) = \frac{1}{\sqrt{4\pi t}} \exp\left[-\frac{x^2}{4t}\right].$$

$K(t,x)$ is the PDF of the normal distribution with mean 0, variance $2t$. The solution of heat equation:

$$\begin{aligned} u(t,x) &= f * K_{kt} \\ &= \int_{-\infty}^{\infty} f(y) K_{kt}(x-y)\, dy \\ &= \frac{1}{\sqrt{4k\pi t}} \int_{-\infty}^{\infty} f(y) \exp\left[-\frac{(x-y)^2}{4kt}\right] dy, \\ v(0,x) &= g * K_{kT} \\ &= \frac{1}{\sqrt{4k\pi T}} \int_{-\infty}^{\infty} g(y) \exp\left[-\frac{(x-y)^2}{4kT}\right] dy. \end{aligned}$$

Note that $*$ denotes the convolution operator.

Remark 2.4.1. We can solve the initial value problem for the heat equation in $\mathbb{R}^n$,

$$\frac{\partial u}{\partial t} = k\nabla u, \quad u(\mathbf{x},0) = f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n.$$

The result is

$$u(\mathbf{x},t) = f * \mathbf{K}_t(\mathbf{x}), \quad \mathbf{K}_t(\mathbf{x}) := (4\pi k t)^{-n/2} e^{-|\mathbf{x}|^2/4kt}.$$

Now, consider the Black-Scholes PDE.

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC = 0.$$

2.4.1 Method 1: $u_s + \frac{1}{2}u_{yy} = 0$.

Let

$$z = \log\frac{S}{K}, \quad \text{i.e. } S = Ke^z.$$

$$\frac{\partial C}{\partial z} = \frac{\partial C}{\partial S} \cdot \frac{\partial S}{\partial z} = S \cdot \frac{\partial C}{\partial S}$$

$$\frac{\partial^2 C}{\partial z^2} = \frac{\partial}{\partial S}\big(S\cdot\frac{\partial C}{\partial S}\big)\cdot\frac{\partial S}{\partial z} = \Big(\frac{\partial C}{\partial S} + S\frac{\partial^2 C}{\partial S^2}\Big)S$$

Then we get

$$\frac{\partial C}{\partial t} + r\frac{\partial C}{\partial z} + \frac{1}{2}\sigma^2\left(\frac{\partial^2 C}{\partial z^2} - \frac{\partial C}{\partial z}\right) - rC = 0.$$

i.e.

$$\frac{\partial C}{\partial t} + (r - \frac{1}{2}\sigma^2)\frac{\partial C}{\partial z} + \frac{1}{2}\sigma^2\left(\frac{\partial^2 C}{\partial z^2}\right) - rC = 0.$$

Change variables: $(t, z) \to (s, y)$

$$\begin{cases} y & := & Az + Bt \\ s & := & t \end{cases}$$

$$\begin{aligned}
\frac{\partial C}{\partial z} &= \frac{\partial C}{\partial y}\frac{\partial y}{\partial z} + \frac{\partial C}{\partial s}\frac{\partial s}{\partial z} = A\frac{\partial C}{\partial y}, \quad \frac{\partial s}{\partial z} = 0,\\
\frac{\partial^2 C}{\partial z^2} &= \frac{\partial}{\partial y}(A\frac{\partial C}{\partial z})\cdot\frac{\partial y}{\partial z} = A^2\frac{\partial^2 C}{\partial y^2},\\
\frac{\partial C}{\partial t} &= \frac{\partial C}{\partial s}\frac{\partial s}{\partial t} + \frac{\partial C}{\partial y}\frac{\partial y}{\partial t} = \frac{\partial C}{\partial s} + B\frac{\partial C}{\partial y}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial C}{\partial s} + B\frac{\partial C}{\partial y} + (r - \frac{1}{2}\sigma^2)A\frac{\partial C}{\partial z} + \frac{1}{2}\sigma^2A^2\frac{\partial^2 C}{\partial y^2} - rC = 0.\\
\frac{\partial C}{\partial s} + \Big(B + (r - \frac{1}{2}\sigma^2)A\Big)\frac{\partial C}{\partial y} + \frac{1}{2}\sigma^2A^2\frac{\partial^2 C}{\partial y^2} - rC = 0.
\end{aligned}$$

Set

$$\begin{cases} A & = & \frac{1}{\sigma} \\ B & = & -\frac{1}{\sigma}\left(r - \frac{1}{2}\sigma^2\right). \end{cases}$$

Then we have

$$\frac{\partial C}{\partial s} + \frac{1}{2}\frac{\partial^2 C}{\partial y^2} - rC = 0.$$

$$S = Ke^z = K\exp\left(\left(r - \frac{1}{2}\sigma^2\right)s + \sigma y\right)$$

Define

$$u(s, y) := e^{-rs}C(s, y).$$

Then we have

$$\frac{\partial u}{\partial s} = -re^{-rs}C + e^{-rs}\frac{\partial C}{\partial s}$$

Hence we get

$$\frac{\partial u}{\partial s} + \frac{1}{2}\frac{\partial^2 u}{\partial y^2} = 0.$$

Hence the BS-equation for call option

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC = 0, \quad C(T,S) = (S-K)^+.$$

is converted to

$$\frac{\partial u}{\partial s} + \frac{1}{2}\frac{\partial^2 u}{\partial y^2} = 0, \quad u(T,y) = e^{-rT}\left(Ke^{(r-\frac{1}{2}\sigma^2)T+\sigma y} - K\right)^+$$

Now let us compute $C(0,S_0) = u(0,x)$.

$$\begin{aligned}
u(0,x) &= \frac{1}{\sqrt{2\pi T}}\int_{-\infty}^{\infty} u(T,y)e^{-\frac{(y-x)^2}{2T}}dy \\
&= \frac{1}{\sqrt{2\pi T}}\int_{-\infty}^{\infty} e^{-rT}(S_T-K)^+ e^{-\frac{(y-x)^2}{2T}}dy \\
&= \frac{1}{\sqrt{2\pi T}}\int_{\{y:y\geq -\frac{1}{\sigma}(r-\frac{1}{2}\sigma^2)T\}} e^{-rT}(S_T-K)e^{-\frac{(y-x)^2}{2T}}dy, \\
&= I_1 - I_2,
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \frac{1}{\sqrt{2\pi T}}\int_{\{y:y\geq -\frac{1}{\sigma}(r-\frac{1}{2}\sigma^2)T\}} e^{-rT}Ke^{(r-\frac{1}{2}\sigma^2)T+\sigma y}e^{-\frac{(y-x)^2}{2T}}dy, \\
I_2 &= \frac{1}{\sqrt{2\pi T}}\int_{\{y:y\geq -\frac{1}{\sigma}(r-\frac{1}{2}\sigma^2)T\}} e^{-rT}Ke^{-\frac{(y-x)^2}{2T}}dy.
\end{aligned}$$

Straightforward calculation leads to the following result

$$\begin{aligned}
I_1 &= \frac{1}{\sqrt{2\pi T}}e^{-rT}Ke^{(r-\frac{1}{2}\sigma^2)T}\int_{\{y:y\geq -\frac{1}{\sigma}(r-\frac{1}{2}\sigma^2)T\}} e^{\sigma y}e^{-\frac{(y-x)^2}{2T}}dy \\
&= \frac{1}{\sqrt{2\pi T}}Ke^{-\frac{1}{2}\sigma^2 T}\int_{-\frac{1}{\sigma}(r-\frac{1}{2}\sigma^2)T}^{\infty} e^{-\frac{(y-x-\sigma T)^2}{2T}+\sigma x+\frac{1}{2}\sigma^2 T}dy \\
&= \frac{1}{\sqrt{2\pi T}}Ke^{\sigma x}\int_{-\frac{1}{\sigma}(r-\frac{1}{2}\sigma^2)T}^{\infty} e^{-\frac{(y-x-\sigma T)^2}{2T}}dy \\
&= \frac{1}{\sqrt{2\pi}}S_0\int_{-\frac{(r+\frac{1}{2}\sigma^2)T+\sigma x}{\sigma\sqrt{T}}}^{\infty} e^{-\frac{z^2}{2}}dz, \qquad S_0 = Ke^{(r-\frac{1}{2}\sigma^2)\cdot 0+\sigma x} = Ke^{\sigma x} \\
&= S_0 N\left(\frac{(r+\frac{1}{2}\sigma^2)T+\ln(S/K)}{\sigma\sqrt{T}}\right)
\end{aligned}$$

and

$$\begin{aligned}
I_2 &= Ke^{-rT}\frac{1}{\sqrt{2\pi T}}\int_{-\frac{1}{\sigma}(r-\frac{1}{2}\sigma^2)T}^{\infty} e^{-\frac{(y-x)^2}{2T}}dy \\
&= Ke^{-rT}\frac{1}{\sqrt{2\pi}}\int_{-\frac{(r-\frac{1}{2}\sigma^2)T+\sigma x}{\sigma\sqrt{T}}}^{\infty} e^{-\frac{z^2}{2}}dz \\
&= Ke^{-rT}N\left(\frac{(r-\frac{1}{2}\sigma^2)T+\ln(S/K)}{\sigma\sqrt{T}}\right)
\end{aligned}$$

2.4.2 Method 2(Time Dependent): $u_s = u_{yy}$.

Let

$$\begin{aligned} u(s,y) &:= e^{r(T-t)}C(t,S), \quad \text{where} \\ s &= T-t, \\ S &= Ke^{(r-\frac{1}{2}\sigma^2)(T-s)+\frac{\sigma}{\sqrt{2}}y}, \quad i.e.\ y = \frac{\sqrt{2}}{\sigma}\ln(S/K) + \left(\frac{\sigma}{\sqrt{2}} - \frac{\sqrt{2}}{\sigma}r\right)t \end{aligned}$$

Then $u(s,y)$ satisfies the heat equation:

$$u_s = u_{yy}$$

1. $s = T$: Then $t = 0$.

$$u(T,y) = e^{rT}C(0,S), \quad S = K\exp\left(\frac{\sigma}{\sqrt{2}}y\right)$$

2. $s = 0$: Then $t = T$.

$$u(0,y) = C(T,S), \quad S = K\exp\left\{\left(r - \frac{1}{2}\sigma^2\right)T + \frac{\sigma}{\sqrt{2}}y\right\}$$

Remark 2.4.2. 위와 같이 변수 변환 후, 시간에 관계없이 y의 범위를 고정하고 FDM을 적용한다. 그러나, S의 범위는 t에 따라 그 범위가 달라진다. 그러나 그 변화의 정도는 그렇게 크지 않을 수 있다. 예를 들어, $t = 0$인 경우와 $t = T$인 경우를 비교해 보면,

$$\exp\left\{\left(r - \frac{1}{2}\sigma^2\right)T\right\}$$

만큼 차이가 난다. $r = 0.1, \sigma = 0.2, T = 1$인 경우, $\exp\{(r - \frac{1}{2}\sigma^2)T\} = 1.0832$ 배 차이가 난다.
반면, subsection (2.4.3)의 방식은 $t = T - \frac{2s}{\sigma^2}$ 로 시간을 변환하여, 시간의 흐름은 왜곡되지만, S의 범위는 시간에 관계없이 변하지 않는다.

Remark 2.4.3 (Boundary Conditions). call option의 경우, deep out of the money인 경우 그 값이 거의 0에 가깝다. 즉,

$$g_1(t,S_t) \approx 0 \approx (S_t - Ke^{-r(T-t)})^+, \quad S_t \ll K.$$

그리고, deep in the money인 경우, option value는 $(S_T - K)$를 할인한 것과 거의 같다. 따라서 $0 \le t \le T$에서 deep in the money call option value는

$$g_2(t,S_t) \approx e^{-r(T-t)}(S_T - K) \approx S_t - Ke^{-r(T-t)} \approx (S_t - Ke^{-r(T-t)})^+, K \ll S_t.$$

가 적절하다. 또한 이 식은 put-call parity로 부터 얻어낼 수도 있다.
$u(s,y) = e^{r(T-t)}g_i(t,S), (i = 1,2)$이므로

$$\begin{aligned} u(s,y_0) &= e^{rs}g_1(T-s,S_0) \\ &= e^{rs}(S_0 - Ke^{-rs})^+, \quad S_0\text{: } s\text{에 따라 값이 달라진다.} \\ &= e^{rs}\left(Ke^{(r-\frac{1}{2}\sigma^2)(T-s)+\frac{\sigma}{\sqrt{2}}y_0} - Ke^{-rs}\right)^+ \\ &\approx 0, \quad \text{if } S_0 \ll K, \end{aligned}$$

$$
\begin{aligned}
u(s, y_N) &= e^{rs} g_2(T-s, S_N) \\
&\approx e^{rs}(S_N - Ke^{-rs})^+, \quad S_N\text{: } s\text{에 따라 값이 달라진다.} \\
&= e^{rs}\left(Ke^{(r-\frac{1}{2}\sigma^2)(T-s)+\frac{\sigma}{\sqrt{2}}y_N} - Ke^{-rs}\right)^+ \\
&\approx e^{rs}\left(Ke^{(r-\frac{1}{2}\sigma^2)(T-s)+\frac{\sigma}{\sqrt{2}}y_N} - Ke^{-rs}\right) \quad \text{if } K \ll S_N.
\end{aligned}
$$

또 다른 방식으로는 deep in the money에서는 $\Delta = 1$, deep out of the money에서는 $\Delta = 0$으로 하여, boundary의 값을 정할 수도 있다. 즉,

$$
\frac{C_t(S_N) - C_t(S_{N-1})}{\Delta S} = 1, \quad \frac{C_t(S_1) - C_t(S_0)}{\Delta S} = 0.
$$

2.4.3 Method 3(Time Independent):

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + (r - D_0) S \frac{\partial C}{\partial S} - rC = 0.$$

Change variables: $(t, z) \to (s, y)$

$$\begin{cases} y &= \log \frac{S}{K} \\ t &= T - \dfrac{s}{\frac{1}{2}\sigma^2} = T - \frac{2s}{\sigma^2} \\ v(s,y) &= \frac{1}{K} C(t, S). \end{cases}$$

$$\begin{aligned}
\frac{\partial C}{\partial t} &= K \frac{\partial v}{\partial s} \cdot \frac{\partial s}{\partial t} = -\frac{1}{2} K \sigma^2 \frac{\partial v}{\partial s}, \\
\frac{\partial v}{\partial y} &= \frac{1}{K} \frac{\partial C}{\partial S} \frac{\partial S}{\partial y} = \frac{1}{K} S \frac{\partial C}{\partial S}, \\
\frac{\partial^2 v}{\partial y^2} &= \frac{\partial}{\partial S}\left(\frac{1}{K} S \frac{\partial C}{\partial S}\right) \frac{\partial S}{\partial y} = \frac{1}{K} S \frac{\partial C}{\partial S} + \frac{1}{K} S^2 \frac{\partial^2 C}{\partial S^2} \\
&= \frac{\partial v}{\partial y} + \frac{1}{K} S^2 \frac{\partial^2 C}{\partial S^2}.
\end{aligned}$$

Hence we get

$$\begin{aligned}
-\frac{1}{2} K \sigma^2 \frac{\partial v}{\partial s} + \frac{1}{2}\sigma^2 K \left(\frac{\partial^2 v}{\partial y^2} - \frac{\partial v}{\partial y}\right) + (r - D_0) K \frac{\partial v}{\partial y} - rKv = 0. \\
\frac{\partial v}{\partial s} = \frac{\partial^2 v}{\partial y^2} + \left(\frac{2(r - D_0)}{\sigma^2} - 1\right) \frac{\partial v}{\partial y} - \frac{2r}{\sigma^2} v = 0.
\end{aligned}$$

Now let

$$v(s,y) = e^{Ay+Bs} u(s,y).$$

$$\begin{aligned}
\frac{\partial v}{\partial s} &= Be^{Ay+Bs}\ u + e^{Ay+Bs}\ \frac{\partial u}{\partial s}, \\
\frac{\partial v}{\partial y} &= Ae^{Ay+Bs}\ u + e^{Ay+Bs}\ \frac{\partial u}{\partial y}, \\
\frac{\partial^2 v}{\partial y^2} &= A^2 e^{Ay+Bs}\ u + 2Ae^{Ay+Bs}\ \frac{\partial u}{\partial y} + e^{Ay+Bs}\ \frac{\partial^2 u}{\partial y^2}
\end{aligned}$$

$$\begin{aligned}
Bu + \frac{\partial u}{\partial t} &= A^2 u + 2A \frac{\partial u}{\partial y} + \frac{\partial^2 u}{\partial y^2} + (k_2 - 1)\left(Au + \frac{\partial u}{\partial y}\right) - k_1 u, \\
\frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial y^2} + (2A + k_2 - 1) \frac{\partial u}{\partial y} + \big(A^2 + A(k_2 - 1) - k_1 - B\big),
\end{aligned}$$

where $k_1 = \frac{2r}{\sigma^2}$, $k_2 = \frac{2(r-D_0)}{\sigma^2}$.

Hence we have

$$\begin{cases} A &= -\frac{1}{2}(k_2-1), \\ B &= A^2 + A(k_2-1) - k_1 = \frac{1}{4}(k_2-1)^2 - \frac{1}{2}(k_2-1)^2 - k_1 = -\frac{1}{4}(k_2-1)^2 - k_1. \end{cases}$$

Note that if $D_0 = 0$ then $k_1 = k_2$ and $B = -\frac{1}{4}(k_1+1)^2$.

Finally we get

$$\begin{aligned} \frac{\partial u}{\partial s} &= \frac{\partial^2 u}{\partial y^2}, \quad t = T - \frac{2s}{\sigma^2}, \ S = Ke^y, \\ u(s,y) &= e^{-Ay-Bs} v(s,y) \\ &= \frac{1}{K} e^{-Ay-Bs} C(t,S) \\ &= \frac{1}{K} e^{-Ay-Bs} C(T - \frac{2s}{\sigma^2}, Ke^y), \\ u(0,y) &= \frac{1}{K} e^{-Ay} C(T, Ke^y) \\ &= \frac{1}{K} e^{-Ay} (Ke^y - K)^+ \\ &= (e^{(1-A)y} - e^{-Ay})^+ \\ &= (e^{\frac{1}{2}(k_1+1)y} - e^{\frac{1}{2}(k_1-1)y})^+, \\ u\left(\frac{\sigma^2}{2}T, y\right) &= \frac{1}{K} e^{-Ay - B\frac{\sigma^2}{2}T} C(0, Ke^y), \\ u(s,y_0) &\approx \frac{1}{K} e^{-Ay_0 - Bs} (Ke^{y_0} - Ke^{-r\frac{2s}{\sigma^2}})^+, \\ u(s,y_N) &\approx \frac{1}{K} e^{-Ay_N - Bs} (Ke^{y_N} - Ke^{-r\frac{2s}{\sigma^2}})^+. \end{aligned}$$

	$\frac{1}{K}C(t,S) = v(s,y) = e^{Ay+Bs}u(s,y)$
PDE	$\frac{\partial u}{\partial s} = \frac{\partial^2 u}{\partial y^2}$
new variables	$y = \log\frac{S}{K}, \quad t = T - \frac{2s}{\sigma^2}$
	$k_1 = \frac{2r}{\sigma^2}$
	$A = -\frac{1}{2}(k_1-1), \quad B = -\frac{1}{4}(k_1+1)^2$

Table 2.1: Change of variables with $D_0 = 0$.

Remark 2.4.4 (Dividend).

$$\begin{aligned} S &= Ke^y, \quad t = T - \frac{2s}{\sigma^2}, \quad 0 \le t \le T, 0 \le s \le \frac{\sigma^2}{2}T, \\ C(t,S) &= K \exp\left\{-\frac{1}{2}(k_2-1)y - \left(\frac{1}{4}(k_2-1)^2 + k_1\right)s\right\} u(s,y), \end{aligned}$$

where

$$k_1 = \frac{2r}{\sigma^2}, \quad k_2 = \frac{2(r-D_0)}{\sigma^2},$$

the Black-Scholes equation

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + (r - D_0) S \frac{\partial C}{\partial S} - rC = 0$$

for any European option paying a constant dividend yield can be transformed into the diffusion equation

$$u_s = u_{yy}.$$

	$\frac{1}{K}C(t,S) = v(s,y) = e^{Ay+Bs}u(s,y)$
PDE	$\frac{\partial u}{\partial s} = \frac{\partial^2 u}{\partial y^2}$
new variables	$y = \log\frac{S}{K}, \quad t = T - \frac{2s}{\sigma^2}$
	$k_1 = \frac{2r}{\sigma^2}, \quad k_2 = \frac{2(r-D_0)}{\sigma^2}$
	$A = -\frac{1}{2}(k_2 - 1), \quad B = -\frac{1}{4}(k_2-1)^2 - k_1$

Table 2.2: Change of variables with dividend

2.4.4 Double Asset(Converting to the heat equation)

In this section we follow Clewlow and Strickland(1998). The underlying processes for S_1 and S_2 are given by

$$\begin{aligned} dS_1 &= (r-\delta_1)S_1dt + \sigma_1 S_1 dz_1, \\ dS_2 &= (r-\delta_2)S_1dt + \sigma_2 S_2 dz_2, \end{aligned}$$

and Black-Scholes PDE is given by

$$\frac{\partial f}{\partial t} + \frac{1}{2}\sigma_1^2 S_1^2 \frac{\partial^2 f}{\partial {S_1}^2} + \frac{1}{2}\sigma_2^2 S_2^2 \frac{\partial^2 f}{\partial {S_2}^2} + \rho\sigma_1\sigma_2 S_1 S_2 \frac{\partial^2 f}{\partial S_1 \partial S_2} + (r-\delta_1)S_1\frac{\partial f}{\partial S_1} + (r-\delta_2)S_2\frac{\partial f}{\partial S_2} - rf = 0.$$

Suppose that $-1 \lneq \rho \lneq 1$.

Step 1. Let $x_1 = \ln S_1, x_2 = \ln S_2$.

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= \frac{\partial f}{\partial S_1}\cdot\frac{\partial S_1}{\partial x_1} = \frac{\partial f}{\partial S_1}\cdot S_1, \\ \frac{\partial f}{\partial x_2} &= \frac{\partial f}{\partial S_2}\cdot S_2, \\ \frac{\partial^2 f}{\partial {x_1}^2} &= \frac{\partial}{\partial S_1}\left(\frac{\partial f}{\partial S_1}\cdot S_1\right)\cdot S_1 = \left(\frac{\partial^2 f}{\partial {S_1}^2}S_1 + \frac{\partial f}{\partial S_1}\right)S_1 \\ &= \frac{\partial^2 f}{\partial {S_1}^2}S_1^2 + \frac{\partial f}{\partial S_1}S_1 = \frac{\partial^2 f}{\partial {S_1}^2}S_1^2 + \frac{\partial f}{\partial x_1}, \\ \frac{\partial^2 f}{\partial {x_2}^2} &= \frac{\partial^2 f}{\partial {S_2}^2}S_2^2 + \frac{\partial f}{\partial S_2}S_2 = \frac{\partial^2 f}{\partial {S_2}^2}S_2^2 + \frac{\partial f}{\partial x_2}, \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} &= \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial S_1}S_1\right) = \frac{\partial}{\partial S_2}\left(\frac{\partial f}{\partial S_1}S_1\right)S_2 = \frac{\partial^2 f}{\partial S_1 \partial S_2}S_1 S_2. \end{aligned}$$

Now, we get

$$\begin{aligned} & \frac{\partial f}{\partial t} + \frac{1}{2}\sigma_1^2 S_1^2 \frac{\partial^2 f}{\partial {S_1}^2} + \frac{1}{2}\sigma_2^2 S_2^2 \frac{\partial^2 f}{\partial {S_2}^2} + \rho\sigma_1\sigma_2 S_1 S_2 \frac{\partial^2 f}{\partial S_1 \partial S_2} + (r-\delta_1)S_1\frac{\partial f}{\partial S_1} + (r-\delta_2)S_2\frac{\partial f}{\partial S_2} - rf \\ = & \frac{\partial f}{\partial t} + \frac{1}{2}\sigma_1^2\left(\frac{\partial^2 f}{\partial {x_1}^2} - \frac{\partial f}{\partial x_1}\right) + \frac{1}{2}\sigma_2^2\left(\frac{\partial^2 f}{\partial {x_2}^2} - \frac{\partial f}{\partial x_2}\right) + \rho\sigma_1\sigma_2\frac{\partial^2 f}{\partial x_1 \partial x_2} + (r-\delta_1)\frac{\partial f}{\partial x_1} + (r-\delta_2)\frac{\partial f}{\partial x_2} - rf \\ = & \frac{\partial f}{\partial t} + \frac{1}{2}\sigma_1^2\frac{\partial^2 f}{\partial {x_1}^2} + \frac{1}{2}\sigma_2^2\frac{\partial^2 f}{\partial {x_2}^2} + \rho\sigma_1\sigma_2\frac{\partial^2 f}{\partial x_1 \partial x_2} + \left(r-\delta_1-\frac{1}{2}\sigma_1^2\right)\frac{\partial f}{\partial x_1} + \left(r-\delta_2-\frac{1}{2}\sigma_2^2\right)\frac{\partial f}{\partial x_2} - rf \\ = & \frac{\partial f}{\partial t} + \frac{1}{2}\sigma_1^2\frac{\partial^2 f}{\partial {x_1}^2} + \frac{1}{2}\sigma_2^2\frac{\partial^2 f}{\partial {x_2}^2} + \rho\sigma_1\sigma_2\frac{\partial^2 f}{\partial x_1 \partial x_2} + \nu_1\frac{\partial f}{\partial x_1} + \nu_2\frac{\partial f}{\partial x_2} - rf, \end{aligned}$$

where $\nu_1 = r - \delta_1 - \frac{1}{2}\sigma_1^2, \nu_2 = r - \delta_2 - \frac{1}{2}\sigma_2^2$.

Step 2. Let

$$\partial X = \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \end{bmatrix}, \quad \partial Y = \begin{bmatrix} \frac{\partial}{\partial y_1} \\ \frac{\partial}{\partial y_2} \end{bmatrix}$$

Diagonalization: Find matrices P and D such that

$$P^t A P = D, \quad \text{where } A = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix},$$

$$P = \begin{bmatrix} e_{11} & e_{21} \\ e_{12} & e_{22} \end{bmatrix}, \quad D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

(a) One solution to P and D is given by

$$P = \begin{bmatrix} \sigma_2 & \sigma_2 \\ \sigma_1 & -\sigma_1 \end{bmatrix}, D = \begin{bmatrix} 2(1+\rho)\sigma_1^2\sigma_2^2 & 0 \\ 0 & 2(1-\rho)\sigma_1^2\sigma_2^2 \end{bmatrix}. \tag{2.4.1}$$

For the details, refer to the appendix(D). In my experience, this method is numerically unstable!!

(b) Another solution can be obtained by eigenvectors and eigenvalues. When $\rho \neq 0$,

$$\begin{aligned} \widetilde{P} &= \begin{bmatrix} \frac{\sigma_1^2-\sigma_2^2-\sqrt{\sigma_1^4+2(-1+2\rho^2)\sigma_1^2\sigma_2^2+\sigma_2^4}}{2\rho\sigma_1\sigma_2} & \frac{\sigma_1^2-\sigma_2^2+\sqrt{\sigma_1^4+2(-1+2\rho^2)\sigma_1^2\sigma_2^2+\sigma_2^4}}{2\rho\sigma_1\sigma_2} \\ 1 & 1 \end{bmatrix}, \\ D &= \begin{bmatrix} \frac{1}{2}\left(\sigma_1^2+\sigma_2^2-\sqrt{\sigma_1^4+2(-1+2\rho^2)\sigma_1^2\sigma_2^2+\sigma_2^4}\right) & 0 \\ 0 & \frac{1}{2}\left(\sigma_1^2+\sigma_2^2+\sqrt{\sigma_1^4+2(-1+2\rho^2)\sigma_1^2\sigma_2^2+\sigma_2^4}\right) \end{bmatrix}. \end{aligned}$$

Note that $\widetilde{P}$ is not normalized, hence $\widetilde{P}^{-1} \neq \widetilde{P}^t$. By normalizing each column vector of $\widetilde{P}$ we can get the orthogonal matrix P.

Remark 2.4.5. If $\rho = 1$ or -1, $\lambda_1 = 0$. we exclude this case from the following consideration.

Let $Y = (y_1, y_2)^t, X = (x_1, x_2)^t$ and

$$\begin{aligned} X &= PY \\ \partial X &= P\partial Y. \end{aligned}$$

Define a partial differential operator $\mathcal{L}$ as follows:

$$\mathcal{L} := \frac{\partial}{\partial t} + \frac{1}{2}\sigma_1^2\frac{\partial^2}{\partial {x_1}^2} + \frac{1}{2}\sigma_2^2\frac{\partial^2}{\partial {x_2}^2} + \rho\sigma_1\sigma_2\frac{\partial^2}{\partial x_1 \partial x_2} + \nu_1\frac{\partial}{\partial x_1} + \nu_2\frac{\partial}{\partial x_2} - r$$

Then we get

$$\begin{aligned} \mathcal{L} &= \frac{\partial}{\partial t} + \frac{1}{2}\sigma_1^2\frac{\partial^2}{\partial {x_1}^2} + \frac{1}{2}\sigma_2^2\frac{\partial^2}{\partial {x_2}^2} + \rho\sigma_1\sigma_2\frac{\partial^2}{\partial x_1 \partial x_2} + \nu_1\frac{\partial}{\partial x_1} + \nu_2\frac{\partial}{\partial x_2} - r \\ &= \frac{\partial}{\partial t} + \frac{1}{2}(\partial X)^t A(\partial X) + (\nu_1, \nu_2)(\partial X) - r \\ &= \frac{\partial}{\partial t} + \frac{1}{2}(\partial Y)^t P^t AP(\partial Y) + (\nu_1, \nu_2)P(\partial Y) - r \\ &= \frac{\partial}{\partial t} + \frac{1}{2}(\partial Y)^t \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}(\partial Y) + \begin{bmatrix} e_{11}\nu_1 + e_{12}\nu_2 & e_{21}\nu_1 + e_{22}\nu_2 \end{bmatrix}(\partial Y) - r \end{aligned}$$

This yields that

$$\mathcal{L}(f) = \frac{\partial f}{\partial t} + \frac{1}{2}\lambda_1\frac{\partial^2 C}{\partial {y_1}^2} + \frac{1}{2}\lambda_2\frac{\partial^2 C}{\partial {y_2}^2} + \alpha_1\frac{\partial f}{\partial y_1} + \alpha_2\frac{\partial f}{\partial y_2} - rf = 0$$

where $(\alpha_1, \alpha_2) = (\nu_1, \nu_2)P$.

Step 3. In this step, let us remove the first order terms as followings: Set g as

$$f(y_1, y_2, t) := \exp(a_1 y_1 + a_2 y_2 + a_3 t) g(y_1, y_2, t).$$

$$\begin{aligned}
\frac{\partial f}{\partial t} &= a_3 f + \exp(a_1 y_1 + a_2 y_2 + a_3 t) \frac{\partial g}{\partial t}, \\
\frac{\partial f}{\partial y_1} &= a_1 f + \exp(a_1 y_1 + a_2 y_2 + a_3 t) \frac{\partial g}{\partial y_1} \\
\frac{\partial f}{\partial y_2} &= a_2 f + \exp(a_1 y_1 + a_2 y_2 + a_3 t) \frac{\partial g}{\partial y_2} \\
\frac{\partial^2 f}{\partial {y_1}^2} &= a_1^2 f + 2a_1 \exp(a_1 y_1 + a_2 y_2 + a_3 t) \frac{\partial g}{\partial y_1} + \exp(a_1 y_1 + a_2 y_2 + a_3 t) \frac{\partial^2 g}{\partial {y_1}^2}, \\
\frac{\partial^2 f}{\partial {y_2}^2} &= a_2^2 f + 2a_2 \exp(a_1 y_1 + a_2 y_2 + a_3 t) \frac{\partial g}{\partial y_2} + \exp(a_1 y_1 + a_2 y_2 + a_3 t) \frac{\partial^2 g}{\partial {y_2}^2}.
\end{aligned}$$

$$\begin{aligned}
0 &= a_3 g + \frac{\partial g}{\partial t} + \alpha_1 (a_1 g + \frac{\partial g}{\partial y_1}) + \alpha_2 (a_2 g + \frac{\partial g}{\partial y_2}) \\
&\quad + \frac{1}{2}\lambda_1 (a_1^2 g + 2a_1 \frac{\partial g}{\partial y_1} + \frac{\partial^2 g}{\partial {y_1}^2}) + \frac{1}{2}\lambda_2 (a_2^2 g + 2a_2 \frac{\partial g}{\partial y_2} + \frac{\partial^2 g}{\partial {y_2}^2}) - rg
\end{aligned}$$

$$\left\{ \begin{array}{ll} 0 = & a_3 + \frac{1}{2}\lambda_1 a_1^2 + \frac{1}{2}\lambda_2 a_2^2 + \alpha_1 a_1 + \alpha_2 a_2 - r, \\ 0 = & \alpha_1 + \lambda_1 a_1, \\ 0 = & \alpha_2 + \lambda_2 a_2. \end{array} \right. \tag{2.4.2}$$

The solution to the system (2.4.2) is given by

$$\begin{aligned}
a_1 &= -\frac{\alpha_1}{\lambda_1}, \\
a_2 &= -\frac{\alpha_2}{\lambda_2}, \\
a_3 &= \frac{\alpha_1^2}{2\lambda_1} + \frac{\alpha_2^2}{2\lambda_2} + r.
\end{aligned}$$

Now we obtain the equation for the function g:

$$\frac{\partial g}{\partial t} + \frac{1}{2}\lambda_1 \frac{\partial^2 g}{\partial {y_1}^2} + \frac{1}{2}\lambda_2 \frac{\partial^2 g}{\partial {y_2}^2} = 0.$$

Finally, we could transform the variables so that the coefficients of the second-order terms are equal using

$$z_1 := y_1, \quad z_2 := \sqrt{\frac{\lambda_1}{\lambda_2}}\, y_2, \quad h(z_1, z_2, s) := g(y_1, y_2, T - s),$$

which gives

$$\frac{\partial h}{\partial s} = \frac{1}{2}\lambda_1 \Big(\frac{\partial^2 h}{\partial {z_1}^2} + \frac{\partial^2 h}{\partial {z_2}^2} \Big).$$

✎ **Summing up change of variables**
When we take P from eigenvectors, we have the followings:

- $$\begin{array}{rll} & (S_1, S_2) & \\ \to & (x_1, x_2) & = (\log S_1, \log S_2) \\ \to & (y_1, y_2) & = (e_{11}x_1 + e_{12}x_2, e_{21}x_1 + e_{22}x_2) \\ & & = (e_{11}\log S_1 + e_{12}\log S_2, e_{21}\log S_1 + e_{22}\log S_2) \\ \to & (z_1, z_2) & = \left(y_1, \sqrt{\dfrac{\lambda_1}{\lambda_2}}\, y_2\right) = \left(e_{11}x_1 + e_{12}x_2, \sqrt{\dfrac{\lambda_1}{\lambda_2}}\, (e_{21}x_1 + e_{22}x_2)\right) \\ & & = \left(e_{11}\log S_1 + e_{12}\log S_2, \sqrt{\dfrac{\lambda_1}{\lambda_2}}\Big(e_{21}\log S_1 + e_{22}\log S_2\Big)\right) \end{array}$$

- $$\begin{array}{rll} & (z_1, z_2) & \\ \to & (y_1, y_2) & = \left(z_1, \sqrt{\dfrac{\lambda_2}{\lambda_1}}\, z_2\right) \\ \to & (x_1, x_2) & = \Big(e_{11}y_1 + e_{21}y_2, e_{12}y_1 + e_{22}y_2\Big) = \left(e_{11}z_1 + e_{21}\sqrt{\dfrac{\lambda_2}{\lambda_1}}\, z_2, e_{12}z_1 + e_{22}\sqrt{\dfrac{\lambda_2}{\lambda_1}}\, z_2\right) \\ \to & (S_1, S_2) & = (e^{x_1}, e^{x_2}) = \left(\exp\left\{e_{11}z_1 + e_{21}\sqrt{\dfrac{\lambda_2}{\lambda_1}}\, z_2\right\}, \exp\left\{e_{12}z_1 + e_{22}\sqrt{\dfrac{\lambda_2}{\lambda_1}}\, z_2\right\}\right) \end{array}$$

✎ **Boundary conditions:**

$$f(S_1, S_2), \quad g(y_1, y_2, t), \quad h(z_1, z_2, s).$$

1. $s = 0, t = T$

$$\left\{\begin{array}{rl} & h(z_1, z_2, 0) = g(y_1, y_2, t)\Big|_{y_1 = z_1, y_2 = \sqrt{\frac{\lambda_2}{\lambda_1}}z_2, t=T} \\ = & \exp\left(\frac{\alpha_1}{\lambda_1}y_1 + \frac{\alpha_2}{\lambda_2}y_2 - (\frac{\alpha_1^2}{2\lambda_1} + \frac{\alpha_2^2}{2\lambda_2} + r)T\right) f(S_1, S_2, T) \\ = & \exp\left(\frac{\alpha_1}{\lambda_1}z_1 + \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}z_2 - (\frac{\alpha_1^2}{2\lambda_1} + \frac{\alpha_2^2}{2\lambda_2} + r)T\right) f(e^{x_1}, e^{x_2}, T) \end{array}\right.$$

where
$$x_1 = e_{11}y_1 + e_{21}y_2, \quad x_2 = e_{12}y_1 + e_{22}y_2, \quad S_1 = e^{x_1}, \quad S_2 = e^{x_2}$$
and
$$(\alpha_1, \alpha_2) = (\nu_1, \nu_2)P.$$

If we take P from eigenvectors,

$$y_2 = \sqrt{\frac{\lambda_2}{\lambda_1}}z_2$$

$$x_1 = e_{11}y_1 + e_{21}y_2, \quad x_2 = e_{12}y_1 + e_{22}y_2,$$
$$\nu_1 = r - \delta_1 - \frac{1}{2}\sigma_1^2, \quad \nu_2 = r - \delta_2 - \frac{1}{2}\sigma_2^2,$$
$$\alpha_1 = e_{11}\nu_1 + e_{12}\nu_2, \quad \alpha_2 = e_{21}\nu_1 + e_{22}\nu_2.$$

Hence we get that

$$\begin{aligned} h(z_1, z_2, 0) &= \exp\left(\frac{\alpha_1}{\lambda_1}z_1 + \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}\,z_2 - \left(\frac{\alpha_1^2}{2\lambda_1} + \frac{\alpha_2^2}{2\lambda_2} + r\right)T\right) \\ &\quad \times f\left(\exp\left\{e_{11}z_1 + e_{21}\sqrt{\frac{\lambda_2}{\lambda_1}}\,z_2\right\}, \exp\left\{e_{12}z_1 + e_{22}\sqrt{\frac{\lambda_2}{\lambda_1}}\,z_2\right\}\right). \end{aligned}$$

2. $0 \le s \le T$

$$\left\{\begin{array}{rl} & h(z_1, z_2, s) = g(y_1, y_2, t)\Big|_{y_1 = z_1, y_2 = \sqrt{\frac{\lambda_2}{\lambda_1}}z_2, t = T-s} \\ = & \exp\left(\frac{\alpha_1}{\lambda_1}y_1 + \frac{\alpha_2}{\lambda_2}y_2 - (\frac{\alpha_1^2}{2\lambda_1} + \frac{\alpha_2^2}{2\lambda_2} + r)(T-s)\right)f(S_1, S_2, T-s) \\ = & \exp\left(\frac{\alpha_1}{\lambda_1}z_1 + \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}z_2 - (\frac{\alpha_1^2}{2\lambda_1} + \frac{\alpha_2^2}{2\lambda_2} + r)(T-s)\right)f(S_1, S_2, T-s) \end{array}\right.$$

where

$$x_1 = e_{11}y_1 + e_{21}y_2, \quad x_2 = e_{12}y_1 + e_{22}y_2, \quad S_1 = e^{x_1}, \quad S_2 = e^{x_2}$$

3. $s = T, t = 0$

$$\left\{\begin{array}{rl} & h(z_1, z_2, s) = g(y_1, y_2, t)\Big|_{y_2 = \sqrt{\frac{\lambda_2}{\lambda_1}}z_2, t = 0} \\ = & \exp\left(\frac{\alpha_1}{\lambda_1}y_1 + \frac{\alpha_2}{\lambda_2}y_2\right)f(S_1, S_2, 0) \\ = & \exp\left(\frac{\alpha_1}{\lambda_1}z_1 + \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}z_2\right)f(S_1, S_2, 0) \end{array}\right.$$

where

$$x_1 = e_{11}y_1 + e_{21}y_2, \quad x_2 = e_{12}y_1 + e_{22}y_2, \quad S_1 = e^{x_1}, \quad S_2 = e^{x_2}.$$

Hence we get that

$$f(e^{x_1}, e^{x_2}, 0) = \exp\left(-\frac{\alpha_1}{\lambda_1}z_1 - \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}\,z_2\right)h(z_1, z_2, T).$$

✎ **Greeks:**
Recall that

$$\begin{aligned} h(z_1, z_2, s) &= \exp\left(\frac{\alpha_1}{\lambda_1}z_1 + \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}z_2 - \left(\frac{\alpha_1^2}{2\lambda_1} + \frac{\alpha_2^2}{2\lambda_2} + r\right)(T-s)\right)f(S_1, S_2, T-s) \\ f(S_1, S_2, t) &= \exp\left(-\frac{\alpha_1}{\lambda_1}z_1 - \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}z_2 + \left(\frac{\alpha_1^2}{2\lambda_1} + \frac{\alpha_2^2}{2\lambda_2} + r\right)(T-s)\right)h(z_1, z_2, s) := F(z_1, z_2, s), \end{aligned}$$

where

$$\begin{aligned} z_1 &= y_1 \\ &= e_{11}x_1 + e_{12}x_2 \end{aligned}$$

$$
\begin{aligned}
&= e_{11}\log S_1 + e_{12}\log S_2,\\
z_2 &= \sqrt{\frac{\lambda_1}{\lambda_2}}\, y_2\\
&= \sqrt{\frac{\lambda_1}{\lambda_2}}\Big(e_{21}\log S_1 + e_{22}\log S_2\Big),\\
s &= T-t.
\end{aligned}
$$

1. Delta:

$$
\begin{aligned}
\Delta(S_1) &= \frac{\partial f}{\partial S_1}\\
&= \frac{\partial F}{\partial z_1}\frac{\partial z_1}{\partial S_1} + \frac{\partial F}{\partial z_2}\frac{\partial z_2}{\partial S_1}\\
&= \left\{-\frac{\alpha_1}{\lambda_1}F + \exp\Big(-\frac{\alpha_1}{\lambda_1}z_1 - \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}z_2 + \Big(\frac{\alpha_1^2}{2\lambda_1}+\frac{\alpha_2^2}{2\lambda_2}+r\Big)(T-s)\Big)\frac{\partial h}{\partial z_1}\right\}\frac{e_{11}}{S_1}\\
&\quad + \left\{-\frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}F + \exp\Big(-\frac{\alpha_1}{\lambda_1}z_1 - \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}z_2 + \Big(\frac{\alpha_1^2}{2\lambda_1}+\frac{\alpha_2^2}{2\lambda_2}+r\Big)(T-s)\Big)\frac{\partial h}{\partial z_2}\right\}\sqrt{\frac{\lambda_1}{\lambda_2}}\frac{e_{21}}{S_1}.
\end{aligned}
$$

Let

$$
A = \exp\Big(-\frac{\alpha_1}{\lambda_1}z_1 - \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}z_2\Big).
$$

$$
\begin{aligned}
\Delta(S_1)|_{t=0} &= \left(-\frac{\alpha_1}{\lambda_1}F + \exp\Big(-\frac{\alpha_1}{\lambda_1}z_1 - \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}z_2\Big)\frac{\partial h}{\partial z_1}\right)\frac{e_{11}}{S_1}\\
&\quad + \left(-\frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}F + \exp\Big(-\frac{\alpha_1}{\lambda_1}z_1 - \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}z_2\Big)\frac{\partial h}{\partial z_2}\right)\sqrt{\frac{\lambda_1}{\lambda_2}}\frac{e_{21}}{S_1}\\
&= \left(-\frac{\alpha_1}{\lambda_1}F + A\frac{\partial h}{\partial z_1}\right)\frac{e_{11}}{S_1} + \left(-\frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}F + A\frac{\partial h}{\partial z_2}\right)\sqrt{\frac{\lambda_1}{\lambda_2}}\frac{e_{21}}{S_1}.
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
\Delta(S_2)|_{t=0} &= \left(-\frac{\alpha_1}{\lambda_1}F + \exp\Big(-\frac{\alpha_1}{\lambda_1}z_1 - \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}z_2\Big)\frac{\partial h}{\partial z_1}\right)\frac{e_{12}}{S_2}\\
&\quad + \left(-\frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}F + \exp\Big(-\frac{\alpha_1}{\lambda_1}z_1 - \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}z_2\Big)\frac{\partial h}{\partial z_2}\right)\sqrt{\frac{\lambda_1}{\lambda_2}}\frac{e_{22}}{S_2}\\
&= \left(-\frac{\alpha_1}{\lambda_1}F + A\frac{\partial h}{\partial z_1}\right)\frac{e_{12}}{S_2} + \left(-\frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}F + A\frac{\partial h}{\partial z_2}\right)\sqrt{\frac{\lambda_1}{\lambda_2}}\frac{e_{22}}{S_2}.
\end{aligned}
$$

Note that we can get $\Delta(S_1)|_{t=0}$ and $\Delta(S_2)|_{t=0}$ from $\frac{\partial h}{\partial z_1}$ and $\frac{\partial h}{\partial z_2}$

2. Gamma:

$$
\begin{aligned}
\Gamma(S_1) &= \frac{\partial^2 f}{\partial {S_1}^2}\\
&= \frac{\partial}{\partial S_1}\left(\frac{\partial f}{\partial S_1}\right)\\
&= \frac{\partial}{\partial S_1}\left(\frac{\partial F}{\partial z_1}\frac{\partial z_1}{\partial S_1} + \frac{\partial F}{\partial z_2}\frac{\partial z_2}{\partial S_1}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\partial}{\partial S_1}\left(\frac{\partial F}{\partial z_1}\right)\frac{\partial z_1}{\partial S_1} + \frac{\partial F}{\partial z_1}\frac{\partial^2 z_1}{\partial {S_1}^2} + \frac{\partial}{\partial S_1}\left(\frac{\partial F}{\partial z_2}\right)\frac{\partial z_2}{\partial S_1} + \frac{\partial F}{\partial z_2}\frac{\partial^2 z_2}{\partial {S_1}^2} \\
&= \frac{\partial^2 F}{\partial {z_1}^2}\left(\frac{\partial z_1}{\partial S_1}\right)^2 + \frac{\partial^2 F}{\partial z_1 \partial z_2}\frac{\partial z_2}{\partial S_1}\frac{\partial z_1}{\partial S_1} + \frac{\partial F}{\partial z_1}\frac{\partial^2 z_1}{\partial {S_1}^2} + \frac{\partial^2 F}{\partial z_1 \partial z_2}\frac{\partial z_1}{\partial S_1}\frac{\partial z_2}{\partial S_1} + \frac{\partial^2 F}{\partial {z_2}^2}\left(\frac{\partial z_2}{\partial S_1}\right)^2 + \frac{\partial F}{\partial z_2}\frac{\partial^2 z_2}{\partial {S_1}^2} \\
&= \frac{\partial^2 F}{\partial {z_1}^2}\left(\frac{\partial z_1}{\partial S_1}\right)^2 + 2\frac{\partial^2 F}{\partial z_1 \partial z_2}\frac{\partial z_2}{\partial S_1}\frac{\partial z_1}{\partial S_1} + \frac{\partial^2 F}{\partial {z_2}^2}\left(\frac{\partial z_2}{\partial S_1}\right)^2 + \frac{\partial F}{\partial z_1}\frac{\partial^2 z_1}{\partial {S_1}^2} + \frac{\partial F}{\partial z_2}\frac{\partial^2 z_2}{\partial {S_1}^2},
\end{aligned}
$$

$$
\Gamma(S_2) = \frac{\partial^2 F}{\partial {z_1}^2}\left(\frac{\partial z_1}{\partial S_2}\right)^2 + 2\frac{\partial^2 F}{\partial z_1 \partial z_2}\frac{\partial z_2}{\partial S_2}\frac{\partial z_1}{\partial S_2} + \frac{\partial^2 F}{\partial {z_2}^2}\left(\frac{\partial z_2}{\partial S_2}\right)^2 + \frac{\partial F}{\partial z_1}\frac{\partial^2 z_1}{\partial {S_2}^2} + \frac{\partial F}{\partial z_2}\frac{\partial^2 z_2}{\partial {S_2}^2}, \quad \text{and}
$$

$$
\begin{aligned}
\Gamma(S_1, S_2) &= \frac{\partial}{\partial S_2}\left(\frac{\partial f}{\partial S_1}\right) \\
&= \frac{\partial}{\partial S_2}\left(\frac{\partial F}{\partial z_1}\frac{\partial z_1}{\partial S_1} + \frac{\partial F}{\partial z_2}\frac{\partial z_2}{\partial S_1}\right) \\
&= \frac{\partial^2 F}{\partial {z_1}^2}\frac{\partial z_1}{\partial S_2}\frac{\partial z_1}{\partial S_1} + \frac{\partial^2 F}{\partial z_1 \partial z_2}\frac{\partial z_2}{\partial S_2}\frac{\partial z_1}{\partial S_1} + \cancel{\frac{\partial F}{\partial z_1}\frac{\partial^2 z_1}{\partial S_1 \partial S_2}} + \frac{\partial^2 F}{\partial {z_2}^2}\frac{\partial z_2}{\partial S_2}\frac{\partial z_2}{\partial S_1} + \frac{\partial^2 F}{\partial z_1 \partial z_2}\frac{\partial z_1}{\partial S_2}\frac{\partial z_2}{\partial S_1} + \cancel{\frac{\partial F}{\partial z_2}\frac{\partial^2 z_2}{\partial S_1 \partial S_2}} \\
&= \frac{\partial^2 F}{\partial {z_1}^2}\frac{\partial z_1}{\partial S_2}\frac{\partial z_1}{\partial S_1} + \frac{\partial^2 F}{\partial {z_2}^2}\frac{\partial z_2}{\partial S_2}\frac{\partial z_2}{\partial S_1} + \frac{\partial^2 F}{\partial z_1 \partial z_2}\left(\frac{\partial z_2}{\partial S_2}\frac{\partial z_1}{\partial S_1} + \frac{\partial z_1}{\partial S_2}\frac{\partial z_2}{\partial S_1}\right).
\end{aligned}
$$

Note that

$$
\begin{aligned}
\frac{\partial^2 F}{\partial {z_1}^2}\Big|_{t=0} &= \left(\frac{\alpha_1}{\lambda_1}\right)^2 F - 2\frac{\alpha_1}{\lambda_1}A\frac{\partial h}{\partial z_1} + A\frac{\partial^2 h}{\partial {z_1}^2}, \\
\frac{\partial^2 F}{\partial z_1 \partial z_2}\Big|_{t=0} &= \frac{\alpha_1\alpha_2}{\lambda_1\sqrt{\lambda_1\lambda_2}}F - \frac{\alpha_1}{\lambda_1}A\frac{\partial h}{\partial z_2} - \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}A\frac{\partial h}{\partial z_1} + A\frac{\partial^2 h}{\partial z_1 \partial z_2}, \\
\frac{\partial^2 F}{\partial {z_2}^2}\Big|_{t=0} &= \left(\frac{\alpha_2^2}{\lambda_1\lambda_2}\right) F - 2\frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}A\frac{\partial h}{\partial z_2} + A\frac{\partial^2 h}{\partial {z_2}^2} \\
\frac{\partial F}{\partial z_1}\Big|_{t=0} &= -\frac{\alpha_1}{\lambda_1}F + A\frac{\partial h}{\partial z_1}, \\
\frac{\partial F}{\partial z_2}\Big|_{t=0} &= -\frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}F + A\frac{\partial h}{\partial z_2}.
\end{aligned}
$$

$$
\begin{aligned}
\Gamma(S_1)|_{t=0} &= \left\{\left(\frac{\alpha_1}{\lambda_1}\right)^2 F - 2\frac{\alpha_1}{\lambda_1}A\frac{\partial h}{\partial z_1} + A\frac{\partial^2 h}{\partial {z_1}^2}\right\}\left(\frac{e_{11}}{S_1}\right)^2 \\
&\quad +2\left\{\frac{\alpha_1\alpha_2}{\lambda_1\sqrt{\lambda_1\lambda_2}}F - \frac{\alpha_1}{\lambda_1}A\frac{\partial h}{\partial z_2} - \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}A\frac{\partial h}{\partial z_1} + A\frac{\partial^2 h}{\partial z_1 \partial z_2}\right\}\left(\frac{e_{11}}{S_1}\right)\sqrt{\frac{\lambda_1}{\lambda_2}}\left(\frac{e_{21}}{S_1}\right) \\
&\quad +\left\{\left(\frac{\alpha_2^2}{\lambda_1\lambda_2}\right) F - 2\frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}A\frac{\partial h}{\partial z_2} + A\frac{\partial^2 h}{\partial {z_2}^2}\right\}\frac{\lambda_1}{\lambda_2}\left(\frac{e_{21}}{S_1}\right)^2 \\
&\quad -\left(-\frac{\alpha_1}{\lambda_1}F + A\frac{\partial h}{\partial z_1}\right)\frac{e_{11}}{S_1^2} - \left(-\frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}F + A\frac{\partial h}{\partial z_2}\right)\sqrt{\frac{\lambda_1}{\lambda_2}}\frac{e_{21}}{S_1^2}.
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
\Gamma(S_2)|_{t=0} &= \left\{\left(\frac{\alpha_1}{\lambda_1}\right)^2 F - 2\frac{\alpha_1}{\lambda_1}A\frac{\partial h}{\partial z_1} + A\frac{\partial^2 h}{\partial {z_1}^2}\right\}\left(\frac{e_{12}}{S_2}\right)^2 \\
&\quad +2\left\{\frac{\alpha_1\alpha_2}{\lambda_1\sqrt{\lambda_1\lambda_2}}F - \frac{\alpha_1}{\lambda_1}A\frac{\partial h}{\partial z_2} - \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}A\frac{\partial h}{\partial z_1} + A\frac{\partial^2 h}{\partial z_1 \partial z_2}\right\}\left(\frac{e_{12}}{S_2}\right)\sqrt{\frac{\lambda_1}{\lambda_2}}\left(\frac{e_{22}}{S_2}\right) \\
&\quad +\left\{\left(\frac{\alpha_2^2}{\lambda_1\lambda_2}\right) F - 2\frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}A\frac{\partial h}{\partial z_2} + A\frac{\partial^2 h}{\partial {z_2}^2}\right\}\frac{\lambda_1}{\lambda_2}\left(\frac{e_{22}}{S_2}\right)^2 \\
&\quad -\left(-\frac{\alpha_1}{\lambda_1}F + A\frac{\partial h}{\partial z_1}\right)\frac{e_{12}}{S_2^2} - \left(-\frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}F + A\frac{\partial h}{\partial z_2}\right)\sqrt{\frac{\lambda_1}{\lambda_2}}\frac{e_{22}}{S_2^2},
\end{aligned}
$$

and the cross gamma

$$\begin{aligned}\Gamma(S_1,S_2)|_{t=0} &= \left\{\left(\frac{\alpha_1}{\lambda_1}\right)^2 F - 2\frac{\alpha_1}{\lambda_1}A\frac{\partial h}{\partial z_1} + A\frac{\partial^2 h}{\partial {z_1}^2}\right\}\frac{e_{11}}{S_1}\frac{e_{12}}{S_2} \\ &+\left\{\left(\frac{\alpha_2^2}{\lambda_1\lambda_2}\right)F - 2\frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}A\frac{\partial h}{\partial z_2} + A\frac{\partial^2 h}{\partial {z_2}^2}\right\}\frac{\lambda_1}{\lambda_2}\frac{e_{21}}{S_1}\frac{e_{22}}{S_2} \\ &+\left\{\frac{\alpha_1\alpha_2}{\lambda_1\sqrt{\lambda_1\lambda_2}}F - \frac{\alpha_1}{\lambda_1}A\frac{\partial h}{\partial z_2} - \frac{\alpha_2}{\sqrt{\lambda_1\lambda_2}}A\frac{\partial h}{\partial z_1} + A\frac{\partial^2 h}{\partial z_1\partial z_2}\right\}\sqrt{\frac{\lambda_1}{\lambda_2}}\frac{e_{11}e_{22}+e_{12}e_{21}}{S_1S_2}.\end{aligned}$$

Example 2.4.6 (2 Stock ELS(Multi-Chances)). Let us evaluate a popular ELS, which is called six chances. The final(maturity) payoff is given by

Valuation Date	Relevant Rate	Early Termination Condition
11 July 2005	104.75	$\min(S_1,S_2)\geq 85$
9 January 2006	109.5	$\min(S_1,S_2)\geq 85$
11 July 2006	114.25	$\min(S_1,S_2)\geq 85$
9 January 2007	119	$\min(S_1,S_2)\geq 85$
11 July 2007	123.75	$\min(S_1,S_2)\geq 85$
9 January 2008(maturity)	128.5	$\min(S_1,S_2)\geq 85$

Table 2.3: Relevant rate

$$\text{maturity payoff} = \begin{cases} 128.5, & \text{if} \quad \min(S_1,S_2)\geq 85 \\ \min(S_1,S_2), & \text{if} \quad \min(S_1,S_2) < 85 \end{cases}$$

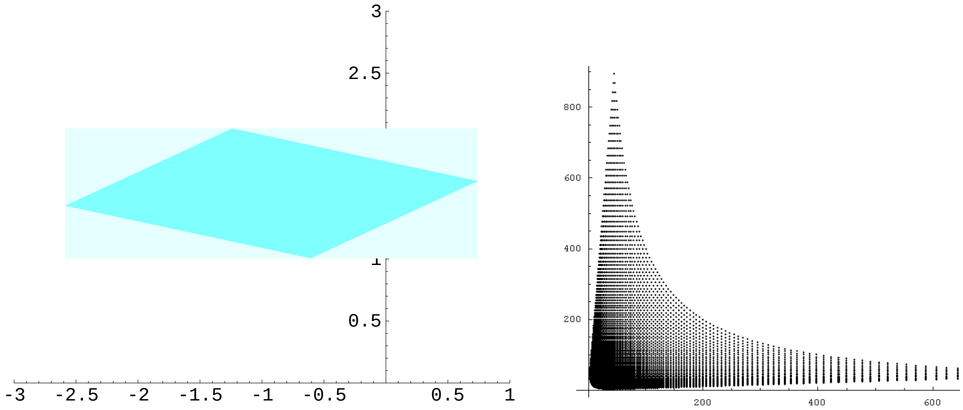
To illustrate how it is calculated, we suppose that the parameters are as shown in Table (2.4). Figure (2.1(a)) demonstrates the extended region after transformation to the heat equation and

	(S_1,S_2)	Interest(r)	Volatility(σ_1,σ_2)	correlation(ρ)	dividend
underlying	$(100,100)$	3.2%	(27%, 30%)	0.8	(1%,2%)

Table 2.4: Parameters

Figure (2.1(b)) demonstrates the transformed solution from the solution of the heat equation. Figure (2.2) demonstrates the FDM solution to the price of ELS.

□



(a) Region for heat equation (b) Restored region

Figure 2.1: Computation regions

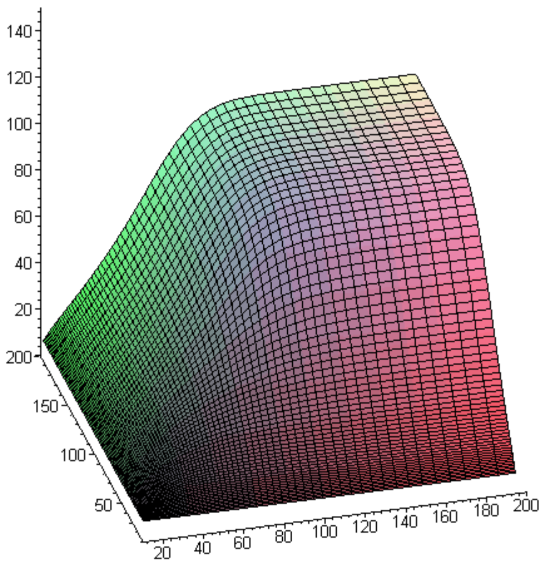


Figure 2.2: Solution for ELS; valuation date-9 January 2005)

2.5 A Generalization of The Black-Scholes PDE

suppose that a function, f, depends on the n variables $x_1, \cdots x_n$ and time t. Suppose further that x_i follows an Ito process

$$dx_i = a_i\,dt + b_i\,dW_i(t), \quad (1 \le i \le n).$$

Assume that the correlation, ρ_{ij}, between dW_i and dW_j is given, we have

$$\begin{aligned} df &= \frac{\partial f}{\partial t}\,dt + \sum_i \frac{\partial f}{\partial x_i}\,dx_i + \frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j} d\langle x_i, x_j\rangle \\ &= \left(\frac{\partial f}{\partial t} + \sum_i \frac{\partial f}{\partial x_i} a_i + \frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j} b_i b_j \rho_{ij}\right) dt + \sum_i \frac{\partial f}{\partial x_i} b_i\,dW_i. \end{aligned}$$

2.5.1 Derivation of Differential Equation with Multi-Asset

Suppose that

$$\frac{dS_i}{S_i} = \mu_i\,dt + \sigma_i\,dW_i, \quad 1 \le i \le n.$$

Let $f_k (1 \le k \le n+1)$ be the price of kth traded security. It follows from Ito's lemma that f_j follow diffusion processes:

$$df_k = \alpha_k f_k\,dt + \sum_{i=1}^{n} \beta_{ik} f_k\,dW_i \quad (1 \le k \le n+1)$$

where

$$\alpha_k f_k = \frac{\partial f_k}{\partial t} + \sum_{i=1}^{n} \frac{\partial f_k}{\partial S_i}\mu_i S_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij}\,\sigma_i\,\sigma_j\,S_i\,S_k\,\frac{\partial^2 f_k}{\partial S_i \partial S_j}, \tag{2.5.1}$$

$$\beta_{ik} f_k = \frac{\partial f_k}{\partial S_i}\sigma_i S_i. \tag{2.5.2}$$

Because there are $n+1$ traded securities and n Wiener processes, it is possible to form an instantaneously riskless portfolio, Π, using the securities. Define Δ_k as the amount of the kth security in the portfolio, so that

$$\begin{aligned} \Pi &= \sum_{k=1}^{n+1} \Delta_k f_k, \\ d\Pi &= \sum_{k=1}^{n+1} \Delta_k df_k \\ &= \sum_{k=1}^{n+1} \Delta_k \left(\alpha_k f_k\,dt + \sum_{i=1}^{n} \beta_{ik} f_k\,dW_i\right) \\ &= \sum_{k=1}^{n+1} \Delta_k \alpha_k f_k\,dt \quad + \quad \sum_{k=1}^{n+1}\sum_{i=1}^{n} \Delta_k \beta_{ik} f_k\,dW_i \end{aligned}$$

$$= \sum_{k=1}^{n+1} \alpha_k f_k \Delta_k \, dt \quad + \quad \sum_{i=1}^{n} \left(\sum_{k=1}^{n+1} \beta_{ik} f_k \Delta_k \right) dW_i$$

Since $d\Pi = r\Pi dt = r \left(\sum_{k=1}^{n+1} \Delta_k f_k \right) dt$, we have that

$$\begin{cases} \sum_{k=1}^{n+1} \beta_{ik} f_k \Delta_k & = 0, \quad 1 \le i \le n, \\ \sum_{k=1}^{n+1} (\alpha_k - r) f_k \Delta_k & = 0. \end{cases}$$

This equations can be regarded as $n+1$ homogeneous linear equations in the Δ_k's. The Δ_k's are not all zero. From a well-known theorem in linear algebra, we can see that for all k, $(\alpha_k - r)f_k$ can be written as a linear combination of $\beta_{1k} f_k, \beta_{2k} f_k, \cdots, \beta_{nk} f_k$. i.e.

$$(\alpha_k - r) f_k \quad = \quad \sum_{i=1}^{n} \lambda_i \beta_{ik} f_k, \quad (1 \le k \le n+1) \tag{2.5.3}$$

or

$$(\alpha_k - r) \quad = \quad \sum_{i=1}^{n} \lambda_i \beta_{ik}, \quad (1 \le k \le n+1)$$

for some $\lambda_i (1 \le i \le n)$ that are dependent only on the state variables and time(independent of k). Substituting from equations (2.5.1) and (2.5.2) into equation (2.5.3), we obtain

$$\frac{\partial f_k}{\partial t} + \sum_{i=1}^{n} \frac{\partial f_k}{\partial S_i} \mu_i S_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} \sigma_i \sigma_j S_i S_k \, \frac{\partial^2 f_k}{\partial S_i \partial S_j} - r f_k \quad = \quad \sum_{i=1}^{n} \lambda_i \frac{\partial f_k}{\partial S_i} \sigma_i S_i$$

that reduced to

$$\frac{\partial f_k}{\partial t} + \sum_{i=1}^{n} S_i (\mu_i - \lambda_i \sigma_i) \frac{\partial f_k}{\partial S_i} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} \sigma_i \sigma_j S_i S_k \, \frac{\partial^2 f_k}{\partial S_i \partial S_j} \quad = \quad r f_k.$$

Dropping the subscripts to f, we deduce that any security whose price, f, is contingent on the underlying $S_i (1 \le i \le n)$ and time, t, satisfies the second-order differential equation

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{n} S_i (\mu_i - \lambda_i \sigma_i) \frac{\partial f}{\partial S_i} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} \sigma_i \sigma_j S_i S_k \, \frac{\partial^2 f}{\partial S_i \partial S_j} \quad = \quad r f. \tag{2.5.4}$$

The particular derivatives security that is obtained is determined by the boundary condition that are imposed on equation (2.5.4).

2.5.2 Time-Dependent Parameters

In this subsection, we follows Wilmott(2000). The Black-Scholes PDE is valid as long as the parameters r, D and σ are known function of time(not random), where r is interest, D dividend rate, and σ volatility, respectively. The PDE that we must solve is now

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2(t) S^2 \frac{\partial^2 V}{\partial S^2} + (r(t) - D(t)) S \frac{\partial V}{\partial S} - r(t) V \quad = \quad 0. \tag{2.5.5}$$

Introduce new variables as follows.

$$\overline{S} = Se^{\alpha(t)}, \quad \bar{t} = \gamma(t), \quad \overline{V} = Ve^{\beta(t)}.$$

We will choose α, β and γ so as to eliminate all time dependent coefficients from Eq(2.5.5). By chain rule, we obtain

$$\begin{aligned}
\frac{\partial V}{\partial t} &= -\dot{\beta}(t)e^{-\beta(t)}\overline{V} + e^{-\beta(t)}\left(\frac{\partial \overline{V}}{\partial \bar{t}}\frac{\partial \bar{t}}{\partial t} + \frac{\partial \overline{V}}{\partial \overline{S}}\frac{\partial \overline{S}}{\partial t}\right), \\
&= -\dot{\beta}(t)e^{-\beta(t)}\overline{V} + e^{-\beta(t)}\dot{\gamma}(t)\frac{\partial \overline{V}}{\partial \bar{t}} + e^{-\beta(t)}\,\overline{S}\,\dot{\alpha}(t)\frac{\partial \overline{V}}{\partial \overline{S}}, \\
\frac{\partial V}{\partial S} &= e^{-\beta(t)}\frac{\partial \overline{V}}{\partial \overline{S}} \\
&= e^{-\beta(t)}\frac{\partial \overline{V}}{\partial \overline{S}}\frac{\partial \overline{S}}{\partial S} \\
&= e^{-\beta(t)}\,e^{\alpha(t)}\frac{\partial \overline{V}}{\partial \overline{S}}, \\
\frac{\partial^2 V}{\partial S^2} &= e^{-\beta(t)}\,e^{2\alpha(t)}\frac{\partial^2 \overline{V}}{\partial \overline{S}^2}.
\end{aligned}$$

After changing variables Eq(2.5.5) becomes

$$\dot{\gamma}(t)\frac{\partial \overline{V}}{\partial \bar{t}} + \frac{1}{2}\sigma^2(t)\overline{S}^2\frac{\partial^2 \overline{V}}{\partial \overline{S}^2} + (r(t) - D(t) + \dot{\alpha}(t))\overline{S}\frac{\partial \overline{V}}{\partial \overline{S}} - (r(t) + \dot{\beta}(t))\overline{V} = 0. \tag{2.5.6}$$

By choosing

$$\beta(t) = \int_t^T r(s)ds, \quad \alpha(t) = \int_t^T (r(s) - D(s))ds, \quad \text{and} \quad \gamma(t) = \int_t^T \sigma^2(s)ds,$$

Eq(2.5.6) becomes

$$\frac{\partial \overline{V}}{\partial \bar{t}} = \frac{1}{2}\overline{S}^2\frac{\partial^2 \overline{V}}{\partial \overline{S}^2}. \tag{2.5.7}$$

If $\overline{V}(\overline{S}, \bar{t})$ is any solution of Eq(2.5.7), then

$$V = e^{-\beta(t)}\,\overline{V}(Se^{\alpha(t)}, \gamma(t)). \tag{2.5.8}$$

Now we use V_{BS} to denote any solution of the Black-Scholes equation for constant interest r_c, dividend rate D_c and volatility σ_c. This solution can be written in the form

$$V_{BS} = e^{-r_c(T-t)}\overline{V}\Big(Se^{(r_c-D_c)(T-t)}, \sigma_c^2(T-t)\Big). \tag{2.5.9}$$

By comparing Eq(2.5.8) and Eq(2.5.9) it follows that the solution of the time dependent parameter problem is the same as the solution of the constant parameter problem if we use the following substitutions:

$$r_c = \frac{1}{T-t}\int_t^T r(s)ds,$$

$$
\begin{aligned}
D_c &= \frac{1}{T-t}\int_t^T D(s)ds, \\
\sigma_c^2 &= \frac{1}{T-t}\int_t^T \sigma^2(s)ds.
\end{aligned}
$$

For example, the formula for a European call option with time dependent parameter is given by

$$
Se^{-\int_t^T D(s)ds}N(d_1) - Ke^{-\int_t^T r(s)ds}N(d_2),
$$

where

$$
\begin{aligned}
d_1 &= \frac{\log\frac{S}{K} + \int_t^T (r(s)-D(s))ds + \frac{1}{2}\int_t^T \sigma^2(s)ds}{\sqrt{\int_t^T \sigma^2(s)ds}}, \\
d_2 &= d_1 - \sqrt{\int_t^T \sigma^2(s)ds}.
\end{aligned}
$$

Remark 2.5.1. If $V(S,T) = (S-K)^+$, the solution of Eq(2.5.7) is given by

$$
\overline{V}(\overline{S},\bar{t}) = \overline{S}N\left(\frac{\log\frac{\overline{S}}{K} + \frac{1}{2}\bar{t}}{\sqrt{\bar{t}}}\right) - KN\left(\frac{\log\frac{\overline{S}}{K} - \frac{1}{2}\bar{t}}{\sqrt{\bar{t}}}\right).
$$

2.6 Markov Process & The Kolmogorov Equations

Consider

$$dX_t = a(X_t)dt + \sigma(X_t)dB_t.$$

Let $h(y)$ be a function, and define

$$v(t,x) = E_{t,x}\Big[h(X_T)\Big], \quad 0 \le t \le T.$$

Then

$$\begin{aligned} v(t,x) &= \int h(y)p(T-t,x,y)dy, \\ v_t(t,x) &= -\int h(y)p_t(T-t,x,y)dy, \\ v_x(t,x) &= \int h(y)p_x(T-t,x,y)dy, \\ v_{xx}(t,x) &= \int h(y)p_{xx}(T-t,x,y)dy, \end{aligned}$$

Theorem 2.6.1. Starting at $X_0 = \xi$, the process $v(t, X_t)$ satisfies the martingale property:

$$E\Big[v(t,X_t)|\mathcal{F}_s\Big] = v(s,X_s), \quad 0 \le s \le t \le T.$$

Proof. According to the Markov property,

$$E\big[h(X_T)|\mathcal{F}_t\big] = E_{t,X_t}\Big[h(X_T)\Big] = v(t,X_t),$$

so

$$\begin{aligned} E\Big[v(t,X_t)|\mathcal{F}_s\Big] &= E\Big[E\big[h(X_T)|\mathcal{F}_t\big]|\mathcal{F}_s\Big] \\ &= E\big[h(X_T)|\mathcal{F}_s\big] \\ &= E_{s,X_s}\big[h(X_T)\big] \\ &= v(s,X_s). \end{aligned}$$

■

Theorem 2.6.2 (The Feynman-Kac Formula). Let X_t be a Ito diffusion.

$$v(t,x) = E_{t,x}\Big[h(X_T)\Big], \quad 0 \le t \le T.$$

where $dX_t = a(X_t)dt + \sigma(X_t)dB_t$. Then

$$v_t(t,x) + a(x)v_x(t,x) + \frac{1}{2}\sigma^2(x)v_{xx}(t,x) = 0, \quad v(T,x) = h(x).$$

2.7 Dividend Paying Assets

The dividend yield is defined as the ratio of the dividend payment to the asset price. Thus the dividend D_0Sdt represents a constant and continuous dividend yield. This dividend structure is a good model for index options; the many of discrete dividend payments on a large index can be approximated by a continuous yield without serious error. The asset price is modeled as

$$\begin{aligned} dS_t &= (\mu - D_0)Sdt + \sigma S dW_t, \\ \Pi &= f - \Delta S, \\ d\Pi &= df - \Delta dS - D_0 S \Delta dt. \end{aligned}$$

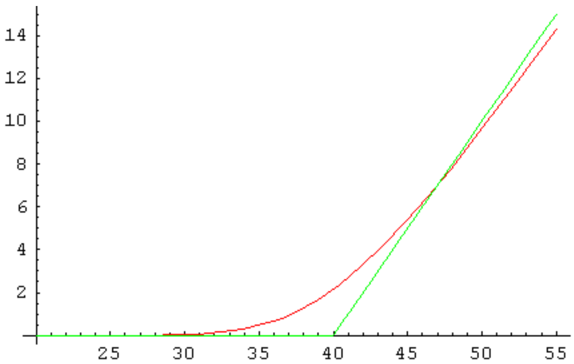
After measure change, we get

$$dS_t = rSdt + \sigma S d\widetilde{W}_t$$

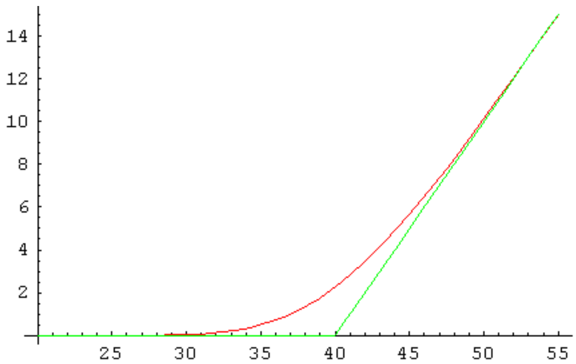
where $\widetilde{W}_t$ is a Brownian motion under risk neutral measure. No arbitrage condition gives us

$$\begin{aligned} 0 &= d\Pi - r\Pi dt \\ &= df - \Delta dS - D_0 S\Delta dt - r(f - \Delta S)dt \\ &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial S}dS + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}(dS)^2 - \Delta dS - D_0 S \Delta dt - r(f - \Delta S)dt \\ &= \Big(\frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} + (r - D_0)S\frac{\partial f}{\partial S} - rf\Big)dt. \end{aligned}$$

This model is also applicable to options on foreign currencies with $D_0 = r_f$.



(a) Call option: $K = 40, T = 0.5, r = 0.1, v = 0.2, d = 0.1$



(b) Call option: $K = 40, T = 0.5, r = 0, v = 0.2$

Figure 2.3: Option with dividend.

2.7.1 Boundary Conditions

Call Option:

$$\begin{aligned} f(0,t) &= 0, \\ f(S,t) &\sim Se^{-D_0(T-t)} \quad \text{as } S \to \infty. \end{aligned}$$

This is because in the limit $S \to \infty$, the option becomes equivalent to the asset but without its dividend income.

Part II

Pricing Financial Instruments

Chapter 3

American Options

3.1 The Perpetual American Put

The perpetual American put contract can be exercised for a put payoff at any time. There is no expiry. We want to find the value of the perpetual American option before exercise.

1. The solution, V is independent of time t. Thus it must satisfy

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0. \tag{3.1.1}$$

2. The option value can never go below the early-exercise payoff. i.e.

$$V \geq \max(K - S, 0).$$

Note that Eq(3.1.1) is the Euler Equation[1] . Hence the solution can be obtained by

$$\begin{aligned} \lambda(\lambda-1) + \frac{2r}{\sigma^2}\lambda - \frac{2r}{\sigma^2} &= 0, \\ (\lambda - 1)\left(\lambda + \frac{2r}{\sigma^2}\right) &= 0. \end{aligned}$$

Thus the general solution of this second-order ordinary equation is

$$V(S) = AS + BS^{-\frac{2r}{\sigma^2}}$$

[1]The Euler equation is given by

$$L[y] = x^2 y'' + \alpha x y' + \beta y = 0.$$

If we assume that Euler equation has a solution of the form

$$y = x^\lambda,$$

then we obtain

$$L[x^\lambda] = x^\lambda \Big\{\lambda(\lambda-1) + \alpha\lambda + \beta\Big\} = 0.$$

where A and B are arbitrary constants.

Clearly, for the perpetual American put the coefficient A must be zero, since as $S \to \infty$ the value of the option must tend to zero.

Suppose $S = S^*$ is the optimal exercise point. Naturally, since we don't exercise when $S \geq K$, $S^* < K$. Then

$$V(S^*) = K - S^*.$$

From this we get

$$\begin{aligned} B(S^*)^{-\frac{2r}{\sigma^2}} &= K - S^*, \\ B &= (K - S^*)(S^*)^{\frac{2r}{\sigma^2}}, \\ \therefore \quad V(S) &= (K - S^*)\left(\frac{S}{S^*}\right)^{-\frac{2r}{\sigma^2}}. \end{aligned}$$

There are two methods to get S^*.

1. **Smooth-pasting condition.** The American option value is maximized by an exercise strategy that makes the option value and option delta continuous. i.e.

   $$V'(S^*) = -1.$$

   Thus we have

   $$\begin{aligned} V'(S) &= -\left(\frac{2r}{\sigma^2}\right)(K - S^*)\left(\frac{S}{S^*}\right)^{-\frac{2r}{\sigma^2}-1}\frac{1}{S^*}, \\ V'(S^*) &= -\left(\frac{2r}{\sigma^2}\right)\left(\frac{K}{S^*} - 1\right) = -1, \\ \therefore \quad S^* &= \frac{K}{1 + \frac{\sigma^2}{2r}}. \end{aligned}$$

2. **Maximization of the Option's Value with respect to S^*.** We find the value by differentiating V with respect to S^*(why? see the following Remark.) and setting the resulting expression equal to zero. For a fixed S, define ϕ_S as

   $$\phi_S(z) = (K - z)\left(\frac{S}{z}\right)^{-\frac{2r}{\sigma^2}}.$$

   We have

   $$\begin{aligned} \phi'_S(z) &= -\left(\frac{S}{z}\right)^{-\frac{2r}{\sigma^2}} + (K - z)\left(\frac{2r}{\sigma^2}\right)\left(\frac{S}{z}\right)^{-\frac{2r}{\sigma^2}+1}\frac{1}{S} \\ &= \frac{1}{z}\left(\frac{S}{z}\right)^{-\frac{2r}{\sigma^2}}\left(-z + \frac{2r}{\sigma^2}(K - z)\right) = 0, \\ \therefore \quad z &= \frac{K}{1 + \frac{\sigma^2}{2r}} := S^*. \end{aligned}$$

Remark 3.1.1. Let $u(S)$ be the perpetual option price. $u(S)$ is given by

$$\begin{aligned} u(S) &= \sup_{\substack{0\le\tau\le\infty \\ \tau:\text{stopping time}}} E[e^{-r\tau}(K-S_\tau)^+] \\ &= E\left[\left(e^{-r\tau_S^*}K - S\exp\left(-\frac{\sigma^2}{2}\tau^* + \sigma W_{\tau_S^*}\right)\right)^+ \mathbf{1}_{\{\tau_S^*<\infty\}}\right], \end{aligned}$$

where τ_S^* is the stopping time defined by

$$\begin{aligned} \tau_S^* &= \inf\{t\ge 0 | u(X_t^S) = (K - X_t^S)^+\}, \\ X_t^S &= S\exp\left(\left(r-\frac{\sigma^2}{2}\right)t + \sigma W_t\right). \end{aligned}$$

Thus we have

$$\tau_S^* = \inf\left\{t\ge 0 | X_t^S \le S^*\right\} = \inf\left\{t \ge 0 \middle| \left(r - \frac{\sigma^2}{2}\right)t + \sigma W_t \le \log\frac{S^*}{S}\right\}.$$

Introduce, for any $z\in\mathbb{R}^+$, the stopping time $\tau_{S,z}$ defined by

$$\tau_{S,z} = \inf\{t\ge 0|X_t^S \le z\}.$$

We fix S and define the following function $\phi(z)$ by

$$\phi(z) = E\left[e^{-r\tau_{S,z}}\mathbf{1}_{\{\tau_{S,z}<\infty\}}\left(K - X^S_{\tau_{S,z}}\right)^+\right].$$

$\phi(z)$ is maximized when $z = S^*$.

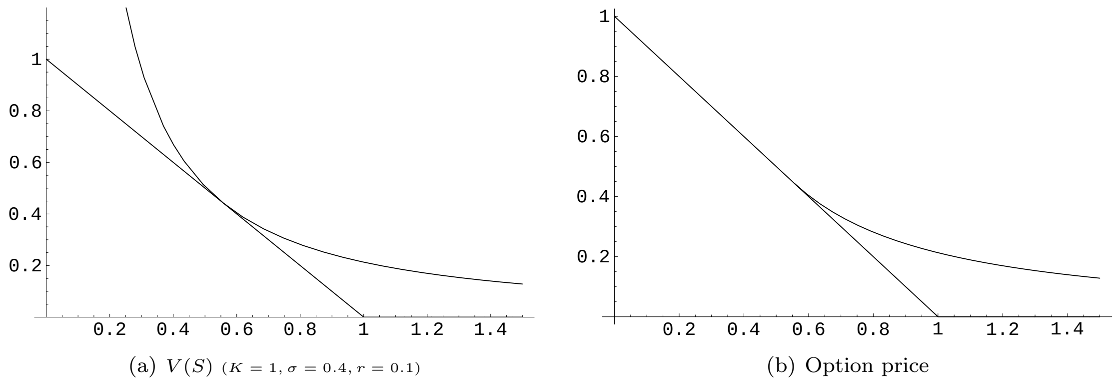


Figure 3.1: The solution to the perpetual American put.

3.2 Mathematical Formulation

When the contract is American the long/short relationship is asymmetrical; it is the holder of the exercise rights who controls the early-exercise feature. The writer of the option can do no more than sit back and enjoy the view. If V is the value of a long position in an American option then all we can say is that we can earn no more than the risk-free rate on our portfolio. Thus we arrive at the inequality

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV \le 0.$$

Suppose that $V = K - S$ for some $S < K$. If this is the case then V most certainly does not satisfy the Black-Scholes equation (unless $r = 0$) since

$$\frac{\partial}{\partial t}(K-S) + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2}(K-S) + rS\frac{\partial}{\partial S}(K-S) - r(K-S) = -rK < 0.$$

But V does satisfy the inequality. When $V = K - S$ the return from the portfolio is less than the return from an equivalent bank deposit, and hence it is optimal to exercise the option.

At any given time t, we must divide the S axis into two distinct regions, one where early exercise is optimal and

$$V = K - S, \quad \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV < 0,$$

and the other where early exercise is not optimal and

$$V > K - S, \quad \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

Let $S_f(t)$ be defined to be the largest value of S, at time t, for which we have $V(t,S) = \max(K-S, 0)$. $S_f(t)$ is called the free boundary.

3.3 Buyer's & Seller's Price

✎ **Notation:**

1. T: expiry.
2. τ: exercise time = stopping time. τ is assumed to satisfy $0 \le \tau \le T$.

In general the payoff is $\mathcal{F}_t$-adapted stochastic process X_t. Its payoff at exercise time τ is X_τ. Fundamental asymmetry between the buyer & seller:

1. Buyer's viewpoint:
   A buyer would be willing to pay A^b if he(or she) can find a hedging strategy (ϕ_t, ψ_t) and a stopping time τ such that

   $$-A^b + \int_0^\tau \phi_s dZ_s + B_\tau^{-1} X_\tau \ge 0, \quad a.s.$$

   This equation means that

(initial payment) + (the present value of trading gain/loss up to stopping time) + (the present value of payoff at the stopping time).

Thus the maximal price the buyer is willing to pay is

$$\begin{aligned}\sup\Big\{A^b \quad : \quad & \exists\tau(\text{stopping time}), \exists\phi(\text{ predictable trading strategy}) \text{ such that} \\ & -A^b + \int_0^\tau \phi_s dZ_s + B_\tau^{-1}X_\tau \geq 0\Big\}\end{aligned}$$

By taking expectation,

$$\begin{aligned} -A^b &\geq -E_Q\left[\int_0^\tau \phi_s dZ_s\right] - E_Q[B_\tau^{-1}X_\tau] = -E_Q[B_\tau^{-1}X_\tau], \\ A^b &\leq E_Q[B_\tau^{-1}X_\tau]. \\ A^b &\leq \sup_{\substack{0\leq\tau\leq T \\ \tau:\text{stopping time}}} E_Q[B_\tau^{-1}X_\tau]. \end{aligned}$$

2. Seller's viewpoint:
A seller would be willing to pay A^s if he(or she) can find a hedging strategy (ϕ_t, ψ_t) such that for any stopping time τ

$$A^s + \int_0^\tau \phi_s dZ_s - B_\tau^{-1}X_\tau \geq 0, \quad a.s.$$

Thus

$$A^s \quad \geq \quad \sup_{\substack{0\leq\tau\leq T \\ \tau:\text{stopping time}}} E_Q[B_\tau^{-1}X_\tau].$$

Theorem 3.3.1. In complete market,

$$A^b = \sup_{\substack{0\leq\tau\leq T \\ \tau:\text{stopping time}}} E_Q[B_\tau^{-1}X_\tau] = A^s.$$

Proof. See Oksendal(2003) page 277. ■

3.4 Optimal Stopping & American Options(discrete)

Since an American option can be exercised at any time between 0 and N, we shall define it as a positive sequence X_n adapted to $\mathcal{F}_n$, where X_n is the immediate profit made by exercising the option at time n.
By induction, we define the American option price for $n = 1, \cdots, N$ by

$$\begin{cases} V_N & := & X_N \\ V_{n-1} & := & \max\Big(X_{n-1}, B_{n-1}E_Q[V_n/B_n|\mathcal{F}_{n-1}]\Big) \end{cases}$$

Let $\tilde{V}_n = V_n/B_n$ be the discounted price of the American option. The market is assumed to be viable and complete and we denote Q the unique probability measure under which the discounted price of financial assets are martingales.

We should note that, as opposed to the European case, the discounted price of the American option is generally not a martingale under risk-neutral measure.

Theorem 3.4.1. The sequence $\{\tilde{V}_n\}_{0\le n\le N}$ is a Q-supermartingale. It is the smallest Q-supermartingale that dominates the sequence $\{\tilde{X}_n\}_{0\le n\le N}$.

Proof. From the definition

$$\tilde{V}_{n-1} = \max\Big(\tilde{X}_{n-1}, E_Q[\tilde{V}_n|\mathcal{F}_{n-1}]\Big)$$

we have $\tilde{V}_n \ge \tilde{X}_n$ and $\tilde{V}_n \ge E_Q[\tilde{V}_n|\mathcal{F}_{n-1}]$. Hence $\tilde{V}_n$ dominates $\tilde{X}_n$ and is a super-martingale.
Let us now consider a super-martingale $\{\tilde{T}_n\}_{0\le n\le N}$ that dominates $\{\tilde{X}_n\}_{0\le n\le N}$. Then

$$\tilde{T}_N \ge \tilde{X}_N = \tilde{V}_N$$

and if $\tilde{T}_n \ge \tilde{V}_n$ we have

$$\tilde{T}_{n-1} \quad \ge \quad E_Q[\tilde{T}_n|\mathcal{F}_{n-1}] \ge E_Q[\tilde{V}_n|\mathcal{F}_{n-1}],$$

and

$$\tilde{T}_{n-1} \quad \ge \quad \max\Big(X_{n-1}, B_{n-1}E_Q[V_n/B_n|\mathcal{F}_{n-1}]\Big) = \tilde{V}_{n-1}.$$

A backward induction proves the assertion that $(\tilde{T}_n)$ dominates $(\tilde{V}_n)$. ■

3.4.1 The Snell Envelope

We consider an adapted sequence $\{Z_n\}_{0\le n\le N}$, and define the sequence $\{U_n\}_{0\le n\le N}$ as follows:

$$\begin{cases} U_N & := & Z_N \\ U_{n-1} & := & \max\Big(Z_{n-1}, E_Q[U_n|\mathcal{F}_{n-1}]\Big), \quad \forall n \le N-1. \end{cases}$$

U_n is called the Snell envelope of Z_n. U_n is a supermartingale by Theorem (3.4.1) and dominates Z_n. It is possible to obtain a martingale process by stopping the sequence (U_n) when $U_n > E[U_{n+1}|\mathcal{F}]$ (i.e. when $U_n = Z_n$).

Theorem 3.4.2. The random variable defined by

$$\tau^* \quad = \quad \inf\{n \geq 0 : U_n = Z_n\}$$

is a stopping time and the stopped sequence $\{U_{n\wedge\tau^*}\}_{0\leq n\leq N}$ is a martingale.

Proof. Since $U_N = Z_N$, τ^* is a well-defined element of $\{0, 1, \cdots, N\}$ and we have

$$\{\tau^* = 0\} = \{U_0 = Z_0\} \in \mathcal{F}_0,$$

and for $k \geq 1$,

$$\{\tau^* = K\} \quad = \quad \left[\Big(\bigcap_{m=0}^{k-1} \{U_m > Z_m\} \Big) \bigcap \{U_k = Z_k\} \right] \in \mathcal{F}_k.$$

We can write that

$$U_{n\wedge\tau^*} - U_{(n-1)\wedge\tau^*} \quad = \quad \mathbf{1}_{\{\tau^*\geq n\}}(U_n - U_{n-1}).$$

By definition of U_n and τ^*, $U_{n-1} = E[U_n|\mathcal{F}_{n-1}]$ on $\{\tau^* \geq n\}$.
(Note that $U_{n-1} = \max\Big(Z_{n-1}, E_Q[U_n|\mathcal{F}_{n-1}]\Big)$ and $U_{n-1} > Z_{n-1}$ on $\{\tau^* \geq n\}$.)

$$\begin{aligned} E[U_{n\wedge\tau^*} - U_{(n-1)\wedge\tau^*}|\mathcal{F}_{n-1}] &= E\big[\mathbf{1}_{\{\tau^*\geq n\}}(U_n - U_{n-1})|\mathcal{F}_{n-1}\big] \\ &= \mathbf{1}_{\{\tau^*\geq n\}} E\big[U_n - U_{n-1}|\mathcal{F}_{n-1}\big] \\ &= 0. \end{aligned}$$

■

Theorem 3.4.3.

$$U_0 = E[Z_{\tau^*}] = \sup_{\substack{0\leq\xi\leq T \\ \xi:\text{stopping time}}} E[Z_\xi].$$

Proof.

$$U_0 \quad = \quad U_{0\wedge\tau^*} = E[U_{N\wedge\tau^*}] = E[U_{\tau^*}] = E[Z_{\tau^*}].$$

Now, let ξ be any stopping time $0 \leq \xi \leq N$. Since $U_{n\wedge\xi}$ is super-martingale,

$$U_0 \quad = \quad U_{0\wedge\xi} \geq E[U_{N\wedge\xi}] = E[U_\xi] \geq E[Z_\xi].$$

Thus we have

$$U_0 \quad \geq \quad \sup_{\substack{0\leq\xi\leq T \\ \xi:\text{stopping time}}} E[Z_\xi].$$

Since τ^* is also a stopping time, we have the equality

$$U_0 \quad = \quad \sup_{\substack{0\leq\xi\leq T \\ \xi:\text{stopping time}}} E[Z_\xi].$$

■

Definition 3.4.4 (Optimal Stopping Time). A stopping time τ is called optimal for the sequence $\{Z_n\}_{0\le n\le N}$ if

$$E[Z_\tau] = \sup_{\substack{0\le\xi\le T\\ \xi:\text{stopping time}}} E[Z_\xi].$$

□

We can see that τ^* is optimal. The following result gives a characterization of optimal stopping times.

Theorem 3.4.5. A stopping time τ is optimal if and only if

$$\begin{cases} Z_\tau = U_\tau & \text{and} \\ (U_{n\wedge\tau})_{0\le n\le N} & \text{is a martingale.} \end{cases}$$

Proof. See Lamberton and Lapeyre(1996) page 20. ■

3.4.2 American & European Option(discrete)

Theorem 3.4.6. Let C_n be the value at time n of an American option described by an adapted sequenced $\{Z_n\}_{0\le n\le N}$ and let c_n be the value at time n of the European option defined by the $\mathcal{F}_N$-measurable random variable $h = Z_N$. Then we have

$$C_n \ge c_n.$$

Moreover, if $c_n \ge Z_n$ for any n, then

$$c_n = C_n, \quad \forall n \in \{0,1,\cdots,N\}.$$

Proof. For the discounted value $(\tilde{C}_n)$ is a supermartingale under risk-neutral measure Q, we have

$$\tilde{C}_n \ \ge\ E_Q[\tilde{C}_N|\mathcal{F}_n] = E_Q[\tilde{c}_N|\mathcal{F}_n] = \tilde{c}_n.$$

Hence $C_n \ge c_n$. If $c_n \ge Z_n$ for any n, c_n is a dominating martingale. By Theorem (3.4.1), $\tilde{C}_n$ is the smallest dominating supermartingale. Hence $\tilde{c}_n \ge \tilde{C}_n$ and $c_n \ge C_n$. Thus

$$C_n = c_n$$

. ■

Example 3.4.7.

1. Now let us check $c_n \ge Z_n = (S_n - K)^+$ when $B_n = (1+r)^n, r \ge 0$.

$$\begin{aligned} \tilde{c}_n &= (1+r)^{-N} E_Q[(S_N - K)^+|\mathcal{F}_n] \\ &\ge E_Q[\tilde{S}_N - K(1+r)^{-N}|\mathcal{F}_n] \\ &= \tilde{S}_n - K(1+r)^{-N}. \end{aligned}$$

   Hence $c_n \ge S_n - K(1+r)^{-(N-n)} \ge S_n - K$. Since $c_n \ge 0$, we also have $c_n \ge (S_n-K)^+ = Z_n$. Therefore European call option price, c_n, equals to the American call option price, C_n.

2. Let $\varphi : \mathbb{R}^+ \to \mathbb{R}$ be a convex function such that $\varphi(S_n) \in L^1$ for $n = 0, 1, \cdots, N$.

$$\begin{aligned}
\tilde{c}_n &= \frac{1}{B_N} E_Q[\varphi(S_N)|\mathcal{F}_n] \\
&\geq \frac{1}{B_N}\varphi\bigg(E_Q[S_N|\mathcal{F}_n]\bigg) \\
&= \frac{1}{B_N}\varphi\bigg(B_N\tilde{S}_n\bigg), \\
c_n &\geq \frac{B_n}{B_N}\varphi\bigg(\frac{B_N}{B_n}\cdot S_n\bigg).
\end{aligned}$$

(a) $r = 0$: Then $c_n \geq \varphi(S_n)$. Hence, when $r = 0$, the American put option price equals to the European put option price.

(b) $\varphi(0) \leq 0, r \geq 0$: First , check that

$$a\varphi(x) \geq \varphi(ax), \quad \text{for } x \geq 0, 0 \leq a \leq 1.$$

Let

$$\begin{aligned}
l_1(y) &= \frac{\varphi(x) - \varphi(0)}{x - 0}\,(y - 0) + \varphi(0), \\
l_2(y) &= \frac{\varphi(x) - 0}{x - 0}\,y.
\end{aligned}$$

Since

$$l_2(y) - l_1(y) \quad = \quad \varphi(0)\left(\frac{y}{x} - 1\right).$$

if $y \leq x$, $l_2(y) \geq l_1(y)$. Thus we get

$$\varphi(y) \leq l_1(y) \leq l_2(y).$$

When we put $y \leftarrow ax$,

$$\varphi(ax) \leq l_1(ax) \leq l_2(ax) = a\varphi(x).$$

From $\frac{B_n}{B_N} \leq 1$, we have

$$c_n \quad \geq \quad \frac{B_n}{B_N}\varphi\bigg(\frac{B_N}{B_n}\cdot S_n\bigg) \geq \varphi(S_n).$$

□

3.4.3 American & European Option(continuous)

Theorem 3.4.8. American call option price, C_0, is the same as the European one, c_0.

Proof. Let S_t be the stock process and B_t risk-less bond price process. Consider call option with strike price $K > 0$ and maturity T. First show that

$$\frac{1}{B_t}$$

is super-martingale. For $s < t$,

$$\begin{aligned} B_s E\left[\frac{1}{B_t}\middle|\mathscr{F}_s\right] &\leq 1, \\ E\left[\frac{1}{B_t}\middle|\mathscr{F}_s\right] &\leq \frac{1}{B_s}. \end{aligned}$$

Since $\frac{S_t}{B_t}$ is a martingale, $-\frac{K}{B_t}$ is sub-martingale and $x \mapsto x^+$ is an increasing convex function, we have

$$\left(\frac{S_t}{B_t} - \frac{K}{B_t}\right)^+$$

is sub-martingale. By the optional sampling theorem, we have

$$E\left[\left(\frac{S_\tau}{B_\tau} - \frac{K}{B_\tau}\right)^+\right] \leq E\left[\left(\frac{S_T}{B_T} - \frac{K}{B_T}\right)^+\right].$$

Then we get

$$C_0 = \sup_{0\leq\tau\leq T} E\left[\left(\frac{S_\tau}{B_\tau} - \frac{K}{B_\tau}\right)^+\right] = E\left[\left(\frac{S_T}{B_T} - \frac{K}{B_T}\right)^+\right] = c_0.$$

■

3.5 Obstacle Problems

This section follows Seydel(2002) and Wilmott(1993).

✎ **Linear Complementarity Problem**

We assume an "obstacle" $f(x)$ with

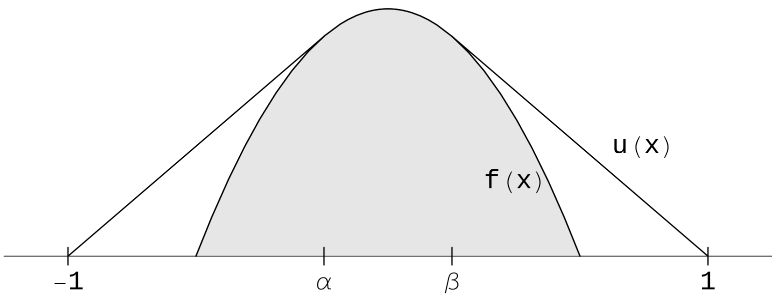


Figure 3.2: Example of an obstacle problem.

$$\begin{aligned}
&f(x) > 0, \quad \text{for } \alpha < x < \beta, \\
&f \in C^2, \quad f'' < 0, \quad f(-1) < 0, \quad f(1) < 0.
\end{aligned}$$

Consider an elastic string held fixed at two ends, $x = \pm 1$, and constrained to lie above the obstacle $f(x)$ as in Figure 3.2. Let $u(x)$ be the displacement of the string. For α and β the curve of u touches the obstacle tangentially. These two values $x = \alpha$ and $x = \beta$ are unknown initially. This obstacle problem is a simple free boundary-value problem.

Our ami is to reformulate the obstacle problem such that the free boundary conditions do not show up explicitly. This may promise computational advantages. The function u is defined by

$$\begin{aligned}
u'' &= 0, \quad \text{for } -1 < x < \alpha, \\
u &= f, \quad \text{for } \alpha < x < \beta, \\
u'' &= 0, \quad \text{for } \beta < x < 1.
\end{aligned}$$

This manifests a complementarity in the sense

$$\begin{aligned}
u'' &= 0, \quad \text{if } u > f, \\
u'' &< 0, \quad \text{if } u = f.
\end{aligned}$$

It is clear that American options are complementary in an analogous way

$$\begin{aligned}
\frac{\partial V}{\partial t} + (r-q)S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV &= 0, \quad \text{if } V > \text{ payoff }, \\
\frac{\partial V}{\partial t} + (r-q)S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV &< 0, \quad \text{if } V = \text{ payoff }.
\end{aligned}$$

The obstacle problem can be reformulated as:

$$\text{Find} \qquad u(x) \in C^1[-1,1]$$

subject to

$$\begin{cases} u''(u-f) = 0, \\ -u'' \geq 0, \quad u - f \geq 0, \\ u(-1) = u(1) = 0. \end{cases} \tag{3.5.1}$$

This formulation does not mention the free boundary explicitly. This will be advantageous because α and β are unknown. If a solution is known, then α and β are read off from the solution.

✎ **Variational Inequality**

Let $\mathcal{K}$ denote the set of functions such that

$$\mathcal{K} = \{v \in C^0[-1,1] : v(-1) = v(1) = 0, v(x) \geq f(x), v \text{ piecewise } C^1\}.$$

We call any function $v(x) \in \mathcal{K}$ a test function and the set $\mathcal{K}$ is called a space of test functions.

For any $v(x) \in \mathcal{K}$ we have $v - f \geq 0$ and because $-u'' \geq 0$

$$-u''(v-f) \geq 0,$$

which gives

$$\int_{-1}^{1} -u''(v-f)dx \geq 0.$$

We also have

$$\int_{-1}^{1} -u''(u-f)dx = 0.$$

Subtracting yields

$$\int_{-1}^{1} -u''(v-u)dx \geq 0 \text{ for any } v \in \mathcal{K}.$$

The obstacle function f does not occur explicitly in this formulation. The obstacle is implicitly defined in $\mathcal{K}$. Integration by parts leads to

$$[-u'(v-u)]_{-1}^{1} + \int_{-1}^{1} u'(v-u)'dx \geq 0,$$

thus we have the variational inequality formulation of free boundary problem, namely

$$\text{Find } u \in \mathcal{K} \text{ such that } \int_{-1}^{1} u'(v-u)'dx \geq 0 \text{ for any } v \in \mathcal{K}.$$

In order to solve the variational inequality formulation we can use the finite element method.

✎ **Discretization of the Obstacle Problem** For a numerical solution, let

$$\Delta x = \frac{2}{M},$$
$$x_i = -1 + i\Delta x, \quad f_i = f(x_i).$$

A finite-difference approximation for (3.5.1) leads to

$$\begin{cases} (w_{i-1} - 2w_i + w_{i+1})(w_i - f_i) = 0, \\ -w_{i-1} + 2w_i - w_{i+1} \geq 0, \quad w_i \geq f_i, \quad 0 < i < M, \end{cases} \tag{3.5.2}$$

where $w_0 = w_M = 0$ and w_i are approximations to $u(x_i)$.

Let

$$\mathbf{B} \quad = \quad \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_{M-1} \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} f_1 \\ \vdots \\ f_{M-1} \end{pmatrix}.$$

To Calculate (3.5.2), one solves $\mathbf{Bw} = \mathbf{0}$ under the side condition $\mathbf{w} \geq \mathbf{f}$.

3.6 Complementarity Problem For The American Put

By means of the transformation

$$\begin{aligned} S &= Ke^y, \quad t = T - \frac{2s}{\sigma^2}, \quad 0 \le t \le T, 0 \le s \le \frac{\sigma^2}{2}T, \\ C(t,S) &= K\exp\left\{-\frac{1}{2}(k_2-1)y - (\frac{1}{4}(k_2-1)^2 + k_1)s\right\}u(s,y). \end{aligned}$$

where

$$k_1 = \frac{2r}{\sigma^2}, \quad k_2 = \frac{2(r-D_0)}{\sigma^2},$$

the Black-Scholes equation

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + (r-D_0)S\frac{\partial C}{\partial S} - rC = 0$$

for any European option paying a constant dividend yield can be transformed into the diffusion equation

$$u_s = u_{yy}.$$

For the American Put, we have

$$\begin{aligned} u_s &= u_{yy}, \\ u(0,y) &= \left(e^{\frac{1}{2}(k_2-1)y} - e^{\frac{1}{2}(k_2+1)y}\right)^+, \\ u(s,y) &\ge e^{\frac{1}{4}\left((k_2-1)^2+4k_1\right)s}\left(e^{\frac{1}{2}(k_2-1)y} - e^{\frac{1}{2}(k_2+1)y}\right)^+, \\ \lim_{y\to\infty} u(s,y) &= 0. \end{aligned}$$

For the call,

$$\begin{aligned} u_s &= u_{yy}, \\ u(0,y) &= \left(e^{\frac{1}{2}(k_2+1)y} - e^{\frac{1}{2}(k_2-1)y}\right)^+, \\ u(s,y) &\ge e^{\frac{1}{4}\left((k_2-1)^2+4k_1\right)s}\left(e^{\frac{1}{2}(k_2+1)y} - e^{\frac{1}{2}(k_2-1)y}\right)^+, \\ \lim_{y\to-\infty} u(s,y) &= 0. \end{aligned}$$

We can write all of these option valuation problems in the more compact linear complementarity form

$$\left\{ \begin{array}{l} u_s - u_{yy} \ge 0, \quad u(s,y) - g(s,y) \ge 0, \\ (u_s - u_{yy})(u-g) = 0. \\ \text{IC: } u(0,y) = g(0,y) \\ \text{BC: } u(s,-\infty) = g(s,-\infty), u(s,\infty) = g(s,\infty). \end{array} \right.$$

Here the transformed payoff constraint functions, $g(s,y)$ is given by

$$g(s,y) = e^{\frac{1}{4}\left((k_2-1)^2+4k_1\right)s}\left(e^{\frac{1}{2}(k_2-1)y} - e^{\frac{1}{2}(k_2+1)y}\right)^+$$

for the put.

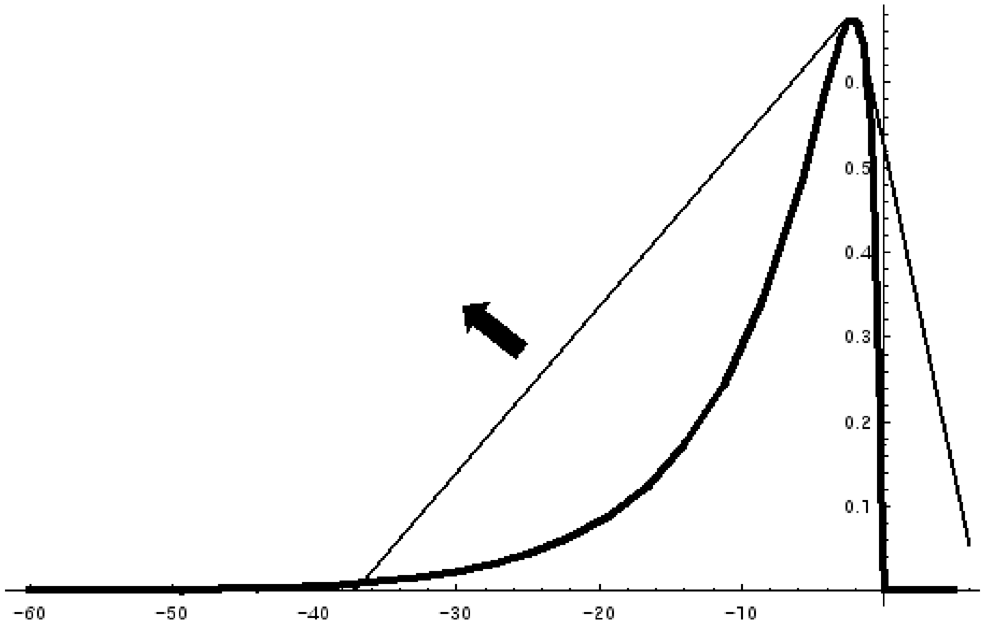


Figure 3.3: Transformed put option's payoff(obstacle)

3.7 FDM for American Option

Finite difference methods(FDM) are no more than a generalization of the binomial method, although we tend to talk about grids and meshes rather than trees. There are many, many ways the FDM can be improved upon, making it faster and more accurate. The binomial method is not so flexible. The main difference between the binomial method and the FDM is that the former contains the diffusion, the volatility, in the tree structure. In the FDM the 'tree' is fixed but parameters change to reflect a changing diffusion [2].

3.7.1 Early Exercise and the Explicit Method

Suppose that we have found v_i^k for all i at time step k, time step to find the option value at $k+1$ using

$$v_i^{k+1} \quad = \quad a_i^k v_{i-1}^k + b_i^k v_i^k + c_i^k v_{i+1}^k.$$

After we have found the option value v_i^{k+1} for all i, check whether the new option values are greater or less than the payoff. If they are less than the payoff then we have arbitrage. Hence we have to replace that value by the payoff at that asset value.

$$v_i^{k+1} \quad \leftarrow \quad \max(v_i^{k+1},\ \text{payoff }).$$

It is clear that this finite-difference solution will converge to something that has a value continuous with the payoff. What is not so clear is that the gradient constraint of continuous delta is also

[2]The main reason that I rarely do any transforming of the equation when I am solving it numerically is that I like to solve in terms of the real financial variables since terms of the contract are specified using these real variables. See Wilmott(1999) page 618.

satisfied.

3.7.2 Early Exercise and Crank-Nicolson

Implementing the American contingent in the Crank-Nicolson method is a bit harder than in the explicit method, but the reward come in the accuracy($O(\Delta t^2, \Delta S^2)$).

The only complication arises because the Crank-Nicolson method is implicit, and every value of the option at the $k+1$ time step is linked to every other value at the time step. It is therefore not good enough to just replace the option value with the payoff after the values have all been calculated[3], the replacement must be done at the same time as the values are found:

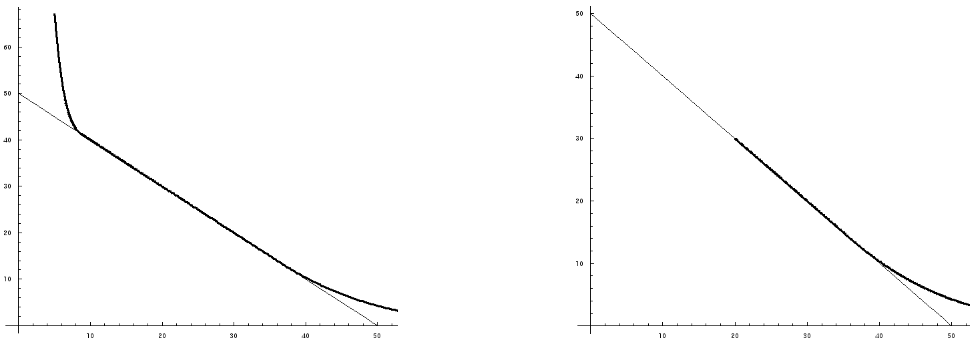
$$v_i^{k+1} \quad \leftarrow \quad \max(v_i^k + \omega(u_i^{k+1} - v_i^k),\ \text{payoff }).$$

This method is called **projected SOR**.

3.7.3 The Projected SOR

$$A\mathbf{x} \geq \mathbf{b}, \quad \mathbf{x} \geq \mathbf{g}, \quad (A\mathbf{x} - \mathbf{b})(\mathbf{x} - \mathbf{g}) = 0.$$

Remark 3.7.1. $g(s, y)$의 값이 0보다 크고, $\lim_{y \to -\infty} g(s, y) = 0$이므로 $g[0]$의 값을 잘못 설정하면 잘못된 결과를 얻을 수 있다. obstacle이 $-\infty$일 때, 0이므로, 주가가 0에 가까이 갈수록 option value도 점점 커지게 된다.



(a) g(0)을 잘못 설정한 경우. (b) g(0)=0으로 설정한 경우.

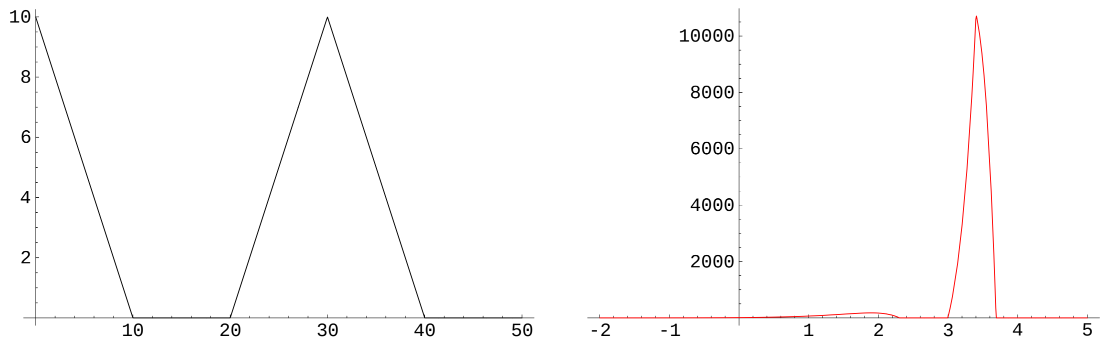
Figure 3.4: American option pricing.

[3]This can be done but the accuracy is then reduced to $O(\Delta t)$.

3.7.4 Implementation Example(Bermudan Style Option)

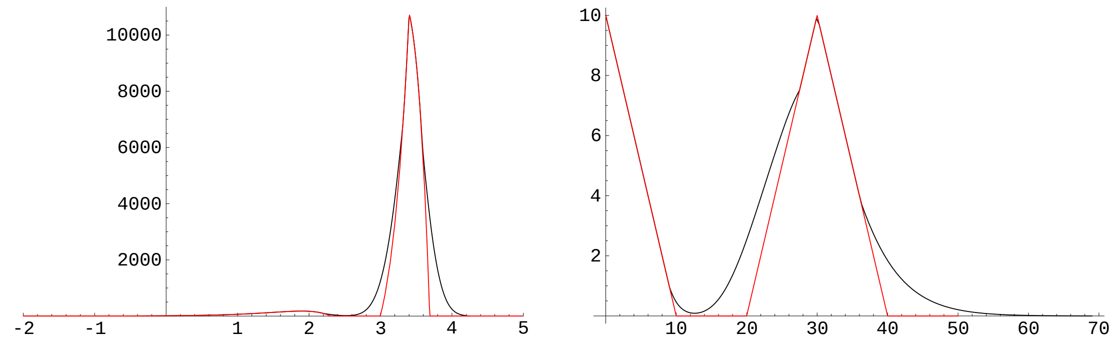
Maturity	1 year
Payoff	Figure 3.5(a)
Exercisable times	$0, 0.2, 0.4, 0.6, 0.8, 1$

Table 3.1: Bermudan Option



(a) Payoff (b) Transformed payoff.

Figure 3.5: Bermudan option pricing.



(a) Before transformation (b) Value function.

Figure 3.6: The price of Bermudan option($r = 10\%, \sigma = 20\%$).

3.8 Analytic Approximation to American Option Prices

3.8.1 Barone, Adesi and Whaley's Approach

We will denote the difference between the American and European option price by ν. Then

$$\nu_t + (r-q)S\nu_S + \frac{1}{2}\sigma^2 S^2 \nu_{SS} = r\nu, \quad \text{Call: } S \le S^*, \text{ put: } S \ge S^{**}.$$

For convenience, we define

$$\begin{aligned} \tau &= T - t \\ h(\tau) &= 1 - e^{-r\tau} \\ \alpha &= \frac{2r}{\sigma^2} \\ \beta &= \frac{2(r-q)}{\sigma^2}. \end{aligned}$$

We also write,

$$\nu = h(\tau) g(S, h).$$

This gives

$$S^2 g_{SS} + \beta S g_S - \frac{\alpha}{h} g - (1-h)\alpha g_h = 0.$$

Up to this point, the analysis has been exact, and no approximation has been made. The approximation involves assuming that the final term on the left-hand side is zero, so that

$$S^2 g_{SS} + \beta S g_S - \frac{\alpha}{h} g = 0. \tag{3.8.1}$$

The ignored term is generally fairly small. When τ is large, $1-h$ is close to zero; when τ is small, g_h is close to zero. Equation (3.8.1) is a second-order ordinary differential equation with two linearly independent solution of the form aS^γ. The solutions of Equation (3.8.1) are

$$\begin{aligned} \gamma_1 &= \left[-(\beta-1) - \sqrt{(\beta-1)^2 + \frac{4\alpha}{h}} \right] \Big/ 2 \\ \gamma_2 &= \left[-(\beta-1) + \sqrt{(\beta-1)^2 + \frac{4\alpha}{h}} \right] \Big/ 2 \\ A_1 &= -\left(\frac{S^{**}}{\gamma_1} \right) \left\{ 1 - e^{-q(T-t)} N[-d_1(S^{**})] \right\} \\ A_2 &= \left(\frac{S^{*}}{\gamma_2} \right) \left\{ 1 - e^{-q(T-t)} N[d_1(S^{*})] \right\} \\ d_1(S) &= \frac{\ln(S/X) + (r - q + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}. \end{aligned}$$

This solutions are obtained from two condition, for example(call option),

$$\begin{aligned} S^* - X &= c(S^*, t) + h a_2 (S^*)^{\gamma_2}, \\ 1 &= e^{-q(T-t)} N[d_1(S^*)] + h\gamma_2 a_2 (S^*)^{\gamma_2 - 1} \end{aligned}$$

1. Call option:

$$C(S,t) = \begin{cases} c(S,t) + A_2\left(\frac{S}{S^*}\right)^{\gamma_2}, & \text{when } S < S^* \\ S - X, & \text{when } S \geq S^*. \end{cases}$$

The S^* is the critical price of the stock above which the option should be exercised. It is estimated by solving the equation

$$S^* - X = c(S^*,t) + \left\{1 - e^{-q(T-t)}N[d_1(S^*)]\right\}\frac{S^*}{\gamma_2}$$

We can find S^* from the Newton-Raphson method. To get S^*, we set

$$\begin{aligned} \text{LHS}(S_i) &= S_i - X, \\ \text{RHS}(S_i) &= c(S_i,t) + \left\{1 - e^{-q(T-t)}N[d_1(S_i)]\right\}\frac{S_i}{\gamma_2}, \\ \frac{\partial \text{RHS}}{\partial S_i} &= b_i = e^{-q(T-t)}N[d_1(S_i)](1 - \frac{1}{\gamma_2}) + \left\{1 - \frac{e^{-q(T-t)}n[d_1(S_i)]}{\sigma\sqrt{T-t}}\right\}\frac{1}{\gamma_2}, \\ S_{i+1} &= S_i - \frac{\text{LHS}(S_i) - \text{RHS}(S_i)}{\text{LHS}'(S_i) - \text{RHS}'(S_i)} = S_i - \frac{S_i - X - \text{RHS}}{1 - b_i} = \frac{X + \text{RHS}(S_i) - b_i S_i}{1 - b_i}. \end{aligned}$$

The iterative procedure should continue until the relative absolute error falls within an acceptable tolerance level. For example,

$$\frac{|\text{LHS}(S_i) - \text{RHS}(S_i)|}{X} < 0.00001.$$

As always with the use of the Newton-Raphson method, we need a seed value. Barone-Adesi and Whaley suggest using

$$S_1^* = X + \left(S^*(\infty) - X\right)\left(1 - e^{h_2}\right), \quad h_2 = -\left((r-q)(T-t) + 2\sigma\sqrt{T-t}\right)\left[\frac{X}{S^*(\infty) - X}\right]$$

$$S^*(\infty) = \frac{X}{1 - \frac{2}{-(\beta-1)+\sqrt{(\beta-1)^2+4\alpha}}}$$

2. Put option:

$$P(S,t) = \begin{cases} p(S,t) + A_1\left(\frac{S}{S^{**}}\right)^{\gamma_1}, & \text{when } S > S^{**} \\ X - S, & \text{when } S \leq S^{**}. \end{cases}$$

The S^{**} is the critical price of the stock above which the option should be exercised. It is estimated by solving the equation

$$X - S^{**} = p(S^{**},t) - \left\{1 - e^{-q(T-t)}N[-d_1(S^{**})]\right\}\frac{S^{**}}{\gamma_1}$$

We can find S^{**} from the Newton-Raphson method. To get S^{**}, we set

$$\text{LHS}(S_i) = X - S_i,$$

$$
\begin{aligned}
\mathrm{RHS}(S_i) &= p(S_i,t) - \left\{1 - e^{-q(T-t)}N[d_1(S_i)]\right\}\frac{S_i}{\gamma_1}, \\
\frac{\partial \mathrm{RHS}}{\partial S_i} &= b_i = -e^{-q(T-t)}N[-d_1(S_i)](1-\frac{1}{\gamma_1}) - \left\{1 + \frac{e^{-q(T-t)}n[-d_1(S_i)]}{\sigma\sqrt{T-t}}\right\}\frac{1}{\gamma_1}, \\
S_{i+1} &= S_i - \frac{\mathrm{RHS}(S_i) - \mathrm{LHS}(S_i)}{\mathrm{RHS}'(S_i) - \mathrm{LHS}'(S_i)} = S_i - \frac{\mathrm{RHS} - X + S_i}{1+b_i} = \frac{X - \mathrm{RHS}(S_i) + b_i S_i}{1+b_i}.
\end{aligned}
$$

The iterative procedure should continue until the relative absolute error falls within an acceptable tolerance level. For example,

$$
\frac{|\mathrm{LHS}(S_i) - \mathrm{RHS}(S_i)|}{X} < 0.00001.
$$

As always with the use of the Newton-Raphson method, we need a seed value. Barone-Adesi and Whaley suggest using

$$
\begin{aligned}
&S_1^{**} = S^{**}(\infty) + \Big(X - S^{**}(\infty)\Big)e^{h_1}, \quad h_2 = \Big((r-q)(T-t) - 2\sigma\sqrt{T-t}\Big)\left[\frac{X}{X - S^{**}(\infty)}\right] \\
&S^{**}(\infty) = \frac{X}{1 - \frac{2}{-(\beta-1)-\sqrt{(\beta-1)^2+4\alpha}}}
\end{aligned}
$$

Chapter 4

Volatility

4.1 Time Dependent Volatility

Let us assume that volatility is a function of time.

$$dS_t = S_t\Big(rdt + \sigma(t)dW_t\Big).$$

The Black-Scholes formulae are still valid when volatility is time dependent provided we use

$$\sqrt{\frac{1}{T-t}\int_t^T \sigma^2(s)ds}$$

in place of σ, i.e. now use

$$d_1 = \frac{\log(S/K) + r(T-t) + \frac{1}{2}\int_t^T \sigma^2(s)ds}{\sqrt{\int_t^T \sigma^2(s)ds}}$$

The implied volatility $\sigma_{\text{imp}}(t,T)$ is defined by

$$\sigma^2_{\text{imp}}(t,T) = \frac{1}{T-t}\int_t^T \sigma^2(s)ds. \tag{4.1.1}$$

Example 4.1.1.

$$S_t = S_0 \exp\Big(rt - \frac{1}{2}\int_0^t \sigma^2(s)ds + \int_0^t \sigma(s)dW_s\Big)$$

Let us calculate the correlation between

$$\int_0^{T_1} \sigma(s)dW_s \quad \text{and} \quad \int_0^{T_2} \sigma(s)dW_s \quad (T_1 < T_2).$$

The covariance between these is given by

$$\text{Cov}\left(\int_0^{T_1} \sigma(s)dW_s\ ,\ \int_0^{T_2} \sigma(s)dW_s\right) = E\left[\Big(\int_0^{T_1} \sigma(s)dW_s\Big)\Big(\int_0^{T_2} \sigma(s)dW_s\Big)\right]$$

$$
\begin{aligned}
&= E\left[\Big(\int_0^{T_1}\sigma(s)dW_s\Big)\Big(\int_0^{T_1}\sigma(s)dW_s+\int_{T_1}^{T_2}\sigma(s)dW_s\Big)\right] \\
&= E\left[\Big(\int_0^{T_1}\sigma(s)dW_s\Big)^2\right] \\
&= \int_0^{T_1}\sigma(s)^2ds, \quad \text{by Ito isometry} \\
&= \sigma^2_{\text{imp}}(0,T_1)\ T_1
\end{aligned}
$$

and the correlation coefficient, ρ, is given by

$$
\begin{aligned}
\rho &= \frac{\sigma^2_{\text{imp}}(0,T_1)\ T_1}{\sigma_{\text{imp}}(0,T_1)\sqrt{T_1}\ \sigma_{\text{imp}}(0,T_2)\ \sqrt{T_2}} \\
&= \frac{\sigma_{\text{imp}}(0,T_1)\ \sqrt{T_1}}{\sigma_{\text{imp}}(0,T_2)\ \sqrt{T_2}}.
\end{aligned}
$$

□

✎ **Volatility from Implied Volatility**
From Eq(4.1.1), we have

$$
\int_t^T \sigma^2(s)ds = (T-t)\sigma^2_{\text{imp}}(t,T).
$$

Differentiation with respect to T implies

$$
\sigma(T) = \sqrt{\sigma^2_{\text{imp}}(t,T)+2(T-t)\sigma_{\text{imp}}(t,T)\frac{\partial\sigma_{\text{imp}}(t,T)}{\partial T}}. \tag{4.1.2}
$$

Practically speaking, we do not have a continuous implied volatility curve. We have only a discrete set of points. We must therefore make some assumption about the term structure of volatility between the date points. Usually one assumes that the function is piecewise constant or linear.

For example, if we have implied volatility for expiries T_i and assume that volatility curve(not implied volatility curve) to be piecewise constant then

$$
\sigma(T) = \sqrt{\frac{(T_i-t)\sigma^2_{\text{imp}}(t,T_i)-(T_{i-1}-t)\sigma^2_{\text{imp}}(t,T_{i-1})}{T_i-T_{i-1}}}, \quad T_{i-1}<T<T_i. \tag{4.1.3}
$$

Example 4.1.2. Suppose that implied volatilities are shown in Table(4.1) and that the volatility curve is piecewise constant. From (4.1.3) we have a piecewise linear volatility curve as in Figure(4.1). Also we have an estimated implied volatility curve by (4.1.1). We can see that the observed discrete set of implied volatility points are on the estimated implied volatility curve. □

Table 4.1: Implied volatilities of options.

Expiry	Implied Volatility
0.5 year	25%
1 year	30%
2 year	38%
3 year	43%

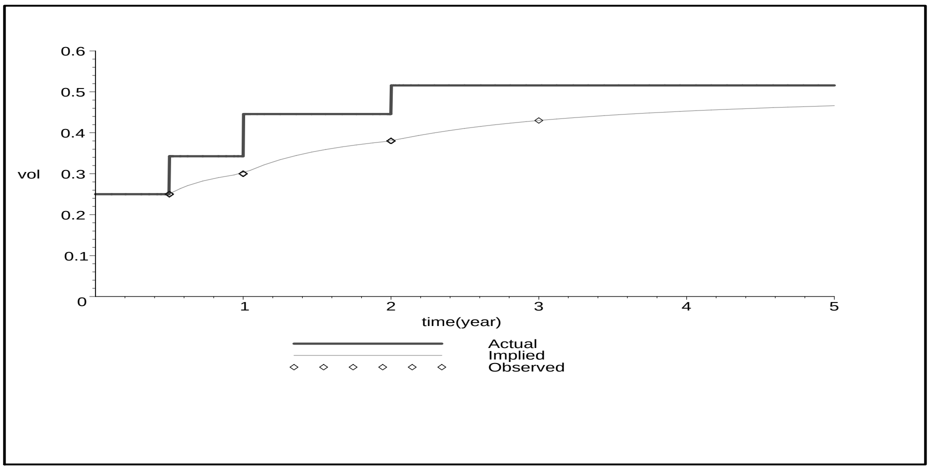


Figure 4.1: Volatility term structure

4.2 Volatility Surface

We will assume that the underlying asset process S_t satisfies

$$dS_t \quad = \quad S_t\Big((r-D)dt + \sigma(S,t)dW_t\Big).$$

In the following, we are going to get $\sigma(S,t)$, precisely $\sigma(K,T)$, from call option price $C(K,T)$, where K is the strike price and T is the option maturity.

1. First, let us computer transition probability density function $p(S_t,t;S_T,T)\Big(=p(S_t,t;S,T)\Big)$. The call option value can be written as

$$\begin{aligned} V(K,T) &= e^{-r(T-t)}\int_K^\infty (S_T-K)\,p(S_t,t;S_T,T)dS_T \\ &= e^{-r(T-t)}\int_K^\infty (S-K)\,p(S_t,t;S,T)dS. \end{aligned} \tag{4.2.1}$$

   After differentiating with respect to K, we get

$$\frac{\partial V}{\partial K} \quad = \quad -e^{-r(T-t)}\int_K^\infty p(S_t,t;S,T)dS. \tag{4.2.2}$$

   After one more differentiation, we arrive at

$$\frac{\partial^2 V}{\partial K^2} \quad = \quad e^{-r(T-t)}p(S_t,t;K,T), \quad \text{i.e. } p(S_t,t;K,T) \quad = \quad e^{r(T-t)}\frac{\partial^2 V}{\partial K^2}. \tag{4.2.3}$$

2. The Fokker-Planck equation(Kolmogorov forward equation) to the transition probability function is given by $\frac{\partial p}{\partial T} = A^* p$, i.e.

$$\frac{\partial p}{\partial T} \quad = \quad \frac{1}{2}\frac{\partial^2}{\partial S^2}\Big(\sigma^2(S,T)S^2 p\Big) - \frac{\partial}{\partial S}\Big((r-D)Sp\Big). \tag{4.2.4}$$

   Note that in this equation, $\sigma(S,T)$ is evaluated at $t=T$.

3. From Eq(4.2.1) and Eq(4.2.4) we have

$$\begin{aligned} \frac{\partial V}{\partial T} &= -rV + e^{-r(T-t)}\int_K^\infty (S-K)\frac{\partial p}{\partial T}dS \\ &= -rV + e^{-r(T-t)}\int_K^\infty (S-K)\left\{\frac{1}{2}\frac{\partial^2\Big(\sigma(S,T)Sp\Big)}{\partial S^2} - \frac{\partial\Big((r-D)Sp\Big)}{\partial S}\right\}dS \\ &= -rV + \frac{1}{2}e^{-r(T-t)}\sigma^2(K,T)K^2 p(S_t,t;K,T) + (r-D)e^{-r(T-t)}\int_K^\infty Sp\,dS \\ &= -rV + \frac{1}{2}\sigma^2(K,T)K^2\frac{\partial^2 V}{\partial K^2} + (r-D)e^{-r(T-t)}\int_K^\infty (S-K)p\,dS + (r-D)Ke^{-r(T-t)}\int_K^\infty p\,dS \\ &= -rV + \frac{1}{2}\sigma^2(K,T)K^2\frac{\partial^2 V}{\partial K^2} + (r-D)V - (r-D)K\frac{\partial V}{\partial K} \\ &= \frac{1}{2}\sigma^2(K,T)K^2\frac{\partial^2 V}{\partial K^2} - DV - (r-D)K\frac{\partial V}{\partial K}. \end{aligned}$$

4. Rearranging this we find that

$$\sigma^2(K,T) = \frac{\frac{\partial V}{\partial T} + (r-D)K\frac{\partial V}{\partial K} + DV}{\frac{1}{2}K^2\frac{\partial^2 V}{\partial K^2}}. \tag{4.2.5}$$

This $\sigma(K,T)$ is called the **local volatility surface**. Also Eq(4.2.5) is termed as Dupire equation.

4.2.1 Local volatility with Black-Scholes Formula

In this subsection, we are going to derive the local volatility formula for the given implied volatilities which are computed from the Black-Scholes formula.

For the given $\sigma_{\text{imp}}(K,T)$, the Black-Scholes formula is given by

$$\begin{aligned} V(K,T) &= Se^{-D(T-t)}N\left(\frac{\log\frac{S}{K} + (r-D+\frac{1}{2}\sigma^2_{\text{imp}}(K,T))(T-t)}{\sigma_{\text{imp}}(K,T)\sqrt{T-t}}\right) \\ &\quad -Ke^{-r(T-t)}N\left(\frac{\log\frac{S}{K} + (r-D-\frac{1}{2}\sigma^2_{\text{imp}}(K,T))(T-t)}{\sigma_{\text{imp}}(K,T)\sqrt{T-t}}\right). \end{aligned}$$

Let us define new variables w, y and τ as follows.

$$\begin{cases} w &:= \sigma^2_{\text{imp}}(T-t), \\ y &:= \log\frac{K}{Se^{(r-D)(T-t)}}, \\ \tau &:= T-t. \end{cases}$$

From these new variables we have

$$V(w,y,\tau) = Se^{-D\tau}\left\{N\left(-\frac{y}{\sqrt{w}} + \frac{\sqrt{w}}{2}\right) - e^y N\left(-\frac{y}{\sqrt{w}} - \frac{\sqrt{w}}{2}\right)\right\}.$$

First, let us computer the numerator of Eq(4.2.5).

1. $\frac{\partial V}{\partial T}$:

$$\begin{aligned} \frac{\partial V}{\partial T} &= \frac{\partial V}{\partial \tau}\frac{\partial \tau}{\partial T} + \frac{\partial V}{\partial w}\frac{\partial w}{\partial T} + \frac{\partial V}{\partial y}\frac{\partial y}{\partial T} \\ &= -DV + \frac{\partial V}{\partial w}\frac{\partial w}{\partial T} + \frac{\partial V}{\partial y}\big(-(r-D)\big). \end{aligned}$$

2. $(r-D)K\frac{\partial V}{\partial K}$:

$$\begin{aligned} (r-D)K\frac{\partial V}{\partial K} &= (r-D)K\left(\frac{\partial V}{\partial y}\frac{\partial y}{\partial K} + \frac{\partial V}{\partial w}\frac{\partial w}{\partial K}\right) \\ &= (r-D)K\left(\frac{\partial V}{\partial y}\frac{\partial y}{\partial K} + \frac{\partial V}{\partial w}\frac{\partial w}{\partial K}\right) \\ &= (r-D)\frac{\partial V}{\partial y} + (r-D)K\frac{\partial V}{\partial w}\frac{\partial w}{\partial K}. \end{aligned}$$

3. $\frac{\partial V}{\partial T} + (r-D)K\frac{\partial V}{\partial K} + DV$:

$$
\begin{aligned}
& \frac{\partial V}{\partial T} + (r-D)K\frac{\partial V}{\partial K} + DV \\
= \;& \frac{\partial V}{\partial w}\frac{\partial w}{\partial T} + (r-D)K\frac{\partial V}{\partial w}\frac{\partial w}{\partial K} \\
= \;& \frac{\partial V}{\partial w}\left(\frac{\partial w}{\partial T} + (r-D)K\frac{\partial w}{\partial K}\right) \\
= \;& \frac{\partial V}{\partial w}\left(\sigma_{\text{imp}}^2 + 2(T-t)\sigma_{\text{imp}}\frac{\partial \sigma_{\text{imp}}}{\partial T} + 2(r-D)K(T-t)\sigma_{\text{imp}}\frac{\partial \sigma_{\text{imp}}}{\partial K}\right).
\end{aligned}
$$

Second, let us computer the denominator of Eq(4.2.5).

1. $\frac{\partial V}{\partial K}$:

$$
\begin{aligned}
\frac{\partial V}{\partial K} &= \frac{\partial V}{\partial y}\frac{\partial y}{\partial K} + \frac{\partial V}{\partial w}\frac{\partial w}{\partial K} \\
&= \frac{\partial V}{\partial y}\frac{1}{K} + \frac{\partial V}{\partial w}\frac{\partial w}{\partial K}.
\end{aligned}
$$

2. $\frac{1}{2}K^2\frac{\partial^2 V}{\partial K^2}$:

$$
\begin{aligned}
\frac{\partial^2 V}{\partial K^2} &= \frac{\partial}{\partial K}\left(\frac{\partial V}{\partial y}\right)\frac{1}{K} - \frac{\partial V}{\partial y}\frac{1}{K^2} + \frac{\partial}{\partial K}\left(\frac{\partial V}{\partial w}\right)\frac{\partial w}{\partial K} + \frac{\partial V}{\partial w}\frac{\partial^2 w}{\partial K^2} \\
&= \frac{\partial^2 V}{\partial y^2}\frac{1}{K^2} + \frac{\partial^2 V}{\partial w \partial y}\frac{\partial w}{\partial K}\frac{1}{K} - \frac{\partial V}{\partial y}\frac{1}{K^2} + \frac{\partial^2 V}{\partial w^2}\left(\frac{\partial w}{\partial K}\right)^2 + \frac{\partial^2 V}{\partial y \partial w}\frac{\partial y}{\partial K}\frac{\partial w}{\partial k} + \frac{\partial V}{\partial w}\frac{\partial^2 w}{\partial K^2} \\
&= \frac{\partial^2 V}{\partial y^2}\frac{1}{K^2} + 2\frac{\partial^2 V}{\partial w \partial y}\frac{\partial w}{\partial K}\frac{1}{K} - \frac{\partial V}{\partial y}\frac{1}{K^2} + \frac{\partial^2 V}{\partial w^2}\left(\frac{\partial w}{\partial K}\right)^2 + \frac{\partial V}{\partial w}\frac{\partial^2 w}{\partial K^2}, \\
\frac{1}{2}K^2\frac{\partial^2 V}{\partial K^2} &= \frac{1}{2}\left(\frac{\partial^2 V}{\partial y^2} - \frac{\partial V}{\partial y}\right) + K\frac{\partial^2 V}{\partial w \partial y}\frac{\partial w}{\partial K} + \frac{1}{2}K^2\frac{\partial^2 V}{\partial w^2}\left(\frac{\partial w}{\partial K}\right)^2 + \frac{1}{2}K^2\frac{\partial V}{\partial w}\frac{\partial^2 w}{\partial K^2}.
\end{aligned}
$$

After tedious calculation, we get the following relation:

$$
\begin{aligned}
\frac{\partial V}{\partial w} &= \frac{1}{2\sqrt{2\pi}}\,\frac{Se^{-D(T-t)}\,e^{-\frac{1}{8}\left(w-4y+\frac{4y^2}{w}\right)}}{\sqrt{w}}, \\
\frac{\partial^2 V}{\partial w^2} &= \frac{1}{16\sqrt{2\pi}}\,\frac{Se^{-D(T-t)}\,e^{-\frac{1}{8}\left(w-4y+\frac{4y^2}{w}\right)}\,(4y^2-4w-w^2)}{\sqrt{w}\,w^2} = \frac{4y^2-4w-w^2}{8w^2}\,\frac{\partial V}{\partial w}, \\
\frac{\partial^2 V}{\partial y^2} - \frac{\partial V}{\partial y} &= \frac{1}{\sqrt{2\pi}}\,\frac{Se^{-D(T-t)}\,e^{-\frac{1}{8}\left(w-4y+\frac{4y^2}{w}\right)}}{\sqrt{w}} = 2\,\frac{\partial V}{\partial w}, \\
\frac{\partial^2 V}{\partial w \partial y} &= \frac{1}{4\sqrt{2\pi}}\,\frac{Se^{-D(T-t)}\,e^{-\frac{1}{8}\left(w-4y+\frac{4y^2}{w}\right)}\,(w-2y)}{\sqrt{w}\,w} = \frac{w-2y}{2w}\,\frac{\partial V}{\partial w}.
\end{aligned}
$$

Thus we have

$$
\frac{1}{2}K^2\frac{\partial^2 V}{\partial K^2} = \frac{\partial V}{\partial w} + K\frac{w-2y}{2w}\,\frac{\partial w}{\partial K}\frac{\partial V}{\partial w} + \frac{1}{2}K^2\,\frac{4y^2-4w-w^2}{8w^2}\,\left(\frac{\partial w}{\partial K}\right)^2\frac{\partial V}{\partial w} + \frac{1}{2}K^2\frac{\partial^2 w}{\partial K^2}\frac{\partial V}{\partial w}
$$

$$
\begin{aligned}
&= \frac{\partial V}{\partial w}\left(1 + K\frac{w-2y}{2w}\frac{\partial w}{\partial K} + \frac{1}{2}K^2\frac{4y^2-4w-w^2}{8w^2}\left(\frac{\partial w}{\partial K}\right)^2 + \frac{1}{2}K^2\frac{\partial^2 w}{\partial K^2}\right)\\
&= \frac{\partial V}{\partial w}\left\{1 + 2K\frac{d_1}{\sqrt{w}}(T-t)\sigma_{\text{imp}}\frac{\partial \sigma_{\text{imp}}}{\partial K} + \frac{1}{2}K^2\frac{4y^2-4w-w^2}{8w^2}4w(T-t)\left(\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2\right.\\
&\qquad \left. +K^2(T-t)\left(\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2 + K^2(T-t)\sigma_{\text{imp}}\frac{\partial^2 \sigma_{\text{imp}}}{\partial K^2}\right\}\\
&= \frac{\partial V}{\partial w}\left\{1 + 2Kd_1\sqrt{T-t}\,\frac{\partial \sigma_{\text{imp}}}{\partial K} + \frac{1}{2}K^2\frac{4y^2-4w-w^2}{8w^2}4w(T-t)\left(\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2\right.\\
&\qquad \left. +K^2(T-t)\left(\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2 + K^2(T-t)\sigma_{\text{imp}}\frac{\partial^2 \sigma_{\text{imp}}}{\partial K^2}\right\}\\
&= \frac{\partial V}{\partial w}\left(1 + Kd_1\sqrt{T-t}\,\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2 + \frac{\partial V}{\partial w}\left(K^2(T-t)\sigma_{\text{imp}}\frac{\partial^2 \sigma_{\text{imp}}}{\partial K^2}\right)\\
&\qquad +K^2(T-t)\left(\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2\left(-d_1^2 + \frac{4y^2-4w-w^2}{4w} + 1\right)\frac{\partial V}{\partial w}\\
&= \frac{\partial V}{\partial w}\left(1 + Kd_1\sqrt{T-t}\,\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2 + \frac{\partial V}{\partial w}\left(K^2(T-t)\sigma_{\text{imp}}\frac{\partial^2 \sigma_{\text{imp}}}{\partial K^2}\right)\\
&\qquad +K^2(T-t)\left(\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2\left(-d_1^2 + d_1^2 + y - \frac{w}{2} - 1 + 1\right)\frac{\partial V}{\partial w}\\
&= \frac{\partial V}{\partial w}\left(1 + Kd_1\sqrt{T-t}\,\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2 + \frac{\partial V}{\partial w}\left(K^2(T-t)\sigma_{\text{imp}}\frac{\partial^2 \sigma_{\text{imp}}}{\partial K^2}\right)\\
&\qquad +K^2(T-t)\left(\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2\left(-\sqrt{w}d_1\right)\frac{\partial V}{\partial w}\\
&= \frac{\partial V}{\partial w}\left(1 + Kd_1\sqrt{T-t}\,\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2 + \frac{\partial V}{\partial w}K^2(T-t)\sigma_{\text{imp}}\left(\frac{\partial^2 \sigma_{\text{imp}}}{\partial K^2} - d_1\left(\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2\sqrt{T-t}\right).
\end{aligned}
$$

where

$$
\begin{aligned}
d_1 &= \frac{-y+\frac{1}{2}w}{\sqrt{w}}\\
&= \frac{\log\frac{S}{K} + \left(r - D + \frac{1}{2}\sigma_{\text{imp}}^2\right)(T-t)}{\sigma_{\text{imp}}\sqrt{T-t}}.
\end{aligned}
$$

Hence we have the following local volatility formula.

$$
\sigma(K,T) = \sqrt{\frac{\sigma_{\text{imp}}^2 + 2(T-t)\sigma_{\text{imp}}\frac{\partial \sigma_{\text{imp}}}{\partial T} + 2(r-D)K(T-t)\sigma_{\text{imp}}\frac{\partial \sigma_{\text{imp}}}{\partial K}}{\left(1 + Kd_1\sqrt{T-t}\,\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2 + K^2(T-t)\sigma_{\text{imp}}\left(\frac{\partial^2 \sigma_{\text{imp}}}{\partial K^2} - d_1\left(\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2\sqrt{T-t}\right)}}. \tag{4.2.6}
$$

By Eq(4.2.3), in terms of the implied volatility the implied risk-neutral probability density function is given by

$$
p(S,t;K,T) = e^{r(T-t)}\frac{\partial^2 V}{\partial K^2} = \frac{2}{K^2}e^{r(T-t)}\frac{\partial V}{\partial w}\,\Xi,
$$

where

$$\Xi = \left(1 + Kd_1\sqrt{T-t}\,\frac{\partial\sigma_{\text{imp}}}{\partial K}\right)^2 + K^2(T-t)\sigma_{\text{imp}}\left(\frac{\partial^2\sigma_{\text{imp}}}{\partial K^2} - d_1\left(\frac{\partial\sigma_{\text{imp}}}{\partial K}\right)^2\sqrt{T-t}\right).$$

Since

$$\frac{\partial V}{\partial w} = \frac{1}{2\sqrt{2\pi}}\,\frac{Se^{-D(T-t)}\,e^{-\frac{1}{8}\left(w-4y+\frac{4y^2}{w}\right)}}{\sqrt{w}} = \frac{1}{2\sqrt{2\pi}}\,\frac{Se^{-D(T-t)}\,e^{-\frac{1}{2}d_1^2}}{\sqrt{w}},$$

we get

$$\begin{aligned}
p(S,t;K,T) &= \frac{2}{K^2}e^{r(T-t)}\frac{1}{2\sqrt{2\pi}}\,\frac{Se^{-D(T-t)}\,e^{-\frac{1}{2}d_1^2}}{\sqrt{w}}\,\Xi \\
&= \frac{2}{K^2}\,\frac{1}{2\sqrt{2\pi}}\,\frac{Ke^{-y}\,e^{-\frac{1}{2}d_1^2}}{\sqrt{w}}\,\Xi \\
&= \frac{e^{-y}e^{-\frac{1}{2}d_1^2}}{K\sigma_{\text{imp}}\sqrt{2\pi(T-t)}}\,\Xi \\
&= \frac{e^{-\frac{1}{2}d_2^2}}{K\sigma_{\text{imp}}\sqrt{2\pi(T-t)}}\,\Xi,
\end{aligned}$$

where

$$\begin{aligned}
d_2 &= d_1 - \sqrt{w} \\
&= \frac{-y-\frac{w}{2}}{\sqrt{w}}.
\end{aligned}$$

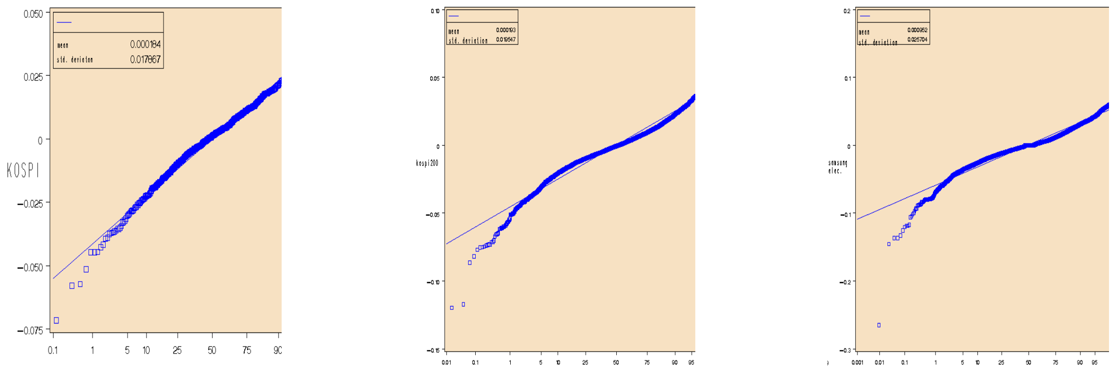
Example 4.2.1. If σ_{imp} is independent to E, i.e $\frac{\partial\sigma_{\text{imp}}}{\partial K} = 0$, the local volatility formula, Eq(4.2.6), reduces to

$$\sigma(T) = \sqrt{\sigma_{\text{imp}}^2 + 2(T-t)\sigma_{\text{imp}}\frac{\partial\sigma_{\text{imp}}}{\partial T}}.$$

This equation is the same as Eq(4.1.2). □

4.3 Stochastic Volatility

There is plenty of evidence that returns on equities, currencies and commodities are not normally distributed. They have higher peaks and fatter tails than predicted by a normal distribution(see Figure 4.2). This has also been cited as evidence for non-constant volatility.



(a) KOSPI (b) KOSPI 200 (c) Samsung Electronics

Figure 4.2: Fat tails

We assume that S satisfies

$$\begin{aligned} dS_t &= \mu S dt + \sigma_t S dW_1, \\ d\sigma_t &= p(S_t, \sigma_t, t)dt + q(S_t, \sigma_t, t)dW_2, \end{aligned}$$

where dW_1 and dW_2 have a correlation ρ.

The new stochastic quantity that we are modeling, the volatility, is not a traded asset. Thus, when volatility is stochastic we are faced with the problem of having a source of randomness that cannot be easily hedged away. Because we have two sources of randomness we must hedge our option with two other contracts, one being the underlying asset as usual, but now we also need another option to hedge the volatility risk. We therefore must set up a portfolio containing one option, with value $V(S, \sigma, t)$, an quantity $-\Delta$ of the asset and a quantity $-\Delta_1$ of another option with value $V_1(S, \sigma, t)$. We have

$$\Pi = V - \Delta S - \Delta_1 V_1.$$

It follows from Ito's formula that

$$\begin{aligned} d\Pi &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial \sigma}d\sigma + \frac{1}{2}\left(\frac{\partial^2 V}{\partial S^2}dS \cdot dS + 2\frac{\partial^2 V}{\partial S \partial \sigma}dS \cdot d\sigma + \frac{\partial^2 V}{\partial \sigma^2}d\sigma \cdot d\sigma\right) - \Delta dS \\ &\quad -\Delta_1\left(\frac{\partial V_1}{\partial t}dt + \frac{\partial V_1}{\partial S}dS + \frac{\partial V_1}{\partial \sigma}d\sigma + \frac{1}{2}\left(\frac{\partial^2 V_1}{\partial S^2}dS \cdot dS + 2\frac{\partial^2 V_1}{\partial S \partial \sigma}dS \cdot d\sigma + \frac{\partial^2 V_1}{\partial \sigma^2}d\sigma \cdot d\sigma\right)\right) \\ &= \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho q \sigma S \frac{\partial^2 V}{\partial S \partial \sigma} + \frac{1}{2}q^2 \frac{\partial^2 V}{\partial \sigma^2}\right) dt \\ &\quad -\Delta_1\left(\frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_1}{\partial S^2} + \rho q \sigma S \frac{\partial^2 V_1}{\partial S \partial \sigma} + \frac{1}{2}q^2 \frac{\partial^2 V_1}{\partial \sigma^2}\right) dt \end{aligned}$$

$$+\left(\frac{\partial V}{\partial S}-\Delta-\Delta_1\frac{\partial V_1}{\partial S}\right)dS+\left(\frac{\partial V}{\partial \sigma}-\Delta_1\frac{\partial V_1}{\partial \sigma}\right)d\sigma.$$

To eliminate all randomness from the portfolio we must choose

$$\begin{cases}\dfrac{\partial V}{\partial S}-\Delta-\Delta_1\dfrac{\partial V_1}{\partial S}=0,\\ \dfrac{\partial V}{\partial \sigma}-\Delta_1\dfrac{\partial V_1}{\partial \sigma}=0.\end{cases}$$

This leave us with

$$\begin{aligned}d\Pi &= \left(\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}+\rho q\sigma S\frac{\partial^2 V}{\partial S\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2 V}{\partial \sigma^2}\right)dt\\ &\quad -\Delta_1\left(\frac{\partial V_1}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V_1}{\partial S^2}+\rho q\sigma S\frac{\partial^2 V_1}{\partial S\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2 V_1}{\partial \sigma^2}\right)dt\\ &= r\Pi dt = r(V-\Delta S-\Delta_1V_1)dt = r\left(V-S\frac{\partial V}{\partial S}+\Delta_1S\frac{\partial V_1}{\partial S}-\Delta_1V_1\right)dt.\end{aligned}$$

Collecting all V terms on the left-hand side and all V_1 terms on the right-hand side we find that

$$\begin{aligned}&\frac{\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}+\rho q\sigma S\frac{\partial^2 V}{\partial S\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2 V}{\partial \sigma^2}+rS\frac{\partial V}{\partial S}-rV}{\frac{\partial V}{\partial \sigma}}\\ =\ &\frac{\frac{\partial V_1}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V_1}{\partial S^2}+\rho q\sigma S\frac{\partial^2 V_1}{\partial S\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2 V_1}{\partial \sigma^2}+rS\frac{\partial V_1}{\partial S}-rV_1}{\frac{\partial V_1}{\partial \sigma}}.\end{aligned}$$

Both sides can only be functions of independent variables, S, σ and t. Thus we have

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}+\rho q\sigma S\frac{\partial^2 V}{\partial S\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2 V}{\partial \sigma^2}+rS\frac{\partial V}{\partial S}-rV=-(p-\lambda q)\frac{\partial V}{\partial \sigma}$$

for some function $\lambda(S,\sigma,t)$.

✎ **The Market Price of Volatility Risk**

Consider the portfolio

$$\Pi = V-\Delta S.$$

By the Ito formula, we have

$$d\Pi = \left(\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}+\rho\sigma qS\frac{\partial^2 V}{\partial S\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2 V}{\partial \sigma^2}\right)dt+\left(\frac{\partial V}{\partial S}-\Delta\right)dS+\frac{\partial V}{\partial \sigma}d\sigma.$$

Because we are delta hedging the coefficient of dS is zero. We find that

$$\begin{aligned}d\Pi-r\Pi dt &= \left(\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}+\rho\sigma qS\frac{\partial^2 V}{\partial S\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2 V}{\partial \sigma^2}+rS\frac{\partial V}{\partial S}-rV\right)dt+\frac{\partial V}{\partial \sigma}d\sigma\\ &= -(p-\lambda q)\frac{\partial V}{\partial \sigma}dt+\frac{\partial V}{\partial \sigma}(pdt+qdW_2)\\ &= q\frac{\partial V}{\partial \sigma}\Big(\lambda dt+dW_2\Big).\end{aligned}$$

There are λ units of extra return, represented by dt. Hence the name 'market price of risk'.

The quantity $p-\lambda q$ is called the **risk-neutral drift rate** of the volatility.

4.4 The GARCH Models

The GARCH(Generalized model of Autoregressive Conditional Heteroscedasticity) is suggested by Engle(1982) and Bollerslev(1986).

The return r_t from stock with price S_t at time t is modeled as

$$r_t \quad := \quad \log \frac{S_{t+1}}{S_t} \quad = \quad \mu + \sigma_t \epsilon_t, \quad t = 0, 1, 2, \cdots$$

where ϵ_t is normally distributed. i.e $\epsilon_t \sim \mathcal{N}(0, 1)$. Note that σ_t is not normalized (short term) volatility. GARCH$(1, 1)$ model is given by

$$\begin{aligned} \sigma^2_{t+1} &= \omega + \alpha(r_t - \mu)^2 + \beta\sigma^2_t \\ &= \omega + \alpha(\sigma_t\epsilon_t)^2 + \beta\sigma^2_t \\ &= \omega + (\alpha\epsilon^2_t + \beta)\sigma^2_t. \end{aligned} \tag{4.4.1}$$

The standard normal variable ϵ_t may be written as

$$\epsilon_t \quad = \quad W(t) - W(t-1),$$

where $W(t)$ is standard Brownian motion. Thus we have

$$\sigma^2_{t+1} \quad = \quad \omega + \Big[\alpha\big(W(t) - W(t-1)\big)^2 + \beta\Big]\sigma^2_t.$$

Theorem 4.4.1 (Nelson(1990)). Let $h = \Delta t$ be the time increment. Choose mappings $\omega_h, \alpha_h, \beta_h$: $\mathbb{R}^+ \to \mathbb{R}^+$, such that

$$\begin{aligned} \frac{\omega_h}{h} &\longrightarrow \omega, \quad \text{as } h \to 0, \\ \frac{1 - \alpha_h h - \beta_h}{h} &\longrightarrow \theta, \quad \text{as } h \to 0, \\ \alpha_h\sqrt{2h} &\longrightarrow \alpha, \quad \text{as } h \to 0. \end{aligned}$$

Then the discrete GARCH$(1, 1)$ model in (4.4.1) with constant return mean converges in distribution to the stochastic differential equations

$$d\sigma^2_t \quad = \quad \Big(\omega - \theta\sigma^2_t\Big)dt + \alpha\sigma^2_t dZ(t)$$

where $Z(t)$ is a standard Brownian motion.

Proof. Rewrite Eq.(4.4.1) with the increment h and plug in the increment dependent parameters $\omega_h, \alpha_h, \beta_h$:

$$\sigma^2_h\Big((n+1)h\Big) \quad = \quad \omega_h + \Big[\alpha_h\big\{W(nh) - W((n-1)h)\big\}^2 + \beta_h\Big]\sigma^2_h(nh).$$

Subtracting $\sigma_h(nh)^2$ from this gives us

$$\begin{aligned} &\sigma^2_h\Big((n+1)h\Big) - \sigma^2_h(nh) \\ = \quad &\omega_h + \Big[\alpha_h\big\{W(nh) - W((n-1)h)\big\}^2 + \beta_h - 1\Big]\sigma^2_h(nh) \end{aligned}$$

$$
\begin{aligned}
&= \omega_h + \Big[\alpha_h h + \beta_h - 1\Big]\sigma_h^2(nh) + \alpha_h\Big[\big\{W(nh) - W((n-1)h)\big\}^2 - h\Big]\sigma_h^2(nh) \\
&= \left[\frac{\omega_h}{h} + \frac{\alpha_h h + \beta_h - 1}{h}\sigma_h^2(nh)\right] h + \sqrt{2h}\,\alpha_h\left[\frac{\big\{W(nh) - W((n-1)h)\big\}^2 - h}{\sqrt{2h}}\right]\sigma_h^2(nh).
\end{aligned}
$$

Let

$$
Z_h(nh) = \sum_{k=1}^{n} \frac{\big\{W(kh) - W((k-1)h)\big\}^2 - h}{\sqrt{2h}}
$$

and denote

$$
\begin{aligned}
X_k &= W(kh) - W((k-1)h), \\
\Xi_k &= \frac{X_k^2 - h}{\sqrt{2}\,h}.
\end{aligned}
$$

Then we have

$$
Z_h(nh) = \sum_{k=1}^{n} \frac{X_k^2 - h}{\sqrt{2h}} = \sqrt{h}\sum_{k=1}^{n}\Xi_k.
$$

Now, we can verify Ξ_k's are i.i.d. random variables with mean zero and variance 1. Since $EX_k^2 = h$, we have $E\Xi_k = 0$. Also, for $k \neq j$,

$$
\begin{aligned}
E\,\Xi_k\Xi_j &= \frac{1}{2h^2}E\Big[X_k^2X_j^2 - hX_k^2 - hX_j^2 + h^2\Big] \\
&= 0, \\
E\Xi_k^2 &= \frac{1}{2h^2}E\Big[X_k^4 - 2hX_k^2 + h^2\Big] \\
&= \frac{1}{2h^2}(3h^2 - 2h^2 + h^2) \\
&= 1.
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
Z_h(nh) &= \sqrt{nh}\,\frac{\sum_{k=1}^{n}\Xi_k}{\sqrt{n}} \\
&= \sqrt{t}\,\frac{\sum_{k=1}^{n}\Xi_k}{\sqrt{n}} \\
&\longrightarrow Z(t),
\end{aligned}
$$

where $Z(t)$ is a standard Brownian motion. Then

$$
\sigma_h^2\Big((n+1)h\Big) - \sigma_h^2(nh) = \left[\frac{\omega_h}{h} + \frac{\alpha_h h + \beta_h - 1}{h}\sigma_h^2(nh)\right] h + \sqrt{2h}\,\alpha_h\Big[Z_h(nh) - Z_h((n-1)h)\Big]\sigma_h^2(nh)
$$

Letting $n \to \infty$ to the both sides, we complete the proof. ■

Remark 4.4.2.

1. For instant

$$\omega_h = \omega h, \quad \alpha_h = \frac{\alpha}{\sqrt{2h}}, \quad \beta_h = 1 - \alpha_h h - (1-\alpha-\beta)h.$$

   satisfy the condition of Theorem (4.4.1). Thus we get

$$\sigma^2\Big((n+1)\Delta t\Big) = \omega\Delta t + \frac{\alpha}{\sqrt{2\Delta t}}(r(n\Delta t)-\mu)^2 + \big\{1 - \frac{\alpha}{\sqrt{2\Delta t}} - (1-\alpha-\beta)\Delta t\big\}\sigma^2(n\Delta t). \quad (4.4.2)$$

2. Replace α by $\sqrt{2}\alpha$ and Δt by 1 in $\frac{\alpha}{\sqrt{2\Delta t}}$ of Eq. (4.4.2). Then we have $\theta = 1-\alpha-\beta$ and

$$\begin{aligned} \sigma^2(n+1) &= \omega + \alpha(r(n)-\mu)^2 + \beta\sigma^2(n), && (4.4.3) \\ d\sigma_t^2 &= \Big(\omega - \theta\sigma_t^2\Big)dt + \sqrt{2}\,\alpha\sigma_t^2 dZ(t) && (4.4.4) \end{aligned}$$

3. Let $V(n) = \sigma^2(n)$ and $u(n) = r(n) - \mu$. Assume that $\alpha+\beta+\gamma = 1$ and

$$V(n+1) = \gamma V_L + \alpha u^2(n) + \beta V(n). \quad (4.4.5)$$

   Then,

$$dV_t = \gamma\Big(V_L - V_t\Big)dt + \sqrt{2}\,\alpha V_t dZ(t).$$

4.5 The Heston Model

In this section, we review the Heston model(1993). Suppose that

$$\begin{aligned} dS(t) &= S(t)\Big(\mu(t)dt + \sigma(t)dW_1(t)\Big), \\ d\sigma(t) &= \alpha(t)\sigma(t)dt + \eta(t)dW_2(t), \\ dW_1(t)\cdot dW_2(t) &= \rho dt. \end{aligned} \tag{4.5.1}$$

Let $v(t) = \sigma^2(t)$. Ito's formula implies

$$\begin{aligned} dv(t) &= 2\sigma(t)d\sigma(t) + \frac{1}{2}\cdot 2(d\sigma(t))^2 \\ &= \Big(2\alpha(t)v(t) + \eta^2(t)\Big)dt + 2\eta(t)\sqrt{v(t)}\,dW_2(t). \end{aligned} \tag{4.5.2}$$

Hence Eq(4.5.1) and Eq(4.5.2) are equivalent. Ito's formula implies that

$$\frac{\partial V}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \rho\eta v\beta S\frac{\partial^2 V}{\partial v\partial S} + \frac{1}{2}\eta^2 v\beta^2\frac{\partial^2 V}{\partial v^2} + rS\frac{\partial V}{\partial S} - rV = -(\alpha - \lambda\eta\beta\sqrt{v})\,\frac{\partial V}{\partial v},$$

where λ is some function of S, v and t. $\lambda(S, v, t)$ is called the market price of (volatility) risk.

내용 채워야 할 곳...

4.6 The Fundamental Transform

We assume that

$$\begin{aligned} dS_t &= (r-\delta)S_t\,dt + \sigma_t\,S_t\,dB_t, \\ dV_t &= b(V_t)\,dt + a(V_t)\,dW_t, \end{aligned}$$

where $V_t = \sigma_t^2$, dB_t and dW_t are correlated Brownian motion with correlation $\rho(V_t)$.

We can write the PDE for generalized European style contingent claim with price $F(S_t, V_t, t)$ and expiration T. That equation becomes

$$\begin{aligned} -\frac{\partial F}{\partial t} &= -rF + \mathcal{A}F, \quad \text{where} \\ \mathcal{A}F &= (r-\delta)S\frac{\partial F}{\partial S} + \frac{1}{2}VS^2\frac{\partial^2 F}{\partial S^2} \\ &\quad + b(V)\frac{\partial F}{\partial V} + \frac{1}{2}a^2(V)\frac{\partial^2 F}{\partial V^2} + \rho(V)a(V)\sqrt{V}S\frac{\partial^2 F}{\partial S\partial V}. \end{aligned}$$

Step 1. The first step is change of variable from S to $x = \log S$, letting $F(S, V, t) = f(x, V, t)$.

$$-f_t = -rf + \left(r - f - \frac{1}{2}V\right) f_x + \frac{1}{2}Vf_{xx} + bf_V + \frac{1}{2}a^2 f_{VV} + \rho a\sqrt{V}f_{xV}.$$

Step 2.

Step 3.

내용 채워야 할 곳...

4.7 Estimating Volatilities

Define σ_n^2 as the volatility of a market variable on day n. The variable u_i is defined as the return during day i:

$$u_i = \log \frac{S_i}{S_{i-1}} \quad \text{or} \quad \frac{S_i - S_{i-1}}{S_{i-1}}.$$

✎ **ARCH(m)**

$$\sigma_n^2 = \gamma V + \sum_{i=1}^{m} \alpha_i u_{n-i}^2, \quad \gamma + \sum_{i=1}^{m} \alpha_i = 1,$$

where V is the long-run volatility and γ is the weight assigned to V.

✎ **EWMA**

$$\begin{aligned} \sigma_n^2 &= \lambda \sigma_{n-1}^2 + (1-\lambda) u_{n-1}^2 \\ &= (1-\lambda) \sum_{i=1}^{\infty} \lambda^{i-1} u_{n-i}^2. \end{aligned}$$

✎ **GARCH($1,1$)**

$$\sigma_n^2 = \gamma V + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2, \quad \gamma + \alpha + \beta = 1.$$

Setting $\omega = \gamma V$, the GARCH$(1,1)$ model can also be written

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2, \quad \gamma + \alpha + \beta = 1.$$

Once ω, α, and β have been estimated, we can calculate

$$\begin{aligned} \gamma &= 1 - \alpha - \beta, \\ V &= \frac{\omega}{\gamma}. \end{aligned}$$

For a stable GARCH$(1,1)$ process we require $\alpha + \beta < 1$. Otherwise the weight γ is negative.

GARCH$(1,1)$ model incorporates mean-reversion whereas the EWMA model does not. The GARCH$(1,1)$ model is, therefore, theoretically more appealing than the EWMA model. In circumstances where the best fit value of ω turns to be negative the GARCH$(1,1)$ model is not stable and it makes sense to switch to the EWMA model.

✎ **GARCH(p,q)**

$$\sigma_n^2 = \gamma V + \sum_{i=1}^{p} \alpha_i u_{n-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{n-i}^2, \quad \gamma + \sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{q} \beta_i = 1.$$

Part III

Exotic Options

Chapter 5

Asian Options

Asian options are especially popular in the currency and commodity markets. An average is less volatile than the underlying asset itself and will lower the price of an average rate option compared with a similar standard option. Asian options are options where the payoff depends on the average price of the underlying asset during some part of the life of the option.

The precise definition of the average used in Asian option depends on two elements:

1. How the data points are combined to form an average.
   i. Arithmetic average.
   ii. Geometric average.
2. Which data points are used.
   i. All quoted prices.
   ii. Just a subset of quoted prices, and over particular time period.

	Discrete	Continuous
Arithmetic	$\frac{1}{n}\sum_{i=1}^{n} S(t_i)$	$\frac{1}{t}\int_0^t S(u)du$
Geometric	$\left(\prod_{i=1}^{n} S(t_i)\right)^{\frac{1}{n}} = \exp\left(\frac{1}{n}\sum_{i=1}^{n}\log S(t_i)\right)$	$\exp\left(\frac{1}{t}\int_0^t \log S(u)du\right)$

Table 5.1: The average rate of Asian options

Note that in both continuous and discrete cases, the geometric average variable is log-normally distributed so that its expectation and variance values may be calculated explicitly.

	A: average, K: strike price
average strike call	$(S_T - A)^+$
average strike put	$(A - S_T)^+$
average rate call	$(A - K)^+$
average rate put	$(K - A)^+$

Table 5.2: Asian options

5.1 Continuous Sampling(Geometric Average): The Pricing Equation

The average is defined as

$$\frac{1}{t}\int_0^t f(S,u)du.$$

If we introduce the new state variable

$$I(t) = \int_0^t f(S,u)du \tag{5.1.1}$$

then,

$$dI(t) = f(S,t)dt.$$

Set up a portfolio containing one of the path-dependent option and short a number of Δ of the underlying asset:

$$\Pi = V(S,I,t) - \Delta S$$

where V is the value of path-dependent option. It follows from Ito's lemma that

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \frac{\partial V}{\partial I} dI + \frac{\partial V}{\partial S} dS - \Delta dS.$$

Choosing

$$\Delta = \frac{\partial V}{\partial S}$$

to hedge the risk, we find that

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + f(S,t)\frac{\partial V}{\partial I}\right) dt.$$

Now, the portfolio earns the risk-free rate of interest r, i.e. $d\Pi = r\Pi dt$, leading to the pricing equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} + f(S,t)\frac{\partial V}{\partial I} - rV = 0. \tag{5.1.2}$$

Theorem 5.1.1 (Geometric Average with Continuous Sampling). If

$$dS_t = S_t(rdt + \sigma dW_t) \tag{5.1.3}$$

and[1]

$$G_t = \exp\left(\frac{1}{T-t_0}\int_{t_0}^{t} \log S_u du\right), \quad t \geq t_0, \tag{5.1.4}$$

where T is fixed constant. The value of Asian option with strike X equals to the vanilla option with volatility

$$\frac{\sigma}{\sqrt{3}}$$

and dividend rate

$$\frac{1}{2}\left(r + q + \frac{1}{6}\sigma^2\right).$$

Proof. We will use the following notation:

$$V_t = \log S_t \quad \text{and} \quad Z_t = \log G_t.$$

From the Ito's lemma it follows that (5.1.3) and (5.1.4) give rise to the following system of stochastic differential equations

$$d\begin{bmatrix} V_t \\ Z_t \end{bmatrix} = \left(\begin{bmatrix} 0 & 0 \\ \beta & 0 \end{bmatrix}\begin{bmatrix} V_t \\ Z_t \end{bmatrix} + \begin{bmatrix} r - \frac{1}{2}\sigma^2 \\ 0 \end{bmatrix}\right)dt + \begin{bmatrix} \sigma \\ 0 \end{bmatrix} dW_t.$$

We can see that $(V_t, Z_t)^t$ is binomially distributed. Hence $\log G_t/G_{t_0}$ is normally distributed. Furthermore, it follows that

$$d\begin{bmatrix} EV_t \\ EZ_t \end{bmatrix} = \left(\begin{bmatrix} 0 & 0 \\ \beta & 0 \end{bmatrix}\begin{bmatrix} EV_t \\ EZ_t \end{bmatrix} + \begin{bmatrix} r - \frac{1}{2}\sigma^2 \\ 0 \end{bmatrix}\right)dt.$$

The covariance matrix of $(V_t, Z_t)^t$ as defined by

$$K(t) = \begin{bmatrix} K_{11}(t) & K_{12}(t) \\ K_{21}(t) & K_{22}(t) \end{bmatrix}$$

is the unique symmetric non-negative definite solution of the following matrix differential equation:

$$\begin{aligned} d\begin{bmatrix} K_{11}(t) & K_{12}(t) \\ K_{21}(t) & K_{22}(t) \end{bmatrix} &= \left(\begin{bmatrix} 0 & 0 \\ \beta & 0 \end{bmatrix}\begin{bmatrix} K_{11}(t) & K_{12}(t) \\ K_{21}(t) & K_{22}(t) \end{bmatrix}\right. \\ &\quad \left. + \begin{bmatrix} K_{11}(t) & K_{12}(t) \\ K_{21}(t) & K_{22}(t) \end{bmatrix}\begin{bmatrix} 0 & \beta \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \sigma \\ 0 \end{bmatrix}\begin{bmatrix} \sigma & 0 \end{bmatrix}\right)dt \\ &= \begin{bmatrix} \sigma^2 & \beta K_{11}(t) \\ \beta K_{11}(t) & \beta K_{12}(t) + \beta K_{21}(t) \end{bmatrix} dt \end{aligned}$$

It is quite straightforward to solve these ODE and thus find that

$$EV_t = \left(r - \frac{1}{2}\sigma^2\right)(t - t_0) + V_{t_0},$$

[1] G_t is neither Markov nor a martingale.

$$EZ_t = \frac{1}{2}\beta(r-\frac{1}{2}\sigma^2)(t-t_0)^2 + \beta V_{t_0}(t-t_0),$$

and

$$\begin{bmatrix} K_{11}(t) & K_{12}(t) \\ K_{21}(t) & K_{22}(t) \end{bmatrix} = \begin{bmatrix} \sigma^2(t-t_0) & \frac{1}{2}\beta\sigma^2(t-t_0)^2 \\ \frac{1}{2}\beta\sigma^2(t-t_0)^2 & \frac{1}{3}\beta^2\sigma^2(t-t_0)^3 \end{bmatrix}.$$

Note that

$$\begin{aligned} Z_{t_0} &= \log G_{t_0} \\ &= \lim_{t\to t_0} \frac{1}{T-t_0}\int_{t_0}^{t} \log S_u du \\ &= \lim_{t\to t_0} \frac{t-t_0}{T-t_0}\frac{1}{t-t_0}\int_{t_0}^{t} \log S_u du \\ &= \lim_{t\to t_0} \frac{t-t_0}{T-t_0} \log S_{t_0} \\ &= 0. \end{aligned}$$

Hence we get that

$$\begin{aligned} \log G_T - \log G_{t_0} &\sim \beta V_{t_0}(T-t_0) + N\left(\frac{1}{2}\beta(r-\frac{1}{2}\sigma^2)(T-t_0)^2,\ \frac{1}{3}\beta^2\sigma^2(T-t_0)^3\right) \\ &= \log S_{t_0} + N\left(\frac{1}{2}(r-\frac{1}{2}\sigma^2)(T-t_0),\ \frac{1}{3}\sigma^2(T-t_0)\right). \end{aligned}$$

In order to use Black-Scholes formula, we should manipulate the mean part.

$$\begin{aligned} \frac{1}{2}\left(r-\frac{1}{2}\sigma^2\right)(T-t_0) &= \frac{1}{2}\left(r-\frac{1}{2}\sigma^2\right)(T-t_0) + \frac{1}{2}\cdot\frac{1}{3}\sigma^2(T-t_0) - \frac{1}{2}\cdot\frac{1}{3}\sigma^2(T-t_0) \\ &= \frac{1}{2}\Big(r-\frac{1}{6}\sigma^2\Big)(T-t_0) - \frac{1}{2}\cdot\frac{1}{3}\sigma^2(T-t_0). \end{aligned}$$

If we take account of the dividend rate, $q(r \leftarrow r-q)$, we could have

$$\frac{1}{2}\Big(r-q-\frac{1}{6}\sigma^2\Big)(T-t_0) - \frac{1}{2}\cdot\frac{1}{3}\sigma^2(T-t_0).$$

The Black-Scholes formula, for example, gives the Asian call option price

$$E[e^{r(T-t_0)}(G_T-X)^+|\mathcal{F}_{t_0}] = S_{t_0}e^{-q_A(T-t_0)}N(d_1) - XN(d_2)$$

where

$$\begin{aligned} q_A &= r - \frac{1}{2}\Big(r-q-\frac{1}{6}\sigma^2\Big) = \frac{1}{2}\Big(r+q+\frac{1}{6}\sigma^2\Big), \\ d_1 &= \frac{\log(S/X) + (r-q_A+\frac{1}{2}\sigma_A^2)(T-t_0)}{\sigma_A\sqrt{T-t_0}}, \\ d_2 &= d_1 - \sigma_A\sqrt{T-t_0}, \\ \sigma_A &= \frac{\sigma}{\sqrt{3}}. \end{aligned}$$

■

5.1.1 Reduction to a One-Dimensional Equation

Solutions to (5.1.2) are defined on the domain

$$S > 0, \quad I > 0, \quad 0 \le t \le T.$$

If $f(x,t) = x$ in (5.1.1), (5.1.2) can be written as

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} + S\frac{\partial V}{\partial I} - rV = 0. \tag{5.1.5}$$

The payoff for the average strike call option is

$$\max\left(S_T - \frac{1}{T}\int_0^T S_\tau d\tau \ , \ 0\right).$$

We can write the running payoff as

$$I_t \max\left(R_t - \frac{1}{t}, 0\right),$$

where

$$I_t = \int_0^t S_\tau d\tau, \quad \text{and} \quad R_t = \frac{S}{I_t}.$$

The payoff at expiry may be written as

$$I_T \max\left(R_T - \frac{1}{T}, 0\right).$$

Theorem 5.1.2. Assume that the option value has the form

$$V(S,I,t) = \frac{S}{R}\, W(R,t),$$

W satisfies

$$\frac{\partial W}{\partial t} + \frac{1}{2}\sigma^2 R^2 \frac{\partial^2 W}{\partial R^2} + R(r-R)\frac{\partial W}{\partial R} - (r-R)W \le 0 \tag{5.1.6}$$

If the option is European we have strict equality in (5.1.6). If it is American we may have inequality in (5.1.6) but the constraint

$$W(R,t) \ge \max\left(R - \frac{1}{t}, 0\right).$$

Proof. From

$$\begin{aligned} R &= \frac{S}{I}, \\ S &= S + 0 \cdot I \end{aligned}$$

we have

$$
\begin{aligned}
\frac{\partial V}{\partial S} &= \frac{\partial V}{\partial R}\frac{\partial R}{\partial S} + \frac{\partial V}{\partial S}\frac{\partial S}{\partial S} \\
&= \frac{\partial V}{\partial R}\frac{1}{I} + \frac{\partial V}{\partial S} \\
&= \frac{1}{I}\left(-\frac{S}{R^2}W + \frac{S}{R}\frac{\partial W}{\partial R}\right) + \frac{1}{R}W \\
&= \cancel{-\frac{1}{R}W} + \frac{\partial W}{\partial R} + \cancel{\frac{1}{R}W} \\
&= \frac{\partial W}{\partial R}, \\
\frac{\partial^2 V}{\partial S^2} &= \frac{\partial^2 W}{\partial R^2}\frac{\partial R}{\partial S} + \cancelto{0}{\frac{\partial}{\partial S}\frac{\partial W}{\partial R}}\frac{\partial S}{\partial S}, \quad (W \text{ is independent to } S), \\
&= \frac{1}{I}\frac{\partial^2 W}{\partial R^2},
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial V}{\partial I} &= \frac{\partial V}{\partial R}\frac{\partial R}{\partial I} \\
&= -\frac{S}{I^2}\left(-\frac{S}{R^2}W + \frac{S}{R}\frac{\partial W}{\partial R}\right) \\
&= W - R\frac{\partial W}{\partial R}.
\end{aligned}
$$

Substituting these equations into (5.1.5), we have

$$
\begin{aligned}
& I\frac{\partial W}{\partial t} + \frac{1}{2}\sigma^2 S^2\,\frac{1}{I}\frac{\partial^2 W}{\partial R^2} + rS\frac{\partial W}{\partial R} + S\left(W - R\frac{\partial W}{\partial R}\right) - rIW \\
= \;& I\left(\frac{\partial W}{\partial t} + \frac{1}{2}\sigma^2 R^2\frac{\partial W}{\partial R} + R(r-R)\frac{\partial W}{\partial R} - (r-R)W\right).
\end{aligned}
$$

■

Remark 5.1.3. Note that one-dimensional PDE for Asian option is difficult to solve numerically since the diffusion term is very small for values of interest on finite difference grid. The dirac delta function also appears as a coefficient of the PDE in the case of the floating strike option.

5.2 Discrete Sampling(Geometric Average): The Pricing Equation

If the underlying asset price is assumed to be log-normally distributed, analytic formulas are available for valuing geometric averaged Asian options.[2]

Theorem 5.2.1. [Discrete Sampling Geometric Average] Suppose that $\{t_1, t_2, \cdots, t_n = T\}$ are equally spaced between 0 and T. i.e. $t_i = iT/n$. The geometric average

$$G_T = \left(\prod_{i=1}^{n} S(t_i)\right)^{\frac{1}{n}} = \exp\left(\frac{1}{n}\sum_{i=1}^{n}\log S(t_i)\right)$$

is log-normally distributed, if S_t is a geometric Brownian motion

$$dS_t = S_t\Big((r-q)dt + \sigma dW_t\Big),$$

and

$$\begin{aligned}\log G_T \quad & \sim \quad N\Big((r-q-\frac{1}{2}\sigma^2)\frac{n+1}{2n}T, \frac{(n+1)(2n+1)}{6n^2}\ \sigma^2 T\Big)\\ & \xrightarrow{\text{as } n\to\infty} \quad N\left(\frac{1}{2}(r-q-\frac{1}{2}\sigma^2)T,\ \frac{\sigma^2}{3}T\right)\end{aligned}$$

Proof. The set of random variables

$$\log\frac{S_{t_1}}{S_0},\ \log\frac{S_{t_2}}{S_0},\cdots,\ \log\frac{S_{t_n}}{S_0}$$

are jointly normally distributed and mean and covariance matrix given by

$$E\left[\log\frac{S_{t_i}}{S_0}\right] = \left(r-q-\frac{1}{2}\sigma^2\right)t_i,$$

and

$$\Sigma = \sigma^2\begin{bmatrix} t_1 & t_1 & \cdots & t_1\\ t_1 & t_2 & \cdots & t_2\\ \vdots & \vdots & \ddots & \vdots\\ t_1 & t_2 & \cdots & t_n\end{bmatrix}, \quad \text{i.e } \Sigma_{ij} = \sigma^2\ t_{i\wedge j}.$$

Since

$$\log G_T = \frac{1}{n}\sum_{i=1}^{n}\log\frac{S_{t_i}}{S_0},$$

we have that

$$E[\log G_T] = \frac{1}{n}\left(r-q-\frac{1}{2}\sigma^2\right)\sum_{i=1}^{n} t_i$$

[2] See A. Kemna and A. Vorst(1990)

$$
\begin{aligned}
&= \frac{1}{n}\left(r-q-\frac{1}{2}\sigma^2\right)\sum_{i=1}^{n}\frac{iT}{n} \\
&= \left(r-q-\frac{1}{2}\sigma^2\right)\frac{n+1}{2n}T \\
&\xrightarrow{\text{as } n\to\infty} \left(r-q-\frac{1}{2}\sigma^2\right)\frac{T}{2}.
\end{aligned}
$$

Also,

$$
\begin{aligned}
\mathrm{Var}\left[\log G_T\right] &= \frac{\sigma^2}{n^2}\sum_{i=1}^{n} t_i(2n+1-2i) \\
&= \frac{\sigma^2 T}{n^3}\sum_{i=1}^{n}\left((2n+1)i-2i^2\right) \\
&= \frac{\sigma^2 T}{n^3}\,\frac{n(n+1)(2n+1)}{6} \\
&= \frac{(n+1)(2n+1)}{6n^2}\,\sigma^2 T \\
&\xrightarrow{\text{as } n\to\infty} \frac{\sigma^2}{3}T.
\end{aligned}
$$

■

From Theorem(5.2.1), we can see that the expected growth rate of geometric average of stock price over period T is

$$
\left(r-q-\frac{1}{2}\sigma^2\right)\frac{T}{2}+\frac{1}{2}\cdot\frac{\sigma^2}{3}T \quad = \quad \left(r-q-\frac{1}{6}\sigma^2\right)\frac{T}{2}
$$

rather than $r-q$. The geometric average Asian option can, therefore, be treated like a regular option with volatility set equal to $\frac{\sigma}{\sqrt{3}}$ and the dividend yield equal to

$$
r-\frac{1}{2}\left(r-q-\frac{1}{6}\sigma^2\right)=\frac{1}{2}\left(r+q+\frac{\sigma^2}{6}\right).
$$

This result coincides with Theorem(5.1.1).

5.3 Arithmetic Average: Turnbull and Wakeman Approximation

✎ **Newly issued Asian option**

The discrete arithmetic average is defined by

$$A_T = \frac{1}{n}\sum_{i=1}^{n} S_{t_i}, \quad t_i = \frac{iT}{n}.$$

The expectation(first moment) of A_T can be calculated as [3]:

$$\begin{aligned}
E[A_T] &= \frac{1}{n}\sum_{i=1}^{n} E[S_{t_i}] \\
&= \frac{1}{n}\sum_{i=1}^{n} S_0 e^{(r-q)t_i} \\
&\xrightarrow{\text{as } n\to\infty} \frac{S_0}{T}\int_0^T e^{(r-q)x}dx = S_0\,\frac{e^{(r-q)T}-1}{(r-q)T} := M_1
\end{aligned}$$

To get the second moment of A_T, first consider $E[S_{t_i}S_{t_j}]$.

$$\begin{aligned}
E[S_{t_i}S_{t_j}] &= S_0^2 E\left[\exp\Big((r-q-\frac{1}{2}\sigma^2)t_i + \sigma W_{t_i} + (r-q-\frac{1}{2}\sigma^2)t_j + \sigma W_{t_j}\Big)\right] \\
&= S_0^2\ \exp\Big((r-q-\frac{1}{2}\sigma^2)(t_i+t_j)\Big)E\left[\exp\Big(\sigma(W_{t_i}+W_{t_j})\Big)\right] \\
&= S_0^2\ \exp\Big((r-q-\frac{1}{2}\sigma^2)(t_i+t_j)\Big)\ \exp\Big(\frac{\sigma^2}{2}(t_i+t_j+2\rho_{ij}\sqrt{t_it_j})\Big) \\
&= S_0^2 \exp\Big((r-q)(t_i+t_j)\Big)\exp\Big(\sigma^2 t_i\Big), \quad \text{if } t_i < t_j \ \left(\rho_{ij} = \frac{\sqrt{t_i}}{\sqrt{t_j}}\right).
\end{aligned}$$

Now, let us calculate the second moment of A_T.

$$\begin{aligned}
E\left[\frac{1}{n}\sum_{i=1}^{n} S_i\right]^2 &= \frac{1}{n^2}\sum_{i=1}^{n} E[S_i^2] + \frac{2}{n^2}\sum_{i<j} E[S_{t_i}S_{t_j}] \\
&= \frac{1}{n^2}\sum_{i=1}^{n} E[S_i^2] + \frac{2}{n^2}\sum_{i<j} S_0^2\exp\Big((r-q)(t_i+t_j)\Big)\exp\Big(\sigma^2 t_i\Big) \\
&\xrightarrow{\text{as } n\to\infty} 0 + \frac{2S_0^2}{T^2}\int_0^T\int_x^T \exp\big((r-q)(x+y)+\sigma^2 x\big)dy\,dx.
\end{aligned}$$

Straightforward but rather cumbersome integration leads to the following result:

$$\begin{aligned}
&\frac{2S_0^2}{T^2}\int_0^T\int_x^T \exp\big((r-q)(x+y)+\sigma^2 x\big)dy\,dx \\
=\ &\frac{2S_0^2}{(r-q)T^2}\int_0^T e^{(r-q+\sigma^2)x}(e^{(r-q)T} - e^{(r-q)x})dx
\end{aligned}$$

[3] See Turnbull and Wakeman(1991).

$$
\begin{aligned}
&= \frac{2S_0^2}{(r-q)T^2}\left[e^{(r-q)T}\,\frac{e^{(r-q+\sigma^2)T}-1}{r-q+\sigma^2}-\frac{e^{(2(r-q)+\sigma^2)T}-1}{2(r-q)+\sigma^2}\right]\\
&= \frac{2S_0^2}{(r-q)T^2}\left[\frac{(r-q)e^{(2(r-q)+\sigma^2)T}}{(r-q+\sigma^2)(2(r-q)+\sigma^2)}+\frac{1}{2(r-q)+\sigma^2}-\frac{e^{(r-q)T}}{r-q+\sigma^2}\right]\\
&= \frac{2e^{(2(r-q)+\sigma^2)T}S_0^2}{((r-q)+\sigma^2)(2(r-q)+\sigma^2)T^2}+\frac{2S_0^2}{(r-q)T^2}\left[\frac{1}{2(r-q)+\sigma^2}-\frac{e^{(r-q)T}}{r-q+\sigma^2}\right] := M_2.
\end{aligned}
$$

If we assume the limit of the average stock price A_T^c is log-normal, we get the followings: (q_A and σ_A are to be determined)

$$
\begin{aligned}
A_T^c &\approx S_0\exp\left((r-q_A-\frac{1}{2}\sigma_A^2)T+\sigma_A W_T\right),\\
E[A_T^c] &= M_1 = S_0\exp((r-q_A)T)\\
M_2 &= S_0^2\exp\left(2(r-q_A-\frac{1}{2}\sigma_A^2)T+2\sigma_A^2T\right)\\
&= S_0^2\exp\left(2(r-q_A)T+\sigma_A^2T\right)\\
&= M_1^2\exp(\sigma_A^2T).
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
q_A &= r-\frac{1}{T}\log\frac{M_1}{S_0},\\
\sigma_A &= \sqrt{\frac{1}{T}\log\frac{M_2}{M_1^2}}.
\end{aligned}
$$

The value of Asian call option(Arithmetic) with strike X is given by

$$
\begin{aligned}
c &= S_0e^{-q_AT}N(d_1)-Xe^{-rT}N(d_2),\\
d_1 &= \frac{\log\frac{S_0}{X}+(r-q_A+\frac{\sigma_A^2}{2})T}{\sigma_A\sqrt{T}},\\
d_2 &= d_1-\sigma_A\sqrt{T}.
\end{aligned}
$$

✎ **Already issued Asian option**

Suppose that the average period is composed of a period of length t_1 over which prices have already been observed and a future period of length t_2. Suppose the average of stock price during past period t_1 is $\overline{S}$. The payoff of (arithmetic) average price call option at maturity is

$$
\left(\frac{\overline{S}\,t_1+S_{\text{ave}}\,t_2}{t_1+t_2} - X\right)^+
$$

where S_{ave} is the average during remaining period t_2. This is the same as

$$
\frac{t_2}{t_1+t_2}\left(S_{\text{ave}} - X^*\right)^+
$$

where

$$
X^*=\frac{t_1+t_2}{t_2}X-\frac{t_1}{t_2}\,\overline{S}.
$$

1. $X^* > 0$: The option can be valued in the same way as a newly issued Asian option.
2. $X^* \le 0$: The option is certain to be excised. Thus the price of Asian option in this case is

$$\frac{t_2}{t_1 + t_2}\left(M_1 e^{-rt_2} \quad - \quad X^* e^{-rt_2}\right).$$

5.4 Arithmetic Average: Levy's Approximation

See Levy(1992).

Chapter 6

Path Dependent Options

6.1 MIN-MAX Distribution of a Brownian Motion with Drift

Let $\{X(t) : 0 \leq t < \infty\}$ be any continuous process.

Definition 6.1.1. 1. For any $y \in \mathbb{R}$, the hitting time of y, $\tau(X, y)$, sometimes denoted by $\tau(y)$ or τ_y, is defined by

$$\tau(y) \quad = \quad \inf\{t \geq 0 | X(t) = y\}.$$

The X-process absorbed at y is defined by

$$X_y(t) \quad = \quad X(t \wedge \tau)$$

where we have used the notation $\alpha \wedge \beta = \min(\alpha, \beta)$.

2. For given process $X(t)$, the maximum and minimum process $M_X(t)$ and $m_X(t)$ are defined by

$$\begin{aligned} M_X(t) &= \sup_{0 \leq s \leq t} X(s) \\ m_X(t) &= \inf_{0 \leq s \leq t} X(s). \end{aligned}$$

□

Recall that S_t is given by

$$S(t) \quad = \quad \exp\Big(\log S_0 + \mu t + \sigma W_t\Big) = e^{X(t)}$$

where

$$X(t) \quad = \quad \log S_0 + \mu t + \sigma W_t,$$

By Ito lemma, we have

$$dX(t) \quad = \quad \mu dt + \sigma dW_t, \quad X(0) = \log S_0 := \alpha. \tag{6.1.1}$$

Theorem 6.1.2. The density, $f_\beta(x;t,\alpha)$, of the absorbed process $X_\beta(t)$, where X is defined by (6.1.1), is given by

$$f_\beta(x;t,\alpha) = \varphi(x;\mu t+\alpha,\sigma^2 t) - \exp\left(-\frac{2\mu(\alpha-\beta)}{\sigma^2}\right)\varphi(x;\mu t-\alpha+2\beta,\sigma^2 t),$$

where

$$\varphi(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

The support of this density is the interval (β,∞) if $\alpha>\beta$, and the interval $(-\infty,\beta)$ if $\alpha<\beta$. Note that the distribution $X_\beta(t)$ is of course a mixed distribution in the sense that it has a point mass at $x=\beta$(the probability that the process is absorbed prior to time t) and the density $f_\beta(x;t,\alpha)$.

Proof. We will prove the theorem under the assumption $\beta<\alpha$. Then we have

$$\begin{aligned} P(\beta < X_\beta(t)\le x) &= P\left(\beta < m_X(t),\ X(t)\le x\right) \\ &= P\left(\frac{\beta-\alpha}{\sigma} < \inf_{0\le s\le t}\left(\frac{\mu s}{\sigma}+W(s)\right),\ \frac{\mu t}{\sigma}+W(t)\le\frac{x-\alpha}{\sigma}\right) \\ &= P\left(-\frac{\beta-\alpha}{\sigma} > \sup_{0\le s\le t}\left(-\frac{\mu s}{\sigma}-W(s)\right),\ \frac{\mu t}{\sigma}+W(t)\le\frac{x-\alpha}{\sigma}\right). \end{aligned}$$

Let $Z(t)=\frac{\mu}{\sigma}t+W(t)$. Now we have

$$\begin{aligned} P(\beta < X_\beta(t)\le x) &= P\left(\sup_{0\le s\le t}\left(-Z(s)\right) < \frac{\alpha-\beta}{\sigma},\ -Z(t)\ge\frac{\alpha-x}{\sigma}\right) \\ &= P\left(M_{-Z}(t) < \frac{\alpha-\beta}{\sigma},\ -Z(t)\ge\frac{\alpha-x}{\sigma}\right). \end{aligned}$$

Then by Girsanov theorem, $Z(t)$ is a Brownian motion under the probability measure Q, where

$$dQ = \exp\left(-\frac{\mu}{\sigma}W_t - \frac{\mu^2}{2\sigma^2}t\right)dP.$$

Note that

$$\begin{aligned} dP &= \exp\left(\frac{\mu}{\sigma}W_t + \frac{\mu^2}{2\sigma^2}t\right)dQ \\ &= \exp\left(\frac{\mu}{\sigma}Z_t - \frac{\mu^2}{2\sigma^2}t\right)dQ \end{aligned}$$

Let $a=\frac{\alpha-x}{\sigma}$ and $b=\frac{\alpha-\beta}{\sigma}$. Thus we have that

$$P(\beta < X_\beta(t)\le x) = \int_a^b\int_z^b \exp\left(\frac{\mu}{\sigma}z-\frac{\mu^2}{2\sigma^2}t\right)f(y,z)dydz,$$

where $f(y,z)$ is the joint density function of $\left(M_{-Z}(t),-Z(t)\right)$ under measure Q. Since $\left(M_{-Z}(t),-Z(t)\right)$ and $\left(M_Z(t),Z(t)\right)$ have the same distribution, by theorem (8.1.4),

$$f(y,z) = \sqrt{\frac{2}{\pi t^3}}(2y-z)\exp\big(-\frac{(2y-z)^2}{2t}\big).$$

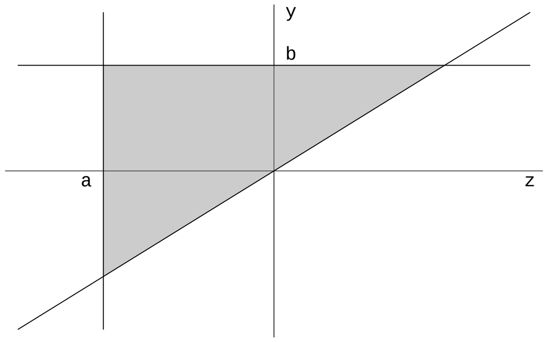


Figure 6.1: Integration region

The definite integration w.r.t. y gives us

$$P(\beta < X_\beta(t) \le x) = \frac{1}{\sqrt{2\pi t}} \int_a^b \exp\left(\frac{\mu}{\sigma}z - \frac{\mu^2}{2\sigma^2}t\right)\left[\exp\left(-\frac{z^2}{2t}\right) - \exp\left(-\frac{(z-2b)^2}{2t}\right)\right] dz.$$

Substituting $\frac{\alpha - w}{\sigma}$ into z yields

$$\begin{aligned}
\exp\left(\frac{\mu}{\sigma}z - \frac{\mu^2}{2\sigma^2}t\right)\exp\left(-\frac{z^2}{2t}\right) &= \exp\left(-\frac{2\mu w - 2\mu\alpha + \mu^2 t}{2\sigma^2}\right)\exp\left(-\frac{(w-\alpha)^2}{2t\sigma^2}\right) \\
&= \exp\left(-\frac{(w-\alpha-t\mu)^2}{2t\sigma^2}\right), \quad \text{and} \\
\exp\left(\frac{\mu}{\sigma}z - \frac{\mu^2}{2\sigma^2}t\right)\exp\left(-\frac{(z-2b)^2}{2t}\right) &= \exp\left(-\frac{2\mu w - 2\mu\alpha + \mu^2 t}{2\sigma^2}\right)\exp\left(-\frac{(w+\alpha-2\beta)^2}{2t\sigma^2}\right) \\
&= \exp\left(-\frac{(w+\alpha-2\beta-\mu t)^2}{2t\sigma^2}\right)\exp\left(-\frac{2\mu(\alpha-\beta)}{\sigma^2}\right).
\end{aligned}$$

Thus we have

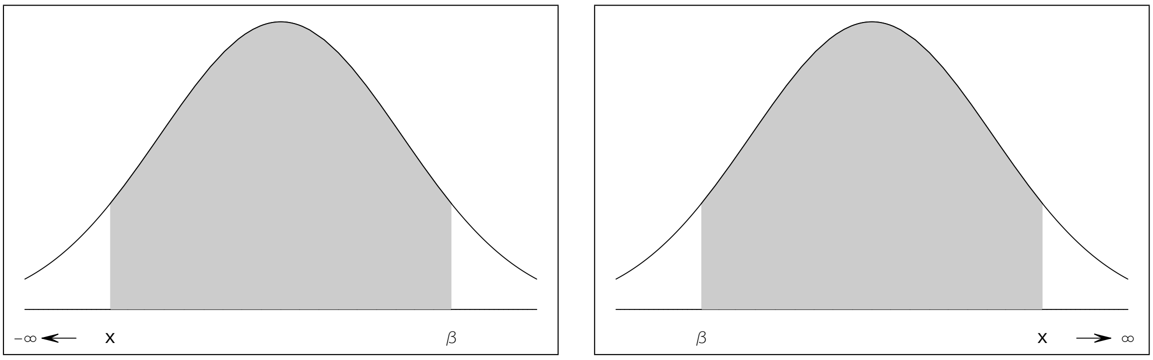
$$\begin{aligned}
& P(\beta < X_\beta(t) \le x) \\
= \; & \frac{1}{\sqrt{2\pi t}\,\sigma}\int_*^x \left[\exp\left(-\frac{(w-\alpha-t\mu)^2}{2t\sigma^2}\right) - \exp\left(-\frac{(w+\alpha-2\beta-\mu t)^2}{2t\sigma^2}\right)\exp\left(-\frac{2\mu(\alpha-\beta)}{\sigma^2}\right)\right] dw \\
= \; & \int_*^x \varphi(w;\mu t+\alpha,\sigma^2 t) - \exp\left(-\frac{2\mu(\alpha-\beta)}{\sigma^2}\right)\varphi(w;\mu t-\alpha+2\beta,\sigma^2 t)dw.
\end{aligned}$$

The theorem can be proved now by differentiating this with respect to x.

■

Corollary 6.1.3. The distribution functions for M_X and m_X are given by the following equations, which hold for $x \ge \alpha$ and $x \le \alpha$ respectively.

$$\begin{aligned}
F_M(x) &= N\left(\frac{x-\alpha-\mu t}{\sigma\sqrt{t}}\right) - \exp\left(\frac{2\mu(x-\alpha)}{\sigma^2}\right)N\left(-\frac{x-\alpha+\mu t}{\sigma\sqrt{t}}\right), \qquad (6.1.2) \\
F_m(x) &= N\left(\frac{x-\alpha-\mu t}{\sigma\sqrt{t}}\right) + \exp\left(\frac{2\mu(x-\alpha)}{\sigma^2}\right)N\left(\frac{x-\alpha+\mu t}{\sigma\sqrt{t}}\right).
\end{aligned}$$



(a) M_X (b) m_X

Figure 6.2: The absorbed process

Proof. The probability of $M_X(t) \le \beta$ is given by

$$\begin{aligned}
P(M_X(t) \le \beta) &= \int_{-\infty}^{\beta} \left[\varphi(x, \mu t + \alpha, \sigma^2 t) - \exp\left(-\frac{2\mu(\alpha - \beta)}{\sigma^2} \right) \varphi(x; \mu t - \alpha + 2\beta, \sigma^2 t)) \right] dx \\
&= \int_{-\infty}^{\beta} \varphi(x, \mu t + \alpha, \sigma^2 t) dx - \exp\left(-\frac{2\mu(\alpha - \beta)}{\sigma^2} \right) \int_{-\infty}^{\beta} \varphi(x; \mu t - \alpha + 2\beta, \sigma^2 t)) dx \\
&= \int_{-\infty}^{\frac{\beta - \mu t - \alpha}{\sigma\sqrt{t}}} \varphi(x, 0, 1) dx - \exp\left(-\frac{2\mu(\alpha - \beta)}{\sigma^2} \right) \int_{-\infty}^{\frac{-\beta - \mu t + \alpha}{\sigma\sqrt{t}}} \varphi(x; 0, 1)) dx \\
&= N\left(\frac{\beta - \mu t - \alpha}{\sigma\sqrt{t}} \right) - \exp\left(-\frac{2\mu(\alpha - \beta)}{\sigma^2} \right) N\left(\frac{-\beta - \mu t + \alpha}{\sigma\sqrt{t}} \right), \\
P(M_X(t) \le x) &= N\left(\frac{x - \mu t - \alpha}{\sigma\sqrt{t}} \right) - \exp\left(\frac{2\mu(x - \alpha)}{\sigma^2} \right) N\left(-\frac{x - \alpha + \mu t}{\sigma\sqrt{t}} \right).
\end{aligned}$$

Similarly, for m_X, we have

$$\begin{aligned}
P(m_X(t) > \beta) &= \int_{\beta}^{\infty} \left[\varphi(x, \mu t + \alpha, \sigma^2 t) - \exp\left(-\frac{2\mu(\alpha - \beta)}{\sigma^2} \right) \varphi(x; \mu t - \alpha + 2\beta, \sigma^2 t)) \right] dx \\
&= \int_{\beta}^{\infty} \varphi(x, \mu t + \alpha, \sigma^2 t) dx - \exp\left(-\frac{2\mu(\alpha - \beta)}{\sigma^2} \right) \int_{\beta}^{\infty} \varphi(x; \mu t - \alpha + 2\beta, \sigma^2 t)) dx \\
&= \int_{\frac{\beta - \mu t - \alpha}{\sigma\sqrt{t}}}^{\infty} \varphi(x, 0, 1) dx - \exp\left(-\frac{2\mu(\alpha - \beta)}{\sigma^2} \right) \int_{\frac{-\beta - \mu t + \alpha}{\sigma\sqrt{t}}}^{\infty} \varphi(x; 0, 1)) dx \\
&= 1 - N\left(\frac{\beta - \mu t - \alpha}{\sigma\sqrt{t}} \right) - \exp\left(-\frac{2\mu(\alpha - \beta)}{\sigma^2} \right) N\left(\frac{\beta + \mu t - \alpha}{\sigma\sqrt{t}} \right), \\
P(m_X(t) \le x) &= N\left(\frac{x - \mu t - \alpha}{\sigma\sqrt{t}} \right) + \exp\left(\frac{2\mu(x - \alpha)}{\sigma^2} \right) N\left(\frac{x - \alpha + \mu t}{\sigma\sqrt{t}} \right).
\end{aligned}$$

■

Remark 6.1.4. Note that $F_M(\alpha) = 0$, $F_m(\alpha) = 1$ and

$$P\Big(M_X(t) = X(0) \Big) = 0.$$

By differentiating $F_M(x)$, we obtain the PDF for the M_X. Thus we have the following corollary.

Corollary 6.1.5. The probability density functions of M_X and m_X are given by.

$$\begin{aligned} f_{M_X}(x) &= \varphi(x;\alpha+\mu t,\sigma^2 t) - \frac{2\mu}{\sigma^2}\exp\left(\frac{2\mu(x-\alpha)}{\sigma^2}\right) N\left(-\frac{x-\alpha+\mu t}{\sigma\sqrt{t}}\right) \\ &\quad + \exp\left(\frac{2\mu(x-\alpha)}{\sigma^2}\right)\varphi(x;\alpha-\mu t,\sigma^2 t), \quad \alpha \le x < \infty \\ f_{m_X}(x) &= \varphi(x;\alpha+\mu t,\sigma^2 t) + \frac{2\mu}{\sigma^2}\exp\left(\frac{2\mu(x-\alpha)}{\sigma^2}\right) N\left(-\frac{x-\alpha+\mu t}{\sigma\sqrt{t}}\right) \\ &\quad + \exp\left(\frac{2\mu(x-\alpha)}{\sigma^2}\right)\varphi(x;\alpha-\mu t,\sigma^2 t), \quad -\infty < x \le \alpha. \end{aligned}$$

Example 6.1.6. Let $X_t = x + W_t$ be a linear Brownian motion starting at $x \in \mathbb{R}$. The PDF of $\inf_{0\le s\le t} X_s$ is given by

$$f(y) = \frac{2}{\sqrt{2\pi t}}\exp\left(-\frac{(y-x)}{2t}\right), \quad -\infty < y \le x.$$

□

Theorem 6.1.7 (Upper Barrier: Distribution of Hitting Time). For $b > X(0)(=\alpha)$, define

$$T_b := \inf\{t : X_t = b\} = \inf\{t : M_X(t) = b\}.$$

Then we have the probability density function of T_b, denoted by f_{T_b} as follows:

$$f_{T_b}(t) = \frac{b-\alpha}{\sqrt{2\pi t^3}\,\sigma}\exp\left(-\frac{(\mu t - b + \alpha)^2}{2t\sigma^2}\right). \tag{6.1.3}$$

Proof. Since $P(T_b \le t) = P\Big(M_X(t) \ge b\Big)$, we have

$$\begin{aligned} P(T_b \le t) &= 1 - P\Big(M_X(t) < b\Big) \\ &= 1 - F_M(b). \end{aligned}$$

Differentiation with respect to t gives us the result. ■

Similarly we have the following theorem.

Theorem 6.1.8 (Lower Barrier: Distribution of Hitting Time). For $a < X(0)(=\alpha)$, define

$$T_a := \inf\{t : X_t = a\} = \inf\{t : m_X(t) = a\}.$$

Then we have the probability density function of T_a, denoted by f_{T_a} as follows:

$$f_{T_a}(t) = \frac{\alpha-a}{\sqrt{2\pi t^3}\,\sigma}\exp\left(-\frac{(\mu t - a + \alpha)^2}{2t\sigma^2}\right). \tag{6.1.4}$$

Proof. The proof is straightforward from the fact that

$$P(T_a \le t) = P\Big(m_X(t) \le a\Big).$$

■

Remark 6.1.9 (The Inverse Gaussian Distribution). Suppose a particle moving along a line tends to move with a uniform velocity v. Suppose also, that the particle subject to linear Brownian motion which causes it to take a variable amount of time to cover a fixed distant(d). It can be shown that the time, X, required to cover the distance is a random variable with probability density function

$$P_X(x) = \frac{d}{\sqrt{2\pi\beta x^3}} e^{-\frac{(vx-d)^2}{2\beta x}}, \quad x > 0,$$

where β is a diffusion constant.

6.2 Double Barrier Hitting Time Distribution of a BM with Drift

In this section we consider three barrier hitting times [1]. We assume that there is an upper barrier and a lower barrier. Let $\{W_t, t \geq 0\}$ be a standard Brownian motion and X_t be a Brownian motion with drift $\mu > 0$ and volatility $\sigma > 0$, i.e.

$$X(t) \quad = \quad \mu t + \sigma W_t.$$

Let $a < 0 < b$. Define

1. $T_{a,b}$ is defined as the first hitting time that X_t hits one of the two barriers. i.e.

$$T_{a,b} = \begin{cases} \inf\{t : X_t = a \text{ or } X_t = b\}, & \text{if such } t \text{ exists,} \\ \infty & \text{if } X_t \text{ never hits the barrier.} \end{cases}$$

Thus we can see that

$$T_{a,b} \quad = \quad \min\{T_a, T_b\}, \quad \text{or} \quad T_a \wedge T_b.$$

2. $T_{a,\tilde{b}}$ is defined as the first hitting time that X_t hits the lower barrier without hitting the upper barrier earlier. i.e.

$$T_{a,\tilde{b}} = \inf\{t : X(t) = a, T_a < T_b\}.$$

3. $T_{\tilde{a},b}$ is defined as the first hitting time that X_t hits the upper barrier without hitting the lower barrier earlier. i.e.

$$T_{\tilde{a},b} = \inf\{t : X(t) = b, T_b < T_a\}.$$

4. $f(t; a)$(drift$=0$): The density function of hitting time distribution to the standard Brownian motion with barrier $a > 0$ is given by

$$f(t; a) \quad = \quad \frac{a}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}}, \quad t > 0, a > 0.$$

The Laplace transform of $f(t; a)$ is given by

$$\int_0^\infty e^{-zt} f(t; a)dt \quad = \quad e^{-a\sqrt{2z}}, \quad z > 0, a > 0.$$

5. Consider a more general function f_{IG}(drift > 0). The density function of an inverse Gaussian distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$ is given by

$$f_{IG}(t; \alpha, \beta) \quad := \quad \frac{\alpha}{\sqrt{2\pi\beta t^3}} e^{-\frac{(\beta t-\alpha)^2}{2\beta t}}, \quad t > 0. \tag{6.2.1}$$

This is a special case of general inverse Gaussian. See Remark(6.1.9). Also, The Laplace transform of $f_{IG}(t)$ is given by

$$\int_0^\infty e^{-zt} f_{IG}(t; \alpha, \beta)dt \quad = \quad e^{\alpha\left(1-\sqrt{1+\frac{2z}{\beta}}\right)}, \quad z > 0, \alpha > 0, \beta > 0. \tag{6.2.2}$$

[1] See Lin(1998).

Example 6.2.1.

1. The cumulative distribution function, denoted by F_{IG}: By Theorem(6.1.7) and Corollary(6.1.3), we obtain

$$
\begin{aligned}
F_{IG}(t;\alpha,\beta) &:= \int_0^t f_{IG}(y;\alpha,\beta)dy \\
&= N\left(\frac{\beta t-\alpha}{\sqrt{\beta t}}\right)+e^{2\alpha}N\left(-\frac{\beta t+\alpha}{\sqrt{\beta t}}\right), \quad \alpha>0,\beta>0.
\end{aligned}
$$

(We set $\mu \leftarrow \beta$, $\sigma \leftarrow \sqrt{\beta}$, $\alpha \leftarrow 0$ and $x \leftarrow \alpha$ in (6.1.2). $1-F$ is F_{IG}.)

2. $\int_0^T e^{-zt}f(t;a),\ z>0$: To calculate this integral, we write the integrand as

$$
\begin{aligned}
e^{-zt}f(t;a) &= e^{-zt}\frac{a}{\sqrt{2\pi t^3}}e^{-\frac{a^2}{2t}} \\
&= e^{-a\sqrt{2z}}\frac{a}{\sqrt{2\pi t^3}}e^{-\frac{(2zt-a\sqrt{2z})^2}{4zt}} \\
&= e^{-a\sqrt{2z}}\frac{a\sqrt{2z}}{\sqrt{4\pi z t^3}}e^{-\frac{(2zt-a\sqrt{2z})^2}{4zt}} \\
&= e^{-a\sqrt{2z}}f_{IG}(t;a\sqrt{2z},2z).
\end{aligned}
$$

Hence we obtain

$$
\begin{aligned}
\int_0^T e^{-zt}f(t;a)dt &= e^{-a\sqrt{2z}}F_{IG}(T;a\sqrt{2z},2z) \\
&= e^{-a\sqrt{2z}}N\left(\frac{\sqrt{2z}\,T-a}{\sqrt{T}}\right)+e^{a\sqrt{2z}}N\left(-\frac{\sqrt{2z}\,T+a}{\sqrt{T}}\right), \quad a>0. \quad (6.2.3)
\end{aligned}
$$

□

The equation (6.2.3) is valid when $a>0$. For $a<0$, we define

$$
f(t;a) = -f\big(t;|a|\big) \ (<0).
$$

Thus, for $a<0$, we have

$$
\begin{aligned}
\int_0^T e^{-zt}f(t;a) &= -e^{-|a|\sqrt{2z}}F_{IG}(T;|a|\sqrt{2z},2z) \\
&= -e^{-|a|\sqrt{2z}}N\left(\frac{\sqrt{2z}\,T-|a|}{\sqrt{T}}\right)-e^{|a|\sqrt{2z}}N\left(-\frac{\sqrt{2z}\,T+|a|}{\sqrt{T}}\right) \\
&= -e^{a\sqrt{2z}}N\left(\frac{\sqrt{2z}\,T+a}{\sqrt{T}}\right)-e^{-a\sqrt{2z}}N\left(-\frac{\sqrt{2z}\,T-a}{\sqrt{T}}\right), \quad a<0. \quad (6.2.4)
\end{aligned}
$$

In the followings, we will identify the defective densities function of $T_{a,\tilde{b}}$ and $T_{\tilde{a},b}$. Note that the density of $T_{a,b}$ is the sum of these two functions.

1. We want to obtain $E[e^{-zT_{a,\bar{b}}}]$ and $E[e^{-zT_{\bar{a},b}}]$. First, find a real λ so that the exponent
$$e^{\lambda X(t)-zt},$$
be a martingale for a given fixed $z > 0$. Since
$$Z_\lambda(t) \quad := \quad e^{\lambda X(t)-\lambda\mu t-\frac{1}{2}\lambda^2\sigma^2 t}$$
is a exponent martingale, it is enough to find λ such that
$$z \quad = \quad \mu\lambda + \frac{1}{2}\lambda^2\sigma^2.$$
Then, we get the following two λ's, say λ_1 and λ_2.
$$\lambda_1 \quad = \quad \frac{-\mu-\sqrt{\mu^2+2\sigma^2 z}}{\sigma^2}, \quad \lambda_2 \quad = \quad \frac{-\mu+\sqrt{\mu^2+2\sigma^2 z}}{\sigma^2}.$$
We obtain two martingales:
$$M_1(t) \quad := \quad Z_{\lambda_1}(t) \quad = \quad e^{\lambda_1 X(t)-zt}, \quad \text{and} \quad M_2(t) \quad := \quad Z_{\lambda_2}(t) \quad = \quad e^{\lambda_2 X(t)-zt}.$$

2. Note that $T_{a,b}$ is a stopping time. Thus, by the *Optional Sampling Theorem* we have
$$E[M_i(T_{a,b})] \quad = \quad E[M_i(0)] \quad = \quad 1, \quad i=1,2.$$
It follows from the law of total probability that
$$\begin{aligned} E[M_i(T_{a,b})] &= E[e^{a\lambda_i - zT_{a,\bar{b}}}] + E[e^{b\lambda_i - zT_{\bar{a},b}}] \\ &= e^{\lambda_i a}E[e^{-zT_{a,\bar{b}}}] + e^{\lambda_i b}E[e^{-zT_{\bar{a},b}}] \\ &= 1. \end{aligned}$$
Hence we have the following linear equations.
$$\begin{aligned} e^{\lambda_1 a}E[e^{-zT_{a,\bar{b}}}] + e^{\lambda_1 b}E[e^{-zT_{\bar{a},b}}] &= 1, \\ e^{\lambda_2 a}E[e^{-zT_{a,\bar{b}}}] + e^{\lambda_2 b}E[e^{-zT_{\bar{a},b}}] &= 1. \end{aligned}$$
Solving the linear equations above yields the Laplace transforms
$$\begin{aligned} E[e^{-zT_{a,\bar{b}}}] &= \frac{e^{b\lambda_2}-e^{b\lambda_1}}{e^{a\lambda_1+b\lambda_2}-e^{b\lambda_1+a\lambda_2}}, \\ E[e^{-zT_{\bar{a},b}}] &= \frac{e^{a\lambda_1}-e^{a\lambda_2}}{e^{a\lambda_1+b\lambda_2}-e^{b\lambda_1+a\lambda_2}}. \end{aligned}$$
Since $e^{-(b-a)(\lambda_2-\lambda_1)} < 1$ and
$$\begin{aligned} \frac{1}{e^{a\lambda_1+b\lambda_2}-e^{b\lambda_1+a\lambda_2}} &= e^{-a\lambda_1-b\lambda_2}\frac{1}{1-e^{-(b-a)(\lambda_2-\lambda_1)}} \\ &= e^{-a\lambda_1-b\lambda_2}\sum_{n=0}^{\infty} e^{-n(b-a)(\lambda_2-\lambda_1)}, \end{aligned}$$
we have
$$E[e^{-zT_{a,\bar{b}}}] \quad = \quad (e^{b\lambda_2}-e^{b\lambda_1})e^{-a\lambda_1-b\lambda_2}\sum_{n=0}^{\infty} e^{-n(b-a)(\lambda_2-\lambda_1)}$$

$$
\begin{aligned}
&= \left(e^{-a\lambda_1} - e^{(b-a)\lambda_1 - b\lambda_2}\right)\sum_{n=0}^{\infty} e^{-n(b-a)(\lambda_2-\lambda_1)} \\
&= \left(e^{-a\lambda_1} - e^{(b-a)\lambda_1 - b\lambda_2}\right)\sum_{n=0}^{\infty} e^{-n(b-a)(\lambda_2-\lambda_1)} \\
&= e^{-a\lambda_1}\sum_{n=0}^{\infty} e^{-n(b-a)(\lambda_2-\lambda_1)} - e^{(b-a)\lambda_1 - b\lambda_2}\sum_{n=0}^{\infty} e^{-n(b-a)(\lambda_2-\lambda_1)} \\
&= \sum_{n=0}^{\infty} e^{-n(b-a)(\lambda_2-\lambda_1)-\lambda_1 a} - e^{-(b-a)(\lambda_2-\lambda_1)-\lambda_2 a}\sum_{n=0}^{\infty} e^{-n(b-a)(\lambda_2-\lambda_1)} \\
&= \sum_{n=0}^{\infty} e^{-n(b-a)(\lambda_2-\lambda_1)-a\lambda_1} - \sum_{n=0}^{\infty} e^{-(n+1)(b-a)(\lambda_2-\lambda_1)-a\lambda_2}, \qquad (6.2.5)
\end{aligned}
$$

and similarly

$$
E[e^{-zT_{\tilde{a},b}}] = \sum_{n=0}^{\infty} e^{-n(b-a)(\lambda_2-\lambda_1)-b\lambda_2} - \sum_{n=0}^{\infty} e^{-(n+1)(b-a)(\lambda_2-\lambda_1)-b\lambda_1}. \qquad (6.2.6)
$$

3. For the notational convenience, we let

$$
\begin{aligned}
a_n &= \frac{2n(b-a)-a}{\sigma}, \quad \text{and} \\
b_n &= \frac{2n(b-a)+b}{\sigma}, \quad n \in \mathbb{Z}.
\end{aligned}
$$

The terms of the first series in (6.2.5) can be rewritten as

$$
\begin{aligned}
& e^{-n(b-a)(\lambda_2-\lambda_1)-a\lambda_1} \\
&= \exp\left(-n(b-a)\,\frac{2\sqrt{\mu^2+2\sigma^2 z}}{\sigma^2} - a\,\frac{-\mu-\sqrt{\mu^2+2\sigma^2 z}}{\sigma^2}\right) \\
&= \exp\left(-(\sigma a_n + a)\,\frac{\sqrt{\mu^2+2\sigma^2 z}}{\sigma^2} + \frac{a\mu}{\sigma^2} + a\frac{\sqrt{\mu^2+2\sigma^2 z}}{\sigma^2}\right) \\
&= \exp\left(\frac{a\mu}{\sigma^2} - \frac{a_n\mu}{\sigma}\sqrt{1+\frac{2\sigma^2 z}{\mu^2}}\right) \\
&= \exp\left\{\frac{a\mu}{\sigma^2} - \frac{a_n\mu}{\sigma}\right\}\exp\left\{\frac{a_n\mu}{\sigma}\left(1-\sqrt{1+\frac{2\sigma^2 z}{\mu^2}}\right)\right\}.
\end{aligned}
$$

From (6.2.2), the second factor above is the Laplace transform of the inverse Gaussian distribution with parameters $\alpha = \frac{a_n\mu}{\sigma}$ and $\beta = \frac{\mu^2}{\sigma^2}$. Note that $\alpha > 0$ and $\beta > 0$ if $n \geq 0$. Thus from (6.2.1), we can see that the corresponding density functions are

$$
\begin{aligned}
& \exp\left\{\frac{a\mu}{\sigma^2} - \frac{a_n\mu}{\sigma}\right\}\frac{\alpha}{\sqrt{2\pi\beta t^3}}e^{-\frac{(\beta t-\alpha)^2}{2\beta t}} \\
&= \exp\left\{\frac{a\mu}{\sigma^2} - \frac{a_n\mu}{\sigma}\right\}\frac{a_n}{\sqrt{2\pi t^3}}\exp\left\{-\frac{(\mu t - a_n\sigma)^2}{2t\sigma^2}\right\} \\
&= \exp\left\{\frac{a\mu}{\sigma^2} - \frac{1}{2}\frac{\mu^2}{\sigma^2}t\right\}\frac{a_n}{\sqrt{2\pi t^3}}\exp\left\{-\frac{a_n^2}{2t}\right\}
\end{aligned}
$$

$$= \exp\left\{\frac{a\mu}{\sigma^2} - \frac{1}{2}\frac{\mu^2}{\sigma^2}t\right\} f(t; a_n), \quad n = 0, 1, 2, \cdots . \tag{6.2.7}$$

For the terms of the second series in (6.2.5), let $m = -(n+1)$. Then we have

$$\begin{aligned}
& e^{-(n+1)(b-a)(\lambda_2-\lambda_1)-a\lambda_2} \\
= \; & e^{m(b-a)(\lambda_2-\lambda_1)-a\lambda_2} \\
= \; & \exp\left((\sigma a_m + a)\frac{\sqrt{\mu^2+2\sigma^2 z}}{\sigma^2} + \frac{a\mu}{\sigma^2} - a\frac{\sqrt{\mu^2+2\sigma^2 z}}{\sigma^2}\right) \\
= \; & \exp\left(\frac{a\mu}{\sigma^2} + \frac{a_m\mu}{\sigma}\sqrt{1+\frac{2\sigma^2 z}{\mu^2}}\right) \\
= \; & \exp\left\{\frac{a\mu}{\sigma^2} + \frac{a_m\mu}{\sigma}\right\}\exp\left\{-\frac{a_m\mu}{\sigma}\left(1-\sqrt{1+\frac{2\sigma^2 z}{\mu^2}}\right)\right\}.
\end{aligned}$$

Thus the functions corresponding to the terms of the second series in (6.2.5) are

$$\begin{aligned}
& \exp\left\{\frac{a\mu}{\sigma^2} - \frac{1}{2}\frac{\mu^2}{\sigma^2}t\right\} f(t; -a_m) \\
= \; & \exp\left\{\frac{a\mu}{\sigma^2} - \frac{1}{2}\frac{\mu^2}{\sigma^2}t\right\} f(t; |a_m|) \\
= \; & -\exp\left\{\frac{a\mu}{\sigma^2} - \frac{1}{2}\frac{\mu^2}{\sigma^2}t\right\} f(t; a_m)\,(>0), \quad m = -1, -2, \cdots .
\end{aligned} \tag{6.2.8}$$

Together with (6.2.7) and (6.2.8), we obtain the density function of $T_{a,\tilde{b}}$:

$$\begin{aligned}
f_{T_{a,\tilde{b}}}(t) \; &:= \; \exp\left(\frac{a\mu}{\sigma^2} - \frac{1}{2}\frac{\mu^2}{\sigma^2}t\right)\sum_{n=-\infty}^{\infty} f(t; a_n) \qquad (6.2.9) \\
&= \; \exp\left(\frac{a\mu}{\sigma^2} - \frac{1}{2}\frac{\mu^2}{\sigma^2}t\right)\sum_{n=-\infty}^{\infty}\frac{2n(b-a)-a}{\sqrt{2\pi t^3}\,\sigma}\exp\left(-\frac{\big(2n(b-a)-a\big)^2}{2t\sigma^2}\right).
\end{aligned}$$

Similarly argument yields the density function of $T_{\tilde{a},b}$:

$$\begin{aligned}
f_{T_{\tilde{a},b}}(t) \; &:= \; \exp\left(\frac{b\mu}{\sigma^2} - \frac{1}{2}\frac{\mu^2}{\sigma^2}t\right)\sum_{n=-\infty}^{\infty} f(t; b_n) \qquad (6.2.10) \\
&= \; \exp\left(\frac{b\mu}{\sigma^2} - \frac{1}{2}\frac{\mu^2}{\sigma^2}t\right)\sum_{n=-\infty}^{\infty}\frac{2n(b-a)+b}{\sqrt{2\pi t^3}\,\sigma}\exp\left(-\frac{\big(2n(b-a)+b\big)^2}{2t\sigma^2}\right).
\end{aligned}$$

Note that half of the coefficients in each series are negative.

Example 6.2.2. Consider the limit

$$\lim_{b\to\infty} f_{T_{a,\tilde{b}}}, \quad \lim_{a\to-\infty} f_{T_{\tilde{a},b}}(t).$$

Note that

$$\lim_{b\to\infty} f(t; a_n) \; = \; 0, \quad \text{if } n \neq 0.$$

Thus we obtain

$$\begin{aligned}\lim_{b\to\infty} f_{T_{a,\bar{b}}} &= \exp\left(\frac{a\mu}{\sigma^2} - \frac{1}{2}\frac{\mu^2}{\sigma^2}t\right)\frac{-a}{\sqrt{2\pi t^3}\,\sigma}\exp\left(-\frac{a^2}{2t\sigma^2}\right)\\ &= \frac{-a}{\sqrt{2\pi t^3}\,\sigma}\exp\left(-\frac{(\mu t - a)^2)}{2t\sigma^2}\right).\end{aligned}$$

This result is identical to Theorem(6.1.8). Similarly, we have

$$\lim_{a\to-\infty} f(t;b_n) = 0, \quad \text{if } n \neq 0.$$

Hence we have

$$\begin{aligned}\lim_{a\to-\infty} f_{T_{\bar{a},b}}(t) &= \exp\left(\frac{b\mu}{\sigma^2} - \frac{1}{2}\frac{\mu^2}{\sigma^2}t\right)\frac{b}{\sqrt{2\pi t^3}\,\sigma}\exp\left(-\frac{b^2}{2t\sigma^2}\right)\\ &= \frac{b}{\sqrt{2\pi t^3}\,\sigma}\exp\left(-\frac{(\mu t - b)^2)}{2t\sigma^2}\right).\end{aligned}$$

□

6.3 Double Barrier Hitting Time Distribution of a GBM

The geometric Brownian motion is given by

$$S(t) = S\exp\left(\mu t + \sigma W(t)\right)$$

where S is initial value at time $t = 0$ and $\mu = r - \frac{\sigma^2}{2}$. We assume that stock pays no dividends.

For the given lower(upper) barrier $L(U)$ with $0 < L < S < U$, denote $\tau_L(\tau_U)$ as the first hitting time to the lower(upper) barrier without hitting the upper(lower) barrier earlier, i.e.

$$\begin{aligned}\tau_L &= \inf\{t; S(t) = L, L < S(u) < U \text{ for all } s < t\},\\ \tau_U &= \inf\{t; S(t) = U, L < S(u) < U \text{ for all } s < t\}.\end{aligned}$$

Let

$$\begin{aligned}a &= \log\frac{L}{S} < 0,\\ b &= \log\frac{U}{S} > 0,\\ g_L(t,S) &= \text{PDF of } \tau_L,\\ g_U(t,S) &= \text{PDF of } \tau_U.\end{aligned}$$

Then the density $g_L(t;S)$ of τ_L is $f_{T_{a,\bar{b}}}(t)$. Here

$$\begin{aligned}a_n &= \frac{2n(b-a)-a}{\sigma} = \frac{1}{\sigma}\log\left\{\left(\frac{U}{L}\right)^{2n}\frac{S}{L}\right\} = \frac{1}{\sigma}\log\frac{U^{2n}S}{L^{2n+1}},\\ b_n &= \frac{2n(b-a)+b}{\sigma} = \frac{1}{\sigma}\log\left\{\left(\frac{U}{L}\right)^{2n}\frac{U}{S}\right\} = \frac{1}{\sigma}\log\frac{U^{2n+1}}{L^{2n}S}.\end{aligned}$$

Note that

$$(a_n > 0 \iff n \geq 0), \quad (b_n > 0 \iff n \geq 0).$$

It follows from (6.2.9) and (6.2.10)that

$$\begin{aligned}
g_L(t;S) &= \exp\left(\frac{a\mu}{\sigma^2} - \frac{1}{2}\frac{\mu^2}{\sigma^2}t\right)\sum_{n=-\infty}^{\infty} f(t;a_n) \\
&= \left(\frac{L}{S}\right)^{\frac{\mu}{\sigma^2}} \exp\left(-\frac{1}{2}\frac{\mu^2}{\sigma^2}t\right)\sum_{n=-\infty}^{\infty} f(t;a_n), \\
g_U(t;S) &= \exp\left(\frac{b\mu}{\sigma^2} - \frac{1}{2}\frac{\mu^2}{\sigma^2}t\right)\sum_{n=-\infty}^{\infty} f(t;b_n) \\
&= \left(\frac{U}{S}\right)^{\frac{\mu}{\sigma^2}} \exp\left(-\frac{1}{2}\frac{\mu^2}{\sigma^2}t\right)\sum_{n=-\infty}^{\infty} f(t;b_n).
\end{aligned}$$

6.3.1 Four Basic integrals

In this subsection we will computer the following four integrals:

1. discounted American down and in - up and out bond (hitting the lower barrier without hitting the upper barrier earlier).
$$\int_0^T e^{-rt} g_L(t;S)dt.$$
2. discounted American up and in - down and out bond. (hitting the upper barrier without hitting the lower barrier earlier).
$$\int_0^T e^{-rt} g_U(t;S)dt.$$
3. un-discounted European down and in- up and out bond. (hitting the lower barrier without hitting the upper barrier earlier).
$$\int_0^T g_L(t;S)dt.$$
4. un-discounted European up and in - down and out bond. (hitting the upper barrier without hitting the lower barrier earlier).
$$\int_0^T g_U(t;S)dt.$$

In the following, we assume that $q = 0$.

1. $\int_0^T e^{-rt} g_L(t;S)dt$:

$$\begin{aligned}
\int_0^T e^{-rt} g_L(t;S)dt &= \left(\frac{L}{S}\right)^{\frac{\mu}{\sigma^2}} \int_0^T e^{-rt} \exp\left(-\frac{1}{2}\frac{\mu^2}{\sigma^2}t\right)\sum_{n=-\infty}^{\infty} f(t;a_n)dt \\
&= \left(\frac{L}{S}\right)^{\frac{\mu}{\sigma^2}} \sum_{n=-\infty}^{\infty}\int_0^T \exp\left(-\frac{(r+\frac{1}{2}\sigma^2)^2}{2\sigma^2}t\right) f(t;a_n)dt.
\end{aligned}$$

By (6.2.3) and (6.2.4) and we have

$$\begin{aligned}\int_0^T e^{-rt} g_L(t;S)dt &= \left(\frac{L}{S}\right)^{\frac{\mu}{\sigma^2}} \sum_{n=0}^{\infty} \left[e^{-a_n\sqrt{2z}} N\left(\frac{\sqrt{2z}\,T - a_n}{\sqrt{T}}\right) + e^{a_n\sqrt{2z}} N\left(-\frac{\sqrt{2z}\,T + a_n}{\sqrt{T}}\right)\right] \\ &\quad - \left(\frac{L}{S}\right)^{\frac{\mu}{\sigma^2}} \sum_{n=-\infty}^{-1} \left[e^{a_n\sqrt{2z}} N\left(\frac{\sqrt{2z}\,T + a_n}{\sqrt{T}}\right) + e^{-a_n\sqrt{2z}} N\left(-\frac{\sqrt{2z}\,T - a_n}{\sqrt{T}}\right)\right],\end{aligned}$$

where $z = \frac{(r+\frac{1}{2}\sigma^2)^2}{2\sigma^2}$. Now we have

$$\begin{aligned}&\int_0^T e^{-rt} g_L(t;S)dt \\ &= \left(\frac{L}{S}\right)^{\frac{\mu}{\sigma^2}} \sum_{n=0}^{\infty} \left[e^{-a_n\left(\frac{r+\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(\frac{(r+\frac{1}{2}\sigma^2)\,T - \sigma a_n}{\sigma\sqrt{T}}\right) + e^{a_n\left(\frac{r+\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(-\frac{(r+\frac{1}{2}\sigma^2)\,T + \sigma a_n}{\sigma\sqrt{T}}\right)\right] \\ &\quad - \left(\frac{L}{S}\right)^{\frac{\mu}{\sigma^2}} \sum_{n=-\infty}^{-1} \left[e^{a_n\left(\frac{r+\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(\frac{(r+\frac{1}{2}\sigma^2)\,T + \sigma a_n}{\sigma\sqrt{T}}\right) + e^{-a_n\left(\frac{r+\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(-\frac{(r+\frac{1}{2}\sigma^2)\,T - \sigma a_n}{\sigma\sqrt{T}}\right)\right].\end{aligned}$$

2. $\int_0^T e^{-rt} g_U(t;S)dt$: Similarly we obtain

$$\begin{aligned}&\int_0^T e^{-rt} g_U(t;S)dt \\ &= \left(\frac{U}{S}\right)^{\frac{\mu}{\sigma^2}} \sum_{n=0}^{\infty} \left[e^{-b_n\left(\frac{r+\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(\frac{(r+\frac{1}{2}\sigma^2)\,T - \sigma b_n}{\sigma\sqrt{T}}\right) + e^{b_n\left(\frac{r+\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(-\frac{(r+\frac{1}{2}\sigma^2)\,T + \sigma b_n}{\sigma\sqrt{T}}\right)\right] \\ &\quad - \left(\frac{U}{S}\right)^{\frac{\mu}{\sigma^2}} \sum_{n=-\infty}^{-1} \left[e^{b_n\left(\frac{r+\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(\frac{(r+\frac{1}{2}\sigma^2)\,T + \sigma b_n}{\sigma\sqrt{T}}\right) + e^{-b_n\left(\frac{r+\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(-\frac{(r+\frac{1}{2}\sigma^2)\,T - \sigma b_n}{\sigma\sqrt{T}}\right)\right].\end{aligned}$$

3. $\int_0^T g_L(t;S)dt$: We have

$$\int_0^T g_L(t;S)dt = \left(\frac{L}{S}\right)^{\frac{\mu}{\sigma^2}} \int_0^T \exp\left(-\frac{1}{2}\frac{\mu^2}{\sigma^2}t\right) \sum_{n=-\infty}^{\infty} f(t;a_n)dt.$$

By (6.2.3) and (6.2.4) and we have

$$\begin{aligned}\int_0^T g_L(t;S)dt &= \left(\frac{L}{S}\right)^{\frac{\mu}{\sigma^2}} \sum_{n=0}^{\infty} \left[e^{-a_n\sqrt{2z}} N\left(\frac{\sqrt{2z}\,T - a_n}{\sqrt{T}}\right) + e^{a_n\sqrt{2z}} N\left(-\frac{\sqrt{2z}\,T + a_n}{\sqrt{T}}\right)\right] \\ &\quad - \left(\frac{L}{S}\right)^{\frac{\mu}{\sigma^2}} \sum_{n=-\infty}^{-1} \left[e^{a_n\sqrt{2z}} N\left(\frac{\sqrt{2z}\,T + a_n}{\sqrt{T}}\right) + e^{-a_n\sqrt{2z}} N\left(-\frac{\sqrt{2z}\,T - a_n}{\sqrt{T}}\right)\right],\end{aligned}$$

where $z = \frac{1}{2}\frac{\mu^2}{\sigma^2} = \frac{(r-\frac{1}{2}\sigma^2)^2}{2\sigma^2}$. Now we have

$$\begin{aligned}&\int_0^T g_L(t;S)dt \\ &= \left(\frac{L}{S}\right)^{\frac{\mu}{\sigma^2}} \sum_{n=0}^{\infty} \left[e^{-a_n\left(\frac{r-\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(\frac{(r-\frac{1}{2}\sigma^2)\,T - \sigma a_n}{\sigma\sqrt{T}}\right) + e^{a_n\left(\frac{r-\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(-\frac{(r-\frac{1}{2}\sigma^2)\,T + \sigma a_n}{\sigma\sqrt{T}}\right)\right] \\ &\quad - \left(\frac{L}{S}\right)^{\frac{\mu}{\sigma^2}} \sum_{n=-\infty}^{-1} \left[e^{a_n\left(\frac{r-\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(\frac{(r-\frac{1}{2}\sigma^2)\,T + \sigma a_n}{\sigma\sqrt{T}}\right) + e^{-a_n\left(\frac{r-\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(-\frac{(r-\frac{1}{2}\sigma^2)\,T - \sigma a_n}{\sigma\sqrt{T}}\right)\right].\end{aligned}$$

4. $\int_0^T g_U(t;S)dt$: Similarly, we have

$$
\begin{aligned}
&\int_0^T g_L(t;S)dt \\
= &\left(\frac{U}{S}\right)^{\frac{\mu}{\sigma^2}} \sum_{n=0}^{\infty} \left[e^{-b_n\left(\frac{r-\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(\frac{(r-\frac{1}{2}\sigma^2)T-\sigma b_n}{\sigma\sqrt{T}}\right) + e^{b_n\left(\frac{r-\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(-\frac{(r-\frac{1}{2}\sigma^2)T+\sigma b_n}{\sigma\sqrt{T}}\right)\right] \\
&-\left(\frac{U}{S}\right)^{\frac{\mu}{\sigma^2}} \sum_{n=-\infty}^{-1} \left[e^{b_n\left(\frac{r-\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(\frac{(r-\frac{1}{2}\sigma^2)T+\sigma b_n}{\sigma\sqrt{T}}\right) + e^{-b_n\left(\frac{r-\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(-\frac{(r-\frac{1}{2}\sigma^2)T-\sigma b_n}{\sigma\sqrt{T}}\right)\right].
\end{aligned}
$$

✎ **Single barrier as a special case of double barrier**
Although the single barrier hitting probabilities can be driven by the reflection principle(Theorem (6.1.7) and (6.1.8)), they are simply special cases of double barrier.

We denote $g_l(t)$ and $g_u(t)$ for the density functions of the hitting time to lower barrier and upper barrier respectively. It is obvious that

$$
\begin{aligned}
g_l(t,S) &= \lim_{U\to\infty} g_L(t,S), && (6.3.1)\\
g_u(t,S) &= \lim_{L\to-\infty} g_U(t,S). && (6.3.2)
\end{aligned}
$$

Note that

$$
\begin{aligned}
f(t;a) &= \frac{|a|}{\sqrt{2\pi t^3}} \exp\left(-\frac{a^2}{2t}\right),\\
a_n = &= \frac{1}{\sigma}\log\frac{U^{2n}S}{L^{2n+1}},\\
b_n = &= \frac{1}{\sigma}\log\frac{U^{2n+1}}{L^{2n}S}.
\end{aligned}
$$

These equations yield that

$$
\begin{cases}
\lim_{U\to\infty} a_n = \infty, \quad \lim_{L\to 0} b_n = \infty, & \text{if } n>0,\\
\lim_{U\to\infty} a_n = -\infty, \quad \lim_{L\to 0} b_n = -\infty, & \text{if } n<0,\\
a_0 = \frac{1}{\sigma}\log\frac{S}{L}, \quad b_0 = \frac{1}{\sigma}\log\frac{U}{S}. & \text{if } n=0.
\end{cases}
$$

Also we obtain

$$
\begin{cases}
\lim_{U\to\infty} f(t,a_n) = 0, & \text{if } n\neq 0,\\
\lim_{L\to 0} f(t;b_n) = 0, & \text{if } n\neq 0.
\end{cases}
$$

Hence previous four basic integrals are simplified as

1. One touch option(down and in bond):

$$
\int_0^T e^{-rt} g_l(t;S) = \left(\frac{L}{S}\right)^{\frac{\mu}{\sigma^2}} \left[e^{-a_0\left(\frac{r+\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(\frac{(r+\frac{1}{2}\sigma^2)T-\sigma a_0}{\sigma\sqrt{T}}\right) + e^{a_0\left(\frac{r+\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(-\frac{(r+\frac{1}{2}\sigma^2)T+\sigma a_0}{\sigma\sqrt{T}}\right)\right]
$$

$$= \left(\frac{L}{S}\right)^{\frac{2r}{\sigma^2}} N\left(\frac{\log\frac{L}{S} + (r+\frac{1}{2}\sigma^2)\,T}{\sigma\sqrt{T}}\right) + \left(\frac{L}{S}\right)^{-1} N\left(\frac{\log\frac{L}{S} - (r+\frac{1}{2}\sigma^2)\,T}{\sigma\sqrt{T}}\right)$$

2. One touch option(up and in bond):

$$\begin{aligned}\int_0^T e^{-rt} g_u(t;S) &= \left(\frac{U}{S}\right)^{\frac{\mu}{\sigma^2}}\left[e^{-b_0\left(\frac{r+\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(\frac{(r+\frac{1}{2}\sigma^2)\,T - \sigma b_0}{\sigma\sqrt{T}}\right) + e^{b_0\left(\frac{r+\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(-\frac{(r+\frac{1}{2}\sigma^2)\,T + \sigma b_0}{\sigma\sqrt{T}}\right)\right]\\ &= \left(\frac{S}{U}\right) N\left(\frac{\log\frac{S}{U} + (r+\frac{1}{2}\sigma^2)\,T}{\sigma\sqrt{T}}\right) + \left(\frac{S}{U}\right)^{-\frac{2r}{\sigma^2}} N\left(\frac{\log\frac{S}{U} - (r+\frac{1}{2}\sigma^2)\,T}{\sigma\sqrt{T}}\right)\end{aligned}$$

An alternative derivation for one touch(up and in) option's formula is given in Section 9.7.

3. European one touch option(down and in, un-discounted):

$$\begin{aligned}\int_0^T g_l(t;S) &= \left(\frac{L}{S}\right)^{\frac{\mu}{\sigma^2}}\left[e^{-a_0\left(\frac{r-\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(\frac{(r-\frac{1}{2}\sigma^2)\,T - \sigma a_0}{\sigma\sqrt{T}}\right) + e^{a_0\left(\frac{r-\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(-\frac{(r-\frac{1}{2}\sigma^2)\,T + \sigma a_0}{\sigma\sqrt{T}}\right)\right]\\ &= \left(\frac{L}{S}\right)^{\frac{2r}{\sigma^2}-1} N\left(\frac{\log\frac{L}{S} + (r-\frac{1}{2}\sigma^2)\,T}{\sigma\sqrt{T}}\right) + N\left(\frac{\log\frac{L}{S} - (r-\frac{1}{2}\sigma^2)\,T}{\sigma\sqrt{T}}\right)\end{aligned}$$

An alternative derivation for one touch(down and in) option's formula is given in Section 8.3.2.

4. European one touch option(up and in, un-discounted):

$$\begin{aligned}\int_0^T g_u(t;S) &= \left(\frac{U}{S}\right)^{\frac{\mu}{\sigma^2}}\left[e^{-b_0\left(\frac{r-\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(\frac{(r-\frac{1}{2}\sigma^2)\,T - \sigma b_0}{\sigma\sqrt{T}}\right) + e^{b_0\left(\frac{r-\frac{1}{2}\sigma^2}{\sigma}\right)} N\left(-\frac{(r-\frac{1}{2}\sigma^2)\,T + \sigma b_0}{\sigma\sqrt{T}}\right)\right]\\ &= N\left(\frac{\log\frac{S}{U} + (r-\frac{1}{2}\sigma^2)\,T}{\sigma\sqrt{T}}\right) + \left(\frac{S}{U}\right)^{1-\frac{2r}{\sigma^2}} N\left(\frac{\log\frac{S}{U} - (r-\frac{1}{2}\sigma^2)\,T}{\sigma\sqrt{T}}\right)\end{aligned}$$

Alternative derivations for one touch(up and in) option's formula is given in Section 8.3.2 and Example 9.7.1. Furthermore this formula can be obtain from Corollary (6.1.3):

$$\begin{aligned}\int_0^T g_u(t;S) &= P(M_X \geq \log U)\\ &= 1 - P(M_X < \log U)\\ &= 1 - F(\log U),\end{aligned}$$

where $F(x)$ is given in (6.1.2). Hence we have

$$\begin{aligned}P\Big(M_S(T) \geq U\Big) &= P\Big(M_X(T) \geq \log U\Big)\\ &= 1 - N\left(\frac{\log U - \log S - \mu t}{\sigma\sqrt{T}}\right) + \exp\left(\frac{2\mu(\log U - \log S)}{\sigma^2}\right) N\left(-\frac{\log U - \log S + \mu T}{\sigma\sqrt{T}}\right)\\ &= N\left(\frac{\log\frac{S}{U} + \mu T}{\sigma\sqrt{T}}\right) + \left(\frac{S}{U}\right)^{-\frac{2\mu}{\sigma^2}} N\left(\frac{\log\frac{S}{U} - \mu T}{\sigma\sqrt{T}}\right). \qquad (6.3.3)\end{aligned}$$

6.4 Distribution of a Brownian Motion with Double Barrier

Fix $T > 0$, and $a < 0 < b$. Let $f_{ab}(x)$ be the defective density function of W_T such that W does not reach either the level a or the level b, where W_t is a standard Brownian motion. Thus, for any Borel set $E \subset (a, b)$,

$$P\Big(W_T \in E, a < m_W(T) < M_W(T) < b\Big) = \int_E f_{ab}(x)dx,$$

where m_W and M_W are minimum and maximum processes of W, respectively. We also define $g_{ab}(x)$ as the defective density function of W_T so that W reaches a or b before time T. In order to derive the density functions $f_{ab}(x)$ and $g_{ab}(x)$, we define the followings recursively:

1. Define

$$\begin{aligned} \tau_1^a &= \inf\{t > 0 : W_t = a\}, \quad \tau_1^b = \inf\{t > \tau_1^a : W_t = b\}, \\ &\vdots \\ \tau_n^a &= \inf\{t > \tau_{n-1}^b : W_t = a\}, \quad \tau_n^b = \inf\{t > \tau_n^a : W_t = b\}, \\ &\vdots \end{aligned}$$

and

$$\begin{aligned} \sigma_1^b &= \inf\{t > 0 : W_t = b\}, \quad \sigma_1^a = \inf\{t > \sigma_1^b : W_t = a\}, \\ &\vdots \\ \sigma_n^b &= \inf\{t > \sigma_{n-1}^b : W_t = b\}, \quad \sigma_n^a = \inf\{t > \sigma_n^b : W_t = a\}, \\ &\vdots \end{aligned}$$

2. Define, for any Borel set $E \subset (a, b)$,

$$\begin{aligned} A_1 &= \{t_1^a < T, W_T \in E\}, \quad A_2 = \{\tau_1^b < T, W_T \in E\}, \\ &\vdots \\ A_{2n-1} &= \{t_n^a < T, W_T \in E\}, \quad A_{2n} = \{\tau_n^b < T, W_T \in E\}, \\ &\vdots \end{aligned}$$

and

$$\begin{aligned} B_1 &= \{t_1^b < T, W_T \in E\}, \quad B_2 = \{\sigma_1^a < T, W_T \in E\}, \\ &\vdots \\ B_{2n-1} &= \{t_n^b < T, W_T \in E\}, \quad B_{2n} = \{\sigma_n^a < T, W_T \in E\}, \\ &\vdots \end{aligned}$$

Note that for any $n > 1$,

$$A_{n-1} \cap B_{n-1} = A_n \cup B_n. \tag{6.4.1}$$

By the reflection principle we have, for $a < x$,

$$\begin{aligned} P\Big(\tau_1^a < T, x < W_T\Big) &= P\Big(\tau_1^a < T, W_T < 2a - x\Big) \\ &= P\Big(W_T < 2a - x\Big). \end{aligned}$$

So, generally we have

$$\begin{aligned} P\Big(\tau_n^a < T, x < W_T\Big) &= P\Big(\tau_n^a < T, W_T < 2a - x\Big) \\ &= P\Big(\tau_{n-1}^b < T, W_T < 2a - x\Big), \quad \text{for } x > a, \end{aligned} \tag{6.4.2}$$

and similarly

$$\begin{aligned} P\Big(\tau_n^b < T, W_T < x\Big) &= P\Big(\tau_n^b < T, 2b - x < W_T\Big) \\ &= P\Big(\tau_n^a < T, 2b - x < W_T\Big), \quad \text{for } x < b. \end{aligned} \tag{6.4.3}$$

We can also obtain

$$\begin{aligned} P\Big(\sigma_n^b < T, W_T < x\Big) &= P\Big(\sigma_n^b < T, 2b - x < W_T\Big) \\ &= P\Big(\sigma_{n-1}^a < T, 2b - x < W_T\Big), \quad \text{for } x < b, \end{aligned} \tag{6.4.4}$$

and

$$\begin{aligned} P\Big(\sigma_n^a < T, x < W_T\Big) &= P\Big(\sigma_n^a < T, W_T < 2a - x\Big) \\ &= P\Big(\sigma_n^b < T, W_T < 2a - x\Big), \quad \text{for } x > a. \end{aligned} \tag{6.4.5}$$

1. Applying (6.4.2) and (6.4.3) recursively, for $x > a$, we can write

$$\begin{aligned} P\Big(\tau_n^a < T, x < W_T\Big) &= P\Big(\tau_{n-1}^b < T, W_T < 2a - x\Big) \\ &= P\Big(\tau_{n-1}^a < T, 2(b-a) + x < W_T\Big) \\ &= P\Big(\tau_{n-2}^b < T, W_T < 2a - 2(b-a) - x\Big) \\ &= P\Big(\tau_{n-2}^a < T, 4(b-a) + x < W_T\Big) \\ &\vdots \\ &= P\Big(\tau_1^a < T, 2(n-1)(b-a) + x < W_T\Big) \\ &= P\Big(W_T < 2b - 2n(b-a) - x\Big) \\ &= \int_{-\infty}^{2b-2n(b-a)-x} \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{y^2}{2T}\right) dy \\ &= \int_x^{\infty} \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{\big(2b - 2n(b-a) - z\big)^2}{2T}\right) dz. \end{aligned}$$

So for any Borel set $E \subset (a,b)$, we have

$$\begin{aligned} P(A_{2n-1}) &= P\Big(\tau_n^a < T, W_T \in E\Big) \\ &= \int_E \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{(z-2b+2n(b-a))^2}{2T}\right) dz. \end{aligned}$$

2. Similarly, applying (6.4.3) and (6.4.2) recursively, for $x < b$, we can write

$$\begin{aligned} P\Big(\tau_n^b < T, W_T < x\Big) &= P\Big(\tau_n^a < T, 2b - x < W_T\Big) \\ &= P\Big(\tau_{n-1}^b < T, W_T < x - 2(b-a)\Big) \\ &= P\Big(\tau_{n-1}^a < T, 2b + 2(b-a) - x < W_T\Big) \\ &= P\Big(\tau_{n-2}^b < T, W_T < x - 4(b-a)\Big) \\ &\ \ \vdots \\ &= P\Big(\tau_1^a < T, 2b + 2(n-1)(b-a) - x < W_T\Big) \\ &= P\Big(W_T < x - 2n(b-a)\Big) \\ &= \int_{-\infty}^{x-2n(b-a)} \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{y^2}{2T}\right) dy \\ &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{\big(2n(b-a)-z\big)^2}{2T}\right) dz. \end{aligned}$$

Hence, for any Borel set $E \subset (a,b)$, we have

$$\begin{aligned} P(A_{2n}) &= P\Big(\tau_n^b < T, W_T \in E\Big) \\ &= \int_E \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{(z-2n(b-a))^2}{2T}\right) dz. \end{aligned}$$

3. Applying (6.4.4) and (6.4.5) recursively, for $x < b$, we can write

$$\begin{aligned} P\Big(\sigma_n^b < T, W_T < x\Big) &= P\Big(\sigma_{n-1}^a < T, 2b - x < W_T\Big) \\ &= P\Big(\sigma_{n-1}^b < T, W_T < x - 2(b-a)\Big) \\ &\ \ \vdots \\ &= P\Big(\sigma_1^b < T, W_T < x - 2(n-1)(b-a)\Big) \\ &= P\Big(2b + 2(n-1)(b-a) - x < W_T\Big) \\ &= \int_{2b+2(n-1)(b-a)-x}^{\infty} \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{y^2}{2T}\right) dy \\ &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{\big(2b+2(n-1)(b-a)-z\big)^2}{2T}\right) dz. \end{aligned}$$

Hence, for any Borel set $E \subset (a,b)$, we have

$$\begin{aligned} P(B_{2n-1}) &= P\Big(\sigma_n^b < T, W_T \in E\Big) \\ &= \int_E \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{\big(z-2b-2(n-1)(b-a)\big)^2}{2T}\right) dz. \end{aligned}$$

4. Finally, applying (6.4.5) and (6.4.4) recursively, for $a < x$, we can write

$$\begin{aligned} P\Big(\sigma_n^a < T, x < W_T\Big) &= P\Big(\sigma_n^b < T, W_T < 2a - x\Big) \\ &= P\Big(\sigma_{n-1}^a < T, 2(b-a) + x < W_T\Big) \\ &\;\;\vdots \\ &= P\Big(\sigma_1^b < T, W_T < 2a - 2(n-1)(b-a) - x\Big) \\ &= P\Big(x + 2n(b-a) < W_T\Big) \\ &= \int_{x+2n(b-a)}^{\infty} \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{y^2}{2T}\right) dy \\ &= \int_x^{\infty} \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{\big(2n(b-a)+z\big)^2}{2T}\right) dz. \end{aligned}$$

Hence, for any Borel set $E \subset (a,b)$, we have

$$\begin{aligned} P(B_{2n}) &= P\Big(\sigma_n^a < T, W_T \in E\Big) \\ &= \int_E \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{\big(z+2n(b-a)\big)^2}{2T}\right) dz. \end{aligned}$$

Now from the definition of $g_{ab}(x)$ and (6.4.1), we can write

$$\begin{aligned} \int_E g_{ab}(x)dx &= P(A_1 \cup B_1) \\ &= P(A_1) + P(B_1) - P(A_1 \cap B_1) \\ &= P(A_1) + P(B_1) - P(A_2 \cup B_2) \\ &= P(A_1) + P(B_1) - P(A_2) - P(B_2) + P(A_1 \cup B_1) \\ &\;\;\vdots \\ &= \sum_{n=1}^{\infty} (-1)^{n-1} \Big\{P(A_n) + P(B_n)\Big\} \\ &= \sum_{n=1}^{\infty} \Big\{P(A_{2n-1}) + P(B_{2n-1}) - P(A_{2n}) - P(B_{2n})\Big\}. \end{aligned}$$

Since

$$f_{ab}(x) = \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{x^2}{2T}\right) - g_{ab}(x),$$

we obtain the following theorem.

Theorem 6.4.1. For any $a < 0 < b$ and $E \subset (a, b)$, we have

$$P\Big(a < m_W(T) < M_W(T) < b, W_T \in E\Big) \quad = \quad \int_E f_{ab}(x)dx,$$

where

$$f_{ab}(x) \quad = \quad \frac{1}{\sqrt{2\pi T}} \sum_{n=-\infty}^{\infty} \left[\exp\left(-\frac{\big(x+2n(b-a)\big)^2}{2T}\right) - \exp\left(-\frac{(x-2b+2n(b-a))^2}{2T}\right)\right]$$

Now let us generalize the previous theorem.

Theorem 6.4.2. Let $X(t) = \mu t + \sigma W_t$. For any $a < 0 < b$, $E \subset (a, b)$, we have

$$P\Big(a < m_X(T) < M_X(T) < b, X_T \in E\Big) \quad = \quad \int_E h_{ab}(x)dx, \tag{6.4.6}$$

where

$$\begin{aligned} h_{a,b} \quad = \quad \frac{1}{\sigma\sqrt{2\pi T}} \sum_{n=-\infty}^{\infty} \Bigg[& \exp\left(\frac{-2\mu n(b-a)}{\sigma^2}\right) \exp\left(-\frac{\big(x+2n(b-a)-\mu T\big)^2}{2\sigma^2 T}\right) \\ & - \exp\left(\frac{-2\mu\big(n(b-a)-b\big)}{\sigma^2}\right) \exp\left(-\frac{\big(x-2b+2n(b-a)-\mu T\big)^2}{2\sigma^2 T}\right)\Bigg]. \end{aligned}$$

Proof. Let $Z(t) := \frac{1}{\sigma} X(t)$. Then

$$Z(t) \quad = \quad \frac{\mu}{\sigma} t + W_t.$$

By the Girsanov theorem Z_t is a Brownian motion under probability measure Q such that

$$dQ \quad = \quad \exp\left(-\frac{\mu}{\sigma} Z_T + \frac{\mu^2}{2\sigma^2} T\right) dP.$$

Now, we have

$$\begin{aligned} & P\Big(a < m_X(T) < M_X(T) < b, X_T \in E\Big) \\ = \quad & P\Big(\frac{a}{\sigma} < m_Z(T) < M_Z(T) < \frac{b}{\sigma}, Z_T \in \frac{E}{\sigma}\Big) \\ = \quad & E_P\left[\mathbf{1}_{\{\frac{a}{\sigma} < m_Z(T) < M_Z(T) < \frac{b}{\sigma}, Z_T \in \frac{E}{\sigma}\}}\right] \\ = \quad & E_P\left[\mathbf{1}_{\{\frac{a}{\sigma} < m_Z(T) < M_Z(T) < \frac{b}{\sigma}, Z_T \in \frac{E}{\sigma}\}} \exp\left(\frac{\mu}{\sigma} Z_T - \frac{\mu^2}{2\sigma^2} T\right)\right] \\ = \quad & \int_{\frac{E}{\sigma}} f_{\frac{a}{\sigma}\frac{b}{\sigma}}(z) \exp\left(\frac{\mu}{\sigma} z - \frac{\mu^2}{2\sigma^2} T\right) dz \\ = \quad & \int_E \frac{1}{\sigma} f_{\frac{a}{\sigma}\frac{b}{\sigma}}\left(\frac{x}{\sigma}\right) \exp\left(\frac{\mu}{\sigma^2} x - \frac{\mu^2}{2\sigma^2} T\right) dx, \end{aligned}$$

where f is defined in theorem(6.4.1). Thus we have

$$h_{a,b} \quad = \quad \frac{1}{\sigma} f_{\frac{a}{\sigma}\frac{b}{\sigma}}\left(\frac{x}{\sigma}\right) \exp\left(\frac{\mu}{\sigma^2} x - \frac{\mu^2}{2\sigma^2} T\right).$$

Since

$$\begin{aligned}
&\exp\left(\frac{\mu}{\sigma^2}x - \frac{\mu^2}{2\sigma^2}T\right)\exp\left(-\frac{\big(x+2n(b-a)\big)^2}{2\sigma^2 T}\right)\\
= \;&\exp\left(\frac{-2\mu n(b-a)}{\sigma^2}\right)\exp\left(-\frac{\big(x+2n(b-a)-\mu T\big)^2}{2\sigma^2 T}\right),
\end{aligned}$$

and

$$\begin{aligned}
&\exp\left(\frac{\mu}{\sigma^2}x - \frac{\mu^2}{2\sigma^2}T\right)\exp\left(-\frac{(x-2b+2n(b-a))^2}{2\sigma^2 T}\right)\\
= \;&\exp\left(\frac{-2\mu\big(n(b-a)-b\big)}{\sigma^2}\right)\exp\left(-\frac{\big(x-2b+2n(b-a)-\mu T\big)^2}{2\sigma^2 T}\right),
\end{aligned}$$

we obtain

$$\begin{aligned}
h_{a,b} = \;&\frac{1}{\sigma\sqrt{2\pi T}}\sum_{n=-\infty}^{\infty}\left[\exp\left(\frac{-2\mu n(b-a)}{\sigma^2}\right)\exp\left(-\frac{\big(x+2n(b-a)-\mu T\big)^2}{2\sigma^2 T}\right)\right.\\
&\left. - \exp\left(\frac{-2\mu\big(n(b-a)-b\big)}{\sigma^2}\right)\exp\left(-\frac{\big(x-2b+2n(b-a)-\mu T\big)^2}{2\sigma^2 T}\right)\right].
\end{aligned}$$

■

Example 6.4.3. From Theorem 6.4.2 we can obtain the conditional probability

$$P\Big(a < m_X(T) < M_X(T) < b \;\Big|\; X_T = x\Big).$$

The conditional probability can be obtain if we divide $h_{a,b}$ defined in (6.4.6) by the probability density function of X_T. Since

$$\begin{aligned}
&\frac{-2\mu n(b-a)}{\sigma^2} - \frac{\big(x+2n(b-a)-\mu T\big)^2}{2\sigma^2 T} + \frac{\big(x-\mu T\big)^2}{2\sigma^2 T}\\
= \;&\frac{-2\mu n(b-a)}{\sigma^2} - \frac{4n(b-a)(x-\mu T)+4n^2(b-a)^2}{2\sigma^2 T}\\
= \;&-\frac{4n(b-a)x+4n^2(b-a)^2}{2\sigma^2 T}\\
= \;&-\frac{2n(b-a)\big(x+n(b-a)\big)}{\sigma^2 T},
\end{aligned}$$

and

$$\begin{aligned}
&\frac{-2\mu\big(n(b-a)-b\big)}{\sigma^2} - \frac{\big(x-2b+2n(b-a)-\mu T\big)^2}{2\sigma^2 T} + \frac{\big(x-\mu T\big)^2}{2\sigma^2 T}\\
= \;&-\frac{2\big(n(b-a)-b\big)\big(x+n(b-a)-b\big)}{\sigma^2 T},
\end{aligned}$$

we have

$$\begin{aligned}
&P\Big(a < m_X(T) < M_X(T) < b \;\Big|\; X_T = x\Big)\\
= \;&\sum_{n=-\infty}^{\infty}\left[\exp\left(-\frac{2n(b-a)\big(x+n(b-a)\big)}{\sigma^2 T}\right) - \exp\left(-\frac{2\big(n(b-a)-b\big)\big(x+n(b-a)-b\big)}{\sigma^2 T}\right)\right].
\end{aligned}$$

Example 6.4.4. Let $\phi(x) = \frac{1}{\sqrt{2\pi T}} \exp\left\{-\frac{(x-\mu T)^2}{2\sigma^2 T}\right\}$, i.e. the probability density function of X_T. We have the conditional probability

$$\begin{aligned}
& P\Big(M_X(T) < b \;\Big|\; m_X(T) > a, X_T = x\Big) \\
= \;& \frac{h_{a,b}}{P\big(m_X(T) > a, X_T = x\big)} \\
= \;& \frac{P\big(a < m_X(T) < M_X(T) < b \;\big|\; X_T = x\big)\; \phi(x)}{P\big(m_X(T) > a \;\big|\; X_T = x\big)\; \phi(x)} \\
= \;& \frac{P\big(a < m_X(T) < M_X(T) < b \;\big|\; X_T = x\big)}{P\big(m_X(T) > a \;\big|\; X_T = x\big)}. \qquad (6.4.7)
\end{aligned}$$

The denominator of (6.4.7) can be obtained from (8.1.1) and the numerator from Example 6.4.3. Note that

$$\begin{aligned}
\lim_{a\to-\infty} P\big(m_X(T) > a \;\big|\; X_T = x\big) &= 1, \\
\lim_{a\to-\infty} \exp\left(-\frac{2n(b-a)\big(x+n(b-a)\big)}{\sigma^2 T}\right) &= \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}, \\
\lim_{a\to-\infty} \exp\left(-\frac{2\big(n(b-a)-b\big)\big(x+n(b-a)-b\big)}{\sigma^2 T}\right) &= \begin{cases} \exp\left(\frac{2b(x-b)}{\sigma^2 T}\right), & n = 0 \\ 0, & n \neq 0 \end{cases}.
\end{aligned}$$

Thus we have

$$\lim_{a\to-\infty} P\Big(M_X(T) < b \;\Big|\; m_X(T) > a, X_T = x\Big) = 1 - \exp\left(\frac{2b(x-b)}{\sigma^2 T}\right).$$

Chapter 7

Lookback Options

Lookback options are contracts which at the delivery time T allow you to take advantage of the realized maximum or minimum of the underlying price process over the entire contract period. Typical examples are

1. lookback call:
$$S_T - \min_{t \le T} S_t$$
2. lookback put:
$$\max_{t \le T} S_t - S_T$$
3. forward lookback call:
$$\left(\max_{t \le T} S_t - K\right)^+$$
4. forward lookback put:
$$\left(K - \min_{t \le T} S_t\right)^+$$

7.1 Lookback Put

The price of the lookback put(not newly issued) at $t = 0$, denoted by $\Pi(0)$, is given by

$$\begin{aligned}
\Pi(0) &= e^{-rT} E\left[\max\left(S_{\max}, \sup_{0 \le t \le T} S(t)\right) - S(T)\right] \\
&= e^{-rT} E\left[\max\left(S_{\max}, \sup_{0 \le t \le T} S(t)\right)\right] - e^{-rT} E\left[S(T)\right] \\
&= e^{-rT} E\left[\max\left(S_{\max}, \sup_{0 \le t \le T} S(t)\right)\right] - S_0 e^{-qT},
\end{aligned}$$

where $S_{\max}(\ge S_0)$ is the maximum stock price achieved to time 0.

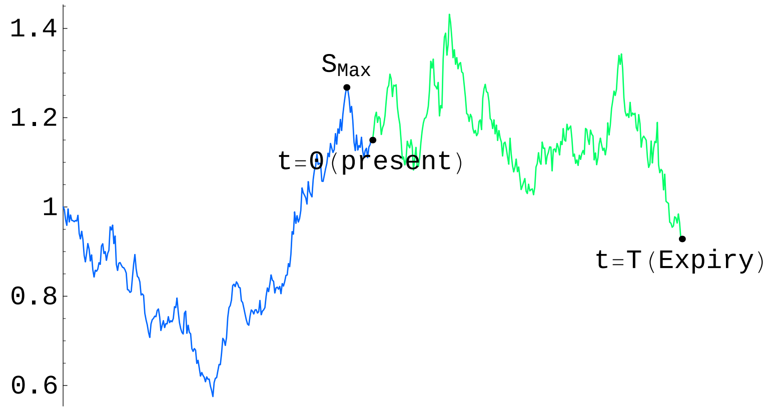


Figure 7.1: Lookback option

From (6.1.1), we see that

$$X(t) = \log S(t), \quad M_S(T) \; = \; e^{M_X(T)}$$

and the distribution for $M_X(t)$, F, is known to us from Corollary (6.1.3). Let f be the density function for $M_X(T)$, then $f = F'$. Now, by differentiating F, we have

$$\begin{aligned} f(x) &= \varphi(x; \log S_0 + \mu T, \sigma^2 T) - \frac{2\mu}{\sigma^2} \exp\left(\frac{2\mu(x - \log S_0)}{\sigma^2}\right) N\left(-\frac{x - \log S_0 + \mu T}{\sigma\sqrt{T}}\right) \\ &\quad + \exp\left(\frac{2\mu(x - \log S_0)}{\sigma^2}\right) \varphi(x; \log S_0 - \mu T, \sigma^2 T), \qquad x \geq \log S_0 \end{aligned}$$

where $\mu = r - q - \frac{1}{2}\sigma^2$ [1].

$$\begin{aligned} \Pi(0) &= e^{-rT}\left[\int_{\log S_0}^{\log S_{\max}} S_{\max} f(x) dx + \int_{\log S_{\max}}^{\infty} e^x f(x) dx\right] - S_0 e^{-qT} \\ &= I + II - S_0 e^{-qT}, \end{aligned}$$

where I and II are the first part and the second part of $\Pi(0)$, respectively.

1.

$$\begin{aligned} I &= e^{-rT} S_{\max}\left[F(\log S_{\max}) - F(\log S_0)\right], \\ F(\log S_{\max}) &= N\left(\frac{\log S_{\max} - \log S_0 - \mu T}{\sigma\sqrt{T}}\right) - \exp\left(\frac{2\mu(\log S_{\max} - \log S_0)}{\sigma^2}\right) N\left(-\frac{\log S_{\max} - \log S_0 + \mu T}{\sigma\sqrt{T}}\right) \\ &= N(b_1) - e^{Y_2} N(-b_3). \end{aligned}$$

[1] Note that

$$\varphi(x; \mu, \sigma^2) \; := \; \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

where

$$\begin{aligned} b_1 &= \frac{\log \frac{S_{\max}}{S_0} - (r - q - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}, \\ Y_2 &= \frac{2(r - q - \frac{1}{2}\sigma^2)\log \frac{S_{\max}}{S_0}}{\sigma^2}, \\ b_3 &= \frac{\log \frac{S_{\max}}{S_0} + (r - q - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}. \end{aligned}$$

Note that

$$\begin{aligned} F(\log S_0) &= N\left(\frac{-\mu T}{\sigma\sqrt{T}}\right) - N\left(\frac{-\mu T}{\sigma\sqrt{T}}\right) \\ &= 0. \end{aligned}$$

Thus, we have

$$I = e^{-rT} S_{\max}\left[N(b_1) - e^{Y_2} N(-b_3)\right].$$

2. Decompose f into f_1, f_2 and f_3 where $f = f_1 + f_2 + f_3$ and

$$\begin{aligned} f_1(x) &= \varphi(x; \log S_0 + \mu T, \sigma^2 T), \\ f_2(x) &= -\frac{2\mu}{\sigma^2}\exp\left(\frac{2\mu(x - \log S_0)}{\sigma^2}\right) N\left(-\frac{x - \log S_0 + \mu T}{\sigma\sqrt{T}}\right), \\ f_3(x) &= \exp\left(\frac{2\mu(x - \log S_0)}{\sigma^2}\right)\varphi(x; \log S_0 - \mu T, \sigma^2 T). \end{aligned}$$

Now, We have

$$\begin{aligned} II &= e^{-rT}\int_{\log S_{\max}}^{\infty} e^x f(x)dx \\ &= e^{-rT}\int_{\log S_{\max}}^{\infty} e^x f_1(x)dx + e^{-rT}\int_{\log S_{\max}}^{\infty} e^x f_2(x)dx + e^{-rT}\int_{\log S_{\max}}^{\infty} e^x f_3(x)dx \\ &= II_1 + II_2 + II_3 \end{aligned}$$

where II_1, II_2 and II_3 are first, second and third part of II.

$$\begin{aligned} II_1 &= e^{-rT}\int_{\log S_{\max}}^{\infty} e^x f_1(x)dx \\ &= e^{-rT}\int_{\log S_{\max}}^{\infty} e^x \varphi(x; \log S_0 + \mu T, \sigma^2 T)dx \\ &\overset{(7.1.2)}{=} e^{-rT} S_0 e^{(r-q)T} N\left(-\frac{\log S_{\max} - \log S_0 - \mu T - \sigma^2 T}{\sigma\sqrt{T}}\right) \\ &= S_0 e^{-qT} N\left(\frac{\log \frac{S_0}{S_{\max}} + (r - q + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) \\ &= S_0 e^{-qT} N(-b_2), \end{aligned}$$

where

$$b_2 = \frac{\log \frac{S_{\max}}{S_0} - (r - q + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} = b_1 - \sigma\sqrt{T}.$$

3. Note that $\frac{1}{\sqrt{2\pi}\sigma}e^{kx}e^{-\frac{(x-m)^2}{2\sigma^2}}$ has an indefinite integral [2].

$$
\begin{aligned}
II_3 &= e^{-rT}\int_{\log S_{\max}}^{\infty} e^x f_3(x)dx \\
&= e^{-rT}\exp\left(-\frac{2\mu\log S_0}{\sigma^2}\right)\int_{\log S_{\max}}^{\infty}\exp\left(\left(\frac{2\mu}{\sigma^2}+1\right)x\right)\varphi(x;\log S_0-\mu T,\sigma^2T)dx \\
&= -e^{-rT}\exp\left(-\frac{2\mu\log S_0}{\sigma^2}\right)\exp\left(\left(\frac{2\mu}{\sigma^2}+1\right)(\log S_0-\mu T)\right) \\
&\quad \exp\left(\frac{1}{2}\sigma^2T\left(\frac{2\mu}{\sigma^2}+1\right)^2\right)\left[N\left(-\frac{x-\log S_0+\mu T-(\frac{2\mu}{\sigma^2}+1)\sigma^2T}{\sigma\sqrt{T}}\right)\right]_{\log S_{\max}}^{\infty} \\
&= S_0e^{-qT}N(-b_2).
\end{aligned}
$$

4. After the partial integration for II_2, we have that

$$
\begin{aligned}
II_2 &= -e^{-rT}\frac{2\mu}{\sigma^2}\exp\left(-\frac{2\mu\log S_0}{\sigma^2}\right)\int_{\log S_{\max}}^{\infty}\exp\left(\left(\frac{2\mu}{\sigma^2}+1\right)x\right)N\left(-\frac{x-\log S_0+\mu T}{\sigma\sqrt{T}}\right)dx \\
&= -e^{-rT}\frac{2\mu}{\sigma^2}\exp\left(-\frac{2\mu\log S_0}{\sigma^2}\right)\frac{\sigma^2}{2\mu+\sigma^2}\exp\left(\left(\frac{2\mu}{\sigma^2}+1\right)x\right)N\left(-\frac{x-\log S_0+\mu T}{\sigma\sqrt{T}}\right)\Bigg|_{\log S_{\max}}^{\infty} \\
&\quad -e^{-rT}\frac{2\mu}{\sigma^2}\exp\left(-\frac{2\mu\log S_0}{\sigma^2}\right)\frac{\sigma^2}{2\mu+\sigma^2}\int_{\log S_{\max}}^{\infty}\exp\left(\left(\frac{2\mu}{\sigma^2}+1\right)x\right)\varphi(x;\log S_0-\mu T,\sigma^2T)dx \\
&= e^{-rT}S_{\max}\frac{2\mu}{2\mu+\sigma^2}e^{Y_2}N(-b_3)-\frac{2\mu}{2\mu+\sigma^2}II_3
\end{aligned}
$$

5. Now we can complete the calculation:

$$
\begin{aligned}
\Pi(0) &= I+II-S_0e^{-qT} \\
&= I+II_1+II_2+II_3-S_0e^{-qT} \\
&= e^{-rT}S_{\max}\left[N(b_1)-e^{Y_2}N(-b_3)\right] \\
&\quad +S_0e^{-qT}N(-b_2)+e^{-rT}S_{\max}\frac{2\mu}{2\mu+\sigma^2}e^{Y_2}N(-b_3)-\frac{2\mu}{2\mu+\sigma^2}II_3+II_3-S_0e^{-qT} \\
&= e^{-rT}S_{\max}\left[N(b_1)-\frac{\sigma^2}{2\mu+\sigma^2}e^{Y_2}N(-b_3)\right]+\frac{\sigma^2}{2\mu+\sigma^2}II_3-S_0e^{-qT}(1-N(-b_2)) \\
&= e^{-rT}S_{\max}\left[N(b_1)-\frac{\sigma^2}{2(r-q)}e^{Y_2}N(-b_3)\right]+S_0e^{-qT}\frac{\sigma^2}{2(r-q)}N(-b_2)-S_0e^{-qT}(1-N(-b_2)) \\
&= e^{-rT}S_{\max}\left[N(b_1)-\frac{\sigma^2}{2(r-q)}e^{Y_2}N(-b_3)\right]+S_0e^{-qT}\frac{\sigma^2}{2(r-q)}N(-b_2)-S_0e^{-qT}N(b_2).
\end{aligned}
$$

[2]

$$
\begin{aligned}
\int\frac{1}{\sqrt{2\pi}\sigma}e^{kx}e^{-\frac{(x-m)^2}{2\sigma^2}}dx &= e^{km+\frac{1}{2}k^2\sigma^2}N\left(\frac{x-m-k\sigma^2}{\sigma}\right)+C, && (7.1.1) \\
&= -e^{km+\frac{1}{2}k^2\sigma^2}N\left(-\frac{x-m-k\sigma^2}{\sigma}\right)+C. && (7.1.2)
\end{aligned}
$$

7.2 Lookback Call

The price of the lookback call(not newly issued) at $t = 0$, denoted by $\Pi(0)$, is given by

$$\begin{aligned}\Pi(0) &= e^{-rT} E\left[S(T) - \min\left(S_{\min},\ \inf_{0\le t\le T} S(t)\right)\right]\\ &= e^{-rT} E\left[S(T)\right] e^{-rT} E\left[\min\left(S_{\min},\ \inf_{0\le t\le T} S(t)\right)\right]\\ &= S_0 e^{-qT} - e^{-rT} E\left[\min\left(S_{\min},\ \inf_{0\le t\le T} S(t)\right)\right],\end{aligned}$$

where $S_{\min}(\le S_0)$ is the minimum stock price achieved to time 0.

From (6.1.1), we see that

$$X(t) = \log S(t), \quad m_S(T) = e^{m_X(T)}$$

and the distribution for $m_X(t)$, F, is known to us from Corollary (6.1.3). Let f be the density function for $m_X(T)$, then $f = F'$. Now, by differentiating F, we have

$$\begin{aligned} f(x) &= \varphi(x; \log S_0 + \mu T, \sigma^2 T) + \frac{2\mu}{\sigma^2}\exp\left(\frac{2\mu(x - \log S_0)}{\sigma^2}\right) N\left(\frac{x - \log S_0 + \mu T}{\sigma\sqrt{T}}\right)\\ &\quad + \exp\left(\frac{2\mu(x-\log S_0)}{\sigma^2}\right)\varphi(x; \log S_0 - \mu T, \sigma^2 T), \qquad x \le \log S_0\end{aligned}$$

where $\mu = r - q - \frac{1}{2}\sigma^2$ [3]. Hence we have

$$\begin{aligned}\Pi(0) &= S_0 e^{-qT} - e^{-rT}\left[\int_{\log S_{\min}}^{\log S_0} S_{\min} f(x)dx + \int_{-\infty}^{\log S_{\min}} e^x f(x)dx\right]\\ &= S_0 e^{-qT} - I - II,\end{aligned}$$

where I and II are the first part and the second part of $\Pi(0)$, respectively.

1.

$$\begin{aligned} I &= e^{-rT} S_{\min}\left[F(\log S_0) - F(\log S_{\min})\right],\\ F(\log S_{\min}) &= N\left(\frac{\log S_{\min} - \log S_0 - \mu T}{\sigma\sqrt{T}}\right) + \exp\left(\frac{2\mu(\log S_{\min} - \log S_0)}{\sigma^2}\right) N\left(\frac{\log S_{\min} - \log S_0 + \mu T}{\sigma\sqrt{T}}\right)\\ &= N(-a_2) + e^{Y_1} N(-a_3).\end{aligned}$$

where

$$\begin{aligned} a_1 &= \frac{\log\frac{S_0}{S_{\min}} + (r - q + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}},\\ a_2 &= \frac{\log\frac{S_{\max}}{S_0} + (r - q - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} = a_1 - \sigma\sqrt{T}\end{aligned}$$

[3] Note that

$$\varphi(x; \mu, \sigma^2) := \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

$$\begin{aligned} a_3 &= \frac{\log \frac{S_0}{S_{\min}} - (r - q - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}, \\ Y_1 &= -\frac{2(r - q - \frac{1}{2}\sigma^2)\log \frac{S_0}{S_{\min}}}{\sigma^2}. \end{aligned}$$

Note that

$$\begin{aligned} F(\log S_0) &= N\left(\frac{-\mu T}{\sigma\sqrt{T}}\right) + N\left(\frac{\mu T}{\sigma\sqrt{T}}\right) \\ &= 1. \end{aligned}$$

Thus, we have

$$I = e^{-rT} S_{\min}\left[N(a_2) - e^{Y_1}N(-a_3)\right].$$

2. Decompose f into f_1, f_2 and f_3 where $f = f_1 + f_2 + f_3$ and

$$\begin{aligned} f_1(x) &= \varphi(x; \log S_0 + \mu T, \sigma^2 T), \\ f_2(x) &= \frac{2\mu}{\sigma^2}\exp\left(\frac{2\mu(x - \log S_0)}{\sigma^2}\right) N\left(\frac{x - \log S_0 + \mu T}{\sigma\sqrt{T}}\right), \\ f_3(x) &= \exp\left(\frac{2\mu(x - \log S_0)}{\sigma^2}\right)\varphi(x; \log S_0 - \mu T, \sigma^2 T). \end{aligned}$$

Now, We have

$$\begin{aligned} II &= e^{-rT}\int_{-\infty}^{\log S_{\min}} e^x f(x)dx \\ &= e^{-rT}\int_{-\infty}^{\log S_{\min}} e^x f_1(x)dx + e^{-rT}\int_{-\infty}^{\log S_{\min}} e^x f_2(x)dx + e^{-rT}\int_{-\infty}^{\log S_{\min}} e^x f_3(x)dx \\ &= II_1 + II_2 + II_3 \end{aligned}$$

where II_1, II_2 and II_3 are first, second and third part of II.

$$\begin{aligned} II_1 &= e^{-rT}\int_{-\infty}^{\log S_{\min}} e^x f_1(x)dx \\ &= e^{-rT}\int_{-\infty}^{\log S_{\min}} e^x \varphi(x; \log S_0 + \mu T, \sigma^2 T)dx \\ &\overset{(7.1.2)}{=} e^{-rT} S_0 e^{(r-q)T}\left[1 - N\left(-\frac{\log S_{\min} - \log S_0 - \mu T - \sigma^2 T}{\sigma\sqrt{T}}\right)\right] \\ &= S_0 e^{-qT} N\left(-\frac{\log \frac{S_0}{S_{\min}} + (r - q + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) \\ &= S_0 e^{-qT} N(-a_1), \end{aligned}$$

3. Since $\frac{1}{\sqrt{2\pi}\sigma}e^{kx}e^{-\frac{(x-m)^2}{2\sigma^2}}$ has an indefinite integral, we have

$$II_3 = e^{-rT}\int_{-\infty}^{\log S_{\min}} e^x f_3(x)dx$$

$$
\begin{aligned}
&= e^{-rT}\exp\left(-\frac{2\mu\log S_0}{\sigma^2}\right)\int_{-\infty}^{\log S_{\min}}\exp\left(\left(\frac{2\mu}{\sigma^2}+1\right)x\right)\varphi(x;\log S_0-\mu T,\sigma^2 T)dx \\
&= -e^{-rT}\exp\left(-\frac{2\mu\log S_0}{\sigma^2}\right)\exp\left(\left(\frac{2\mu}{\sigma^2}+1\right)(\log S_0-\mu T)\right) \\
&\quad \exp\left(\frac{1}{2}\sigma^2 T\left(\frac{2\mu}{\sigma^2}+1\right)^2\right)\left[N\left(-\frac{x-\log S_0+\mu T-(\frac{2\mu}{\sigma^2}+1)\sigma^2 T}{\sigma\sqrt{T}}\right)\right]_{-\infty}^{\log S_{\min}} \\
&= S_0 e^{-qT}N(-a_1).
\end{aligned}
$$

4. After the partial integration for II_2, we have that

$$
\begin{aligned}
II_2 &= e^{-rT}\frac{2\mu}{\sigma^2}\exp\left(-\frac{2\mu\log S_0}{\sigma^2}\right)\int_{-\infty}^{\log S_{\min}}\exp\left(\left(\frac{2\mu}{\sigma^2}+1\right)x\right)N\left(\frac{x-\log S_0+\mu T}{\sigma\sqrt{T}}\right)dx \\
&= e^{-rT}\frac{2\mu}{\sigma^2}\exp\left(-\frac{2\mu\log S_0}{\sigma^2}\right)\frac{\sigma^2}{2\mu+\sigma^2}\exp\left(\left(\frac{2\mu}{\sigma^2}+1\right)x\right)N\left(\frac{x-\log S_0+\mu T}{\sigma\sqrt{T}}\right)\Bigg|_{-\infty}^{\log S_{\min}} \\
&\quad -e^{-rT}\frac{2\mu}{\sigma^2}\exp\left(-\frac{2\mu\log S_0}{\sigma^2}\right)\frac{\sigma^2}{2\mu+\sigma^2}\int_{\log S_{\max}}^{\infty}\exp\left(\left(\frac{2\mu}{\sigma^2}+1\right)x\right)\varphi(x;\log S_0-\mu T,\sigma^2 T)dx \\
&= e^{-rT}S_{\min}\frac{2\mu}{2\mu+\sigma^2}e^{Y_1}N(-a_3)-\frac{2\mu}{2\mu+\sigma^2}II_3
\end{aligned}
$$

5. Now we can complete the calculation:

$$
\begin{aligned}
\Pi(0) &= S_0e^{-qT}-I-II \\
&= S_0e^{-qT}-I-II_1-II_2-II_3 \\
&= S_0e^{-qT}-e^{-rT}S_{\min}\left[N(a_2)-e^{Y_1}N(-a_3)\right] \\
&\quad -S_0e^{-qT}N(-a_1)-e^{-rT}S_{\min}\frac{2\mu}{2\mu+\sigma^2}e^{Y_1}N(-a_3)+\frac{2\mu}{2\mu+\sigma^2}II_3-II_3 \\
&= S_0e^{-qT}N(a_1)-\frac{\sigma^2}{2\mu+\sigma^2}II_3-e^{-rT}S_{\min}\left[N(a_2)-e^{Y_1}N(-a_3)+\frac{2\mu}{2\mu+\sigma^2}e^{Y_1}N(-a_3)\right] \\
&= S_0e^{-qT}N(a_1)-S_0e^{-qT}\frac{\sigma^2}{2(r-q)}N(-a_1)-e^{-rT}S_{\min}\left[N(a_2)-\frac{\sigma^2}{2(r-q)}e^{Y_1}N(-a_3)\right].
\end{aligned}
$$

7.3 Forward Lookback Call

Immediately Issued
The price of forward lookback call option, denoted by $\Pi(0)$ is

$$\Pi(0) = e^{-rT} E\left[\left(\sup_{0\le t\le T} S(t) - K\right)^+\right].$$

To evaluate, we need to identify the distribution of $\sup_{0\le t\le T} S(t)$. Define

$$G(U) = \begin{cases} \int_0^T g_u(t, S_0)dt = P\Big(M_S(T) \ge U\Big) = P\Big(M_X(T) \ge \log U\Big), & \text{if } U \ge S_0 \\ 1, & \text{if } U < S_0 \end{cases}$$

where g_u is defined in (6.3.2).

By Theorem (1.10.2) we have

$$\begin{aligned} \Pi(0) &= e^{-rT}\int_K^\infty (x-K)f(x)dx \\ &= e^{-rT}\int_K^\infty G(x)dx, \end{aligned}$$

where $f(x)$ is the probability density function of $\sup_{0\le t\le T} S(t)$.

1. $K \ge S_0$: Now, to simplify notation, we denote

$$\begin{aligned} d_1(u,v) &= \frac{\log\frac{u}{v} + \mu T}{\sigma\sqrt{T}} + \sigma\sqrt{T} = \frac{\log\frac{u}{v} + (r - q + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}, \\ d_2(u,v) &= \frac{\log\frac{u}{v} + \mu T}{\sigma\sqrt{T}} = \frac{\log\frac{u}{v} + (r - q - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}, \\ d_3(u,v) &= d_2(u,v) - \frac{2\mu}{\sigma}\sqrt{T} = \frac{\log\frac{u}{v} - \mu T}{\sigma\sqrt{T}} = \frac{\log\frac{u}{v} - (r - q - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}. \end{aligned}$$

From (6.3.3) we have

$$\begin{aligned} \Pi(0) &= e^{-rT}\int_K^\infty N\left(\frac{\log\frac{S_0}{x} + \mu T}{\sigma\sqrt{T}}\right)dx + \int_K^\infty \left(\frac{S_0}{x}\right)^{-\frac{2\mu}{\sigma^2}} N\left(\frac{\log\frac{S_0}{x} - \mu T}{\sigma\sqrt{T}}\right)dx \\ &= e^{-rT}\int_K^\infty N\Big(d_2(S_0,x)\Big)dx + \int_K^\infty \left(\frac{S_0}{x}\right)^{-\frac{2\mu}{\sigma^2}} N\left(d_2(S_0,x) - \frac{2\mu}{\sigma}\sqrt{T}\right)dx \qquad (7.3.1) \\ &= e^{-rT}(I + II), \end{aligned}$$

where I and II are the first and second part of (7.3.1) respectively.
The integrations yield that

$$\begin{aligned} e^{-rT} I &= e^{-rT}\int_K^\infty N\Big(d_2(S_0,x)\Big)dx \\ &= S_0 e^{-qT} N\Big(d_1(S_0,K)\Big) - K e^{-rT} N\Big(d_2(S_0,K)\Big), \quad \text{by Theorem (1.10.2)}, \end{aligned}$$

and

$$
\begin{aligned}
& II \\
= \;& S_0^{-\frac{2\mu}{\sigma^2}} \frac{1}{\sqrt{2\pi}} \int_K^\infty x^{\frac{2\mu}{\sigma^2}} \int_{-\infty}^{d_2(S_0,x)-\frac{2\mu}{\sigma}\sqrt{T}} e^{-\frac{y^2}{2}} \, dy \, dx \\
= \;& S_0^{-\frac{2\mu}{\sigma^2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_2(S_0,K)-\frac{2\mu}{\sigma}\sqrt{T}} \int_K^{S_0 \exp\left(-\sigma\sqrt{T}y-\mu T\right)} x^{\frac{2\mu}{\sigma^2}} e^{-\frac{y^2}{2}} \, dx \, dy \\
= \;& \frac{\sigma^2}{2\mu+\sigma^2} S_0^{-\frac{2\mu}{\sigma^2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_2(S_0,K)-\frac{2\mu}{\sigma}\sqrt{T}} \left\{ S_0^{\frac{2\mu+\sigma^2}{\sigma^2}} \exp\left(-\frac{2\mu+\sigma^2}{\sigma}\sqrt{T}y - \frac{2\mu+\sigma^2}{\sigma^2}\mu T \right) - K^{\frac{2\mu+\sigma^2}{\sigma^2}} \right\} e^{-\frac{y^2}{2}} dy \\
= \;& \frac{\sigma^2}{2\mu+\sigma^2} S_0 e^{(\mu+\frac{\sigma^2}{2})T} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_2(S_0,K)-\frac{2\mu}{\sigma}\sqrt{T}} \exp\left(-\frac{(y+\frac{2\mu+\sigma^2}{\sigma}\sqrt{T})^2}{2} \right) dy \\
& - \frac{\sigma^2}{2\mu+\sigma^2} S_0^{-\frac{2\mu}{\sigma^2}} K^{\frac{2\mu+\sigma^2}{\sigma^2}} N\left(d_2(S_0,K) - \frac{2\mu}{\sigma}\sqrt{T} \right) \\
= \;& \frac{\sigma^2}{2\mu+\sigma^2} S_0 e^{(\mu+\frac{\sigma^2}{2})T} N\left(d_2(S_0,K) + \sigma\sqrt{T} \right) - \frac{\sigma^2}{2\mu+\sigma^2} S_0^{-\frac{2\mu}{\sigma^2}} K^{\frac{2\mu+\sigma^2}{\sigma^2}} N\left(d_2(S_0,K) - \frac{2\mu}{\sigma}\sqrt{T} \right) \\
= \;& \frac{\sigma^2}{2\mu+\sigma^2} S_0 e^{(\mu+\frac{\sigma^2}{2})T} N\Big(d_1(S_0,K) \Big) - \frac{\sigma^2}{2\mu+\sigma^2} S_0^{-\frac{2\mu}{\sigma^2}} K^{\frac{2\mu+\sigma^2}{\sigma^2}} N\Big(d_3(S_0,K) \Big).
\end{aligned}
$$

Hence we obtain

$$
\begin{aligned}
\Pi(0) \;=\;& \left(1 + \frac{\sigma^2}{2\mu+\sigma^2} \right) S_0 e^{-qT} N\Big(d_1(S_0,K) \Big) - K e^{-rT} N\Big(d_2(S_0,K) \Big) \\
& - \frac{\sigma^2}{2\mu+\sigma^2} e^{-rT} S_0^{-\frac{2\mu}{\sigma^2}} K^{\frac{2\mu+\sigma^2}{\sigma^2}} N\Big(d_3(S_0,K) \Big) \quad \Big(:= \Psi(K) \Big). \qquad (7.3.2)
\end{aligned}
$$

2. $K < S_0$: The price of forward lookback call is

$$
\begin{aligned}
\Pi(0) \;=\;& e^{-rT} \int_K^\infty G(x) dx \\
=\;& e^{-rT}(S_0 - K) + e^{-rT} \int_{S_0}^\infty G(x) dx. \qquad (7.3.3)
\end{aligned}
$$

The second term on the right-hand side of (7.3.3) is given by (7.3.2) with $K = S_0$ [4]. Hence we have

$$
\Pi(0) \;=\; e^{-rT}(S_0 - K) + \Psi(S_0).
$$

[4] Alternatively,

$$
\begin{aligned}
\Pi(0) \;=\;& e^{-rT} \int_{S_0}^\infty (x-K) f(x) dx \\
=\;& e^{-rT} \int_{S_0}^\infty (x - S_0 + S_0 - K) f(x) dx \\
=\;& e^{-rT} \int_{S_0}^\infty (S_0 - K) f(x) dx + e^{-rT} \int_{S_0}^\infty (x - S_0) f(x) dx \\
=\;& e^{-rT}(S_0 - K) + e^{-rT} \int_{S_0}^\infty (x - S_0) f(x) dx.
\end{aligned}
$$

Therefore we have

$$\Pi(0) = \begin{cases} \Psi(K), & K \geq S_0, \\ e^{-rT}(S_0 - K) + \Psi(S_0), & K < S_0. \end{cases} \tag{7.3.4}$$

Already Issued
At time 0, the price of a European forward lookback call that is issued already, denoted by $\widetilde{\Pi}(0)$, is given by

$$\widetilde{\Pi}(0) = e^{-rT} E\left[\left(\max\left(S_{\max}, \sup_{0 \leq t \leq T} S(t)\right) - K\right)^+\right],$$

where $S_{\max}$ is the maximum stock price achieved to time 0.

Let $f(x)$ be the probability density function of $\sup_{0 \leq t \leq T} S(t)$.

$$\widetilde{\Pi}(0) = e^{-rT} \int_{S_0}^{S_{\max}} (S_{\max} - K)^+ f(x)dx + e^{-rT} \int_{S_{\max}}^{\infty} (x - K)^+ f(x)dx$$

1. $S_{\max} \geq K$: We have

$$\begin{aligned} \widetilde{\Pi}(0) &= e^{-rT} \int_{S_0}^{S_{\max}} (S_{\max} - K) f(x)dx + e^{-rT} \int_{S_{\max}}^{\infty} (x - K) f(x)dx \\ &= e^{-rT} \int_{S_0}^{S_{\max}} (S_{\max} - K) f(x)dx + e^{-rT} \int_{S_{\max}}^{\infty} (x - S_{\max} + S_{\max} - K) f(x)dx \\ &= e^{-rT} \int_{S_0}^{\infty} (S_{\max} - K) f(x)dx + e^{-rT} \int_{S_{\max}}^{\infty} (x - S_{\max}) f(x)dx \\ &= e^{-rT}(S_{\max} - K) G(S_0) + \Psi\Big(S_{\max}\Big) \\ &= e^{-rT}(S_{\max} - K) + \Psi\Big(S_{\max}\Big), \quad \text{since } G(S_0) = 1. \end{aligned}$$

2. $S_{\max} < K$: We have

$$\widetilde{\Pi}(0) = e^{-rT} \int_{K}^{\infty} (x - K) f(x)dx = \Pi(0),$$

where $\Pi(0)$ is given in (7.3.4). Since $S_0 \leq S_{\max} < K$, $\Pi(0) = \Psi(K)$.

Hence these result and (7.3.4) can be summarized as

$$\Pi(0) = \begin{cases} \Psi(K), & K \geq S_{\max}, \\ e^{-rT}(S_{\max} - K) + \Psi(S_{\max}), & K \leq S_{\max}. \end{cases} \tag{7.3.5}$$

If the forward lookback call option has just been issued, $S_{\max} = S_0$.

7.4 Forward Lookback Put

Immediately Issued
The price of forward lookback put option, denoted by $\Pi(0)$ is

$$\Pi(0) = e^{-rT} E\Big[\Big(K - \inf_{0\le t\le T} S(t)\Big)^+\Big].$$

To evaluate, we need to identify the distribution of $\inf_{0\le t\le T} S(t)$. Define

$$G(L) = \begin{cases} \int_0^T g_l(t,S_0)dt = P\Big(m_S(T)\le L\Big) = P\Big(m_X(T)\le \log L\Big), & \text{if } S_0 > L \\ 1, & \text{if } S_0 \le L \end{cases}$$

where g_l is defined in (6.3.1).

By Theorem (1.10.3) we have

$$\begin{aligned}\Pi(0) &= e^{-rT}\int_0^K (K-x)f(x)dx \\ &= e^{-rT}\int_0^K G(x)dx,\end{aligned}$$

where $f(x)$ is the probability density function of $\inf_{0\le t\le T} S(t)$.

1. $K \le S_0$: From Corollary (6.1.3) we have

$$\begin{aligned}\Pi(0) &= e^{-rT}\int_0^K N\left(-\frac{\log\frac{S_0}{x}+\mu T}{\sigma\sqrt{T}}\right)dx + \int_0^K \left(\frac{S_0}{x}\right)^{-\frac{2\mu}{\sigma^2}} N\left(-\frac{\log\frac{S_0}{x}-\mu T}{\sigma\sqrt{T}}\right)dx \\ &= e^{-rT}\int_0^K N\Big(-d_2(S_0,x)\Big)dx + \int_0^K\left(\frac{S_0}{x}\right)^{-\frac{2\mu}{\sigma^2}} N\Big(-d_3(S_0,x)\Big)dx \qquad (7.4.1) \\ &= e^{-rT}(I+II),\end{aligned}$$

where I and II are the first and second part of (7.4.1) respectively.
The integrations yield that

$$\begin{aligned} e^{-rT}I &= e^{-rT}\int_0^K N\Big(-d_2(S_0,x)\Big)dx \\ &= Ke^{-rT}N\Big(-d_2(S_0,K)\Big) - S_0e^{-qT}N\Big(-d_1(S_0,K)\Big), \quad \text{by Theorem (1.10.3)},\end{aligned}$$

and

$$\begin{aligned} & II \\ =\ & S_0^{-\frac{2\mu}{\sigma^2}}\frac{1}{\sqrt{2\pi}}\int_0^K x^{\frac{2\mu}{\sigma^2}}\int_{-\infty}^{-d_3(S_0,x)} e^{-\frac{y^2}{2}}\,dy\,dx \\ =\ & S_0^{-\frac{2\mu}{\sigma^2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-d_3(S_0,K)}\int_{S_0\exp\left(\sigma\sqrt{T}y-\mu T\right)}^K x^{\frac{2\mu}{\sigma^2}} e^{-\frac{y^2}{2}}\,dx\,dy \\ =\ & \frac{\sigma^2}{2\mu+\sigma^2}S_0^{-\frac{2\mu}{\sigma^2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-d_3(S_0,K)}\left\{K^{\frac{2\mu+\sigma^2}{\sigma^2}} - S_0^{\frac{2\mu+\sigma^2}{\sigma^2}}\exp\left(\frac{2\mu+\sigma^2}{\sigma}\sqrt{T}y - \frac{2\mu+\sigma^2}{\sigma^2}\mu T\right)\right\}e^{-\frac{y^2}{2}}\,dy\end{aligned}$$

$$
\begin{aligned}
&= \frac{\sigma^2}{2\mu+\sigma^2} S_0^{-\frac{2\mu}{\sigma^2}} K^{\frac{2\mu+\sigma^2}{\sigma^2}} N\Big(-d_3(S_0,K)\Big) \\
&\quad -\frac{\sigma^2}{2\mu+\sigma^2} S_0 e^{(\mu+\frac{\sigma^2}{2})T} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-d_3(S_0,K)} \exp\left(-\frac{(y-\frac{2\mu+\sigma^2}{\sigma}\sqrt{T})^2}{2}\right) dy \\
&= \frac{\sigma^2}{2\mu+\sigma^2} S_0^{-\frac{2\mu}{\sigma^2}} K^{\frac{2\mu+\sigma^2}{\sigma^2}} N\Big(-d_3(S_0,K)\Big) - \frac{\sigma^2}{2\mu+\sigma^2} S_0 e^{(\mu+\frac{\sigma^2}{2})T} N\left(-d_2(S_0,K)-\sigma\sqrt{T}\right) \\
&= \frac{\sigma^2}{2\mu+\sigma^2} S_0^{-\frac{2\mu}{\sigma^2}} K^{\frac{2\mu+\sigma^2}{\sigma^2}} N\Big(-d_3(S_0,K)\Big) - \frac{\sigma^2}{2\mu+\sigma^2} S_0 e^{(\mu+\frac{\sigma^2}{2})T} N\Big(-d_1(S_0,K)\Big).
\end{aligned}
$$

Hence we obtain

$$
\begin{aligned}
\Pi(0) &= Ke^{-rT} N\Big(-d_2(S_0,K)\Big) - \left(1+\frac{\sigma^2}{2\mu+\sigma^2}\right) S_0 e^{-qT} N\Big(-d_1(S_0,K)\Big) \\
&\quad + \frac{\sigma^2}{2\mu+\sigma^2} e^{-rT} S_0^{-\frac{2\mu}{\sigma^2}} K^{\frac{2\mu+\sigma^2}{\sigma^2}} N\Big(-d_3(S_0,K)\Big) \quad \Big(:= \Phi(K)\Big). \qquad (7.4.2)
\end{aligned}
$$

2. $K > S_0$: In this case the price of forward lookback put is

$$
\begin{aligned}
\Pi(0) &= e^{-rT} \int_0^K G(x)dx \\
&= e^{-rT}(K-S_0) + e^{-rT} \int_0^{S_0} G(x)dx. \qquad (7.4.3)
\end{aligned}
$$

The second term on the right-hand side of (7.4.3) is given by (7.4.2) with $K = S_0$ [5]. Hence we have

$$
\Pi(0) = e^{-rT}(K-S_0) + \Phi(S_0).
$$

Therefore we have

$$
\Pi(0) = \begin{cases} \Phi(K), & K \le S_0 \\ e^{-rT}(K-S_0) + \Phi(S_0), & K > S_0. \end{cases} \qquad (7.4.4)
$$

Already Issued

At time 0, the price of a European forward lookback put that is issued already, denoted by $\widetilde{\Pi}(0)$, is given by

$$
\widetilde{\Pi}(0) = e^{-rT} E\left[\left(K - \min\left(S_{\min}, \inf_{0\le t\le T} S(t)\right)\right)^+\right],
$$

[5] Alternatively,

$$
\begin{aligned}
\Pi(0) &= e^{-rT} \int_0^{S_0} (K-x) f(x)dx \\
&= e^{-rT} \int_0^{S_0} (K-S_0+S_0-x) f(x)dx \\
&= e^{-rT} \int_0^{S_0} (K-S_0) f(x)dx + e^{-rT} \int_0^{S_0} (S_0-x) f(x)dx \\
&= e^{-rT}(K-S_0) + e^{-rT} \int_0^{S_0} (S_0-x) f(x)dx.
\end{aligned}
$$

where $S_{\min}$ is the minimum stock price achieved to time 0.

Let $f(x)$ be the probability density function of $\inf_{0\le t\le T} S(t)$.

$$\widetilde{\Pi}(0) \;=\; e^{-rT}\int_{S_{\min}}^{S_0}(K-S_{\min})^+ f(x)dx + e^{-rT}\int_0^{S_{\min}}(K-x)^+ f(x)dx$$

1. $S_{\min}\le K$: We have

$$\begin{aligned}
\widetilde{\Pi}(0) &= e^{-rT}\int_{S_{\min}}^{S_0}(K-S_{\min})f(x)dx + e^{-rT}\int_0^{S_{\min}}(K-x)f(x)dx \\
&= e^{-rT}\int_{S_{\min}}^{S_0}(K-S_{\min})f(x)dx + e^{-rT}\int_0^{S_{\min}}(K-S_{\min}+S_{\min}-x)f(x)dx \\
&= e^{-rT}\int_0^{S_0}(K-S_{\min})f(x)dx + e^{-rT}\int_0^{S_{\min}}(S_{\min}-x)f(x)dx \\
&= e^{-rT}(K-S_{\min})G(S_0) + \Phi\Big(S_{\min}\Big) \\
&= e^{-rT}(K-S_{\min}) + \Phi\Big(S_{\min}\Big).
\end{aligned}$$

2. $S_{\min}>K$: We have

$$\widetilde{\Pi}(0) \;=\; e^{-rT}\int_0^K (K-x)f(x)dx \;=\; \Pi(0),$$

   where $\Pi(0)$ is given in (7.4.4). Since $S_0\ge S_{\min}>K$, $\Pi(0)=\Phi(K)$.

Hence these result and (7.4.4) can be summarized as

$$\Pi(0) \;=\; \begin{cases}\Phi(K), & K\le S_{\min}, \\ e^{-rT}(K-S_{\min})+\Phi(S_{\min}), & K\ge S_{\min}.\end{cases} \tag{7.4.5}$$

If the forward lookback put option has just been issued, $S_{\min}=S_0$.

Chapter 8

Barrier Options

This chapter discusses barrier options. First let us introduce notations.

✎ **Notation:**

1. S_t: stock price. $S = S_0$
2. T: the time to maturity.
3. $Z_t = \ln(S_t/S_0)$: the continuously compounded return process for S_t.
4. H: the barrier.
5. K: strike price. $\overline{K} = \max(K, H), k = \log(K/S), \overline{k} = \log(\overline{K}/S), b = \log(H/S)$.
6. $\mu := r - q - \frac{\sigma^2}{2}$, the risk-neutral drift for Z_t.

$$D(X, \lambda, \sigma, T) = \frac{\log X + \lambda T}{\sigma\sqrt{T}}.$$

$$\begin{aligned}
d_1 &= D\left(\frac{S}{K}, \sigma^2 + \mu, \sigma, T\right) = \frac{\log(S/K) + (r - q + \sigma^2/2)T}{\sigma\sqrt{T}}, \quad d_2 = d_1 - \sigma\sqrt{T}. \\
\overline{d}_1 &= D\left(\frac{S}{\overline{K}}, \sigma^2 + \mu, \sigma, T\right) = \frac{\log(S/\overline{K}) + (r - q + \sigma^2/2)T}{\sigma\sqrt{T}}, \quad \overline{d}_2 = \overline{d}_1 - \sigma\sqrt{T}. \\
\underline{d}_1 &= D\left(\frac{S}{\underline{K}}, \sigma^2 + \mu, \sigma, T\right) = \frac{\log(S/K) + (r - q + \sigma^2/2)T}{\sigma\sqrt{T}}, \quad \underline{d}_2 = \underline{d}_1 - \sigma\sqrt{T}. \\
k_1 &= D\left(\frac{H^2}{SK}, \sigma^2 + \mu, \sigma, T\right) = \frac{\log(H^2/(SK)) + (r - q + \sigma^2/2)T}{\sigma\sqrt{T}}, \quad k2 = k_1 - \sigma\sqrt{T}. \\
\overline{k}_1 &= D\left(\frac{H^2}{S\overline{K}}, \sigma^2 + \mu, \sigma, T\right) = \frac{\log(H^2/(SK)) + (r - q + \sigma^2/2)T}{\sigma\sqrt{T}}, \quad \overline{k}2 = \overline{k}_1 - \sigma\sqrt{T}. \\
\underline{k}_1 &= D\left(\frac{H^2}{S\underline{K}}, \sigma^2 + \mu, \sigma, T\right) = \frac{\log(H^2/(SK)) + (r - q + \sigma^2/2)T}{\sigma\sqrt{T}}, \quad \underline{k}2 = \underline{k}_1 - \sigma\sqrt{T}.
\end{aligned}$$

8.1 Reflection Principle

Let W_t be a standard Brownian motion($W_0 = 0$), and define

$$M_t := \sup_{0\le s\le t} W_s, \quad m_t := \inf_{0\le s\le t} W_s,$$
$$T_a := \inf\{t : W_t = a\}, \quad a > 0.$$

Since Brownian motion has continuous sample paths, it can be easily shown that

$$T_a = \inf\{t : M_t = a\} \qquad M_t = \inf\{a : T_a \ge t\}.$$

Moreover, for each a, T_a is a stopping time and the map $a \mapsto T_a$ is increasing and left-continuous.

Theorem 8.1.1 (Reflection Principle). For any $a > 0, h > 0$, we have

$$\begin{aligned} P(M_t \ge a, W_t \ge a+h) &= P(M_t \ge a, W_t \le a-h), \\ P(M_t \ge a) &= P(T_a \le t) = 2P(W_t \ge a) = P(|W_t| \ge a). \end{aligned}$$

Proof.

$$\begin{aligned} P(M_t \ge a) &= P(M_t \ge a, W_t \ge a) + P(M_t \ge a, W_t < a) \\ &= 2P(M_t \ge a, W_t \ge a) \\ &= 2P(W_t \ge a) \\ &= P(|W_t| \ge a). \end{aligned}$$

■

Example 8.1.2. Show that

$$P\Big(M_t = 0\Big) = 0.$$

Sol. For $\epsilon > 0$, by reflection principle we have

$$P\Big(M_t \ge \epsilon\Big) = 2P(W_t \ge \epsilon).$$

As ϵ goes to 0, we have

$$P\Big(M_t > 0\Big) = 2P(W_t > 0) = 1.$$

Hence we have done the proof. □

Theorem 8.1.3. For $a > 0$,

$$\begin{aligned} P(M_t \le a) &= \int_0^a \frac{2}{\sqrt{2\pi t}} \exp\left(-\frac{y^2}{2t}\right) dy, \\ P(T_a \le t) &= \int_a^\infty \frac{2}{\sqrt{2\pi t}} \exp\left(-\frac{y^2}{2t}\right) dy. \end{aligned}$$

Also, the PDF for T_a is given by

$$f_{T_a}(t) = \frac{a}{\sqrt{2\pi t^3}} \exp\left(-\frac{a^2}{2t}\right).$$

The distribution of T_a is called the inverse Gaussian or Wald distribution.

Proof.

$$\begin{aligned}P(M_t \le a) &= 1 - 2P(W_t \ge a) \\ &= 1 - 2\Big(\frac{1}{2} - P(0 \le W_t \le a)\Big) \\ &= 2P(0 \le W_t \le a) \\ &= \frac{2}{\sqrt{2\pi t}} \int_0^a e^{-\frac{y^2}{2t}}\, dy.\end{aligned}$$

Since $P(T_a \le t) = P(M_t \ge a)$,

$$\begin{aligned}P(T_a \le t) &= 1 - P(M_t \le a) \\ &= \frac{2}{\sqrt{2\pi t}} \int_a^\infty e^{-\frac{y^2}{2t}}\, dy \\ &= \frac{2}{\sqrt{2\pi}} \int_{\frac{a}{\sqrt{t}}}^\infty e^{-\frac{z^2}{2}}\, dz.\end{aligned}$$

Upon differentiation with respect to t,we can derive the density function for T_a as

$$\begin{aligned}f_{T_a}(t) &= \frac{d}{dt} P(T_a \le t) \\ &= \frac{d}{dt}\left(\frac{2}{\sqrt{2\pi}} \int_{\frac{a}{\sqrt{t}}}^\infty e^{-\frac{z^2}{2}}\, dz\right) \\ &= \frac{2}{\sqrt{2\pi}}\left(0 - e^{-\frac{a^2}{2t}}\left(-\frac{1}{2}\frac{a}{\sqrt{t^3}}\right)\right) \\ &= \frac{a}{\sqrt{2\pi t^3}}\, e^{-\frac{a^2}{2t}}.\end{aligned}$$

■

Note that this is a special case of Corollary(6.1.3) and Eq(6.1.3) with $\mu = 0, \alpha = 0$ and $\sigma = 1$.

Theorem 8.1.4. The joint density of (W_t, M_t) is given by on $\{(a, b) : a \le b, b \ge 0\}$ [1] by

$$f_{W_t,M_t}(a, b) = \sqrt{\frac{2}{\pi t^3}}(2b - a) \exp\left(-\frac{(2b - a)^2}{2t}\right)$$

and the joint density of (W_t, m_t) is given by on $\{(a, b) : a \ge b, b \le 0\}$ by

$$f_{W_t,m_t}(a, b) = \sqrt{\frac{2}{\pi t^3}}(a - 2b) \exp\left(-\frac{(2b - a)^2}{2t}\right).$$

Proof. Applying the reflection principle, we can write, for $a < b$,

$$P(W_t \le a, M_t > b) = P(W_t \le b - (b - a), M_t > b)$$

[1] Since $W_0 = M_0 = 0$, we only consider the case $b > 0$.

$$\begin{aligned} &= P(W_t \geq b + (b-a), M_t > b) \\ &= P(W_t \geq 2b - a), \quad \text{by } \{W_t \geq 2b-a\} \subset \{M_t > b\}. \end{aligned}$$

Therefore

$$\begin{aligned} P(W_t \leq a, M_t > b) &= \int_{2b-a}^{\infty} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right) dx \\ &= \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(2b-y)^2}{2t}\right) dy, \quad y := 2b - x. \end{aligned}$$

By differentiation w.r.t a and b, we complete the proof:

$$f_{W_t,M_t}(a,b) = -\frac{\partial^2}{\partial a \partial b} P(W_t \leq a, M_t > b).$$

Similarly, we have

$$\begin{aligned} P(W_t \geq a, m_t < b) &= P(W_t \geq b + (a-b), m_t < b) \\ &= P(W_t \leq b - (a-b), m_t < b) \\ &= P(W_t \leq 2b - a), \quad \text{by } \{W_t \leq 2b-a\} \subset \{m_t < b\}. \end{aligned}$$

Therefore

$$P(W_t \geq a, m_t < b) = \int_{-\infty}^{2b-a} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right) dx.$$

By differentiation w.r.t a and b, we complete the proof.

$$f_{W_t,m_t}(a,b) = -\frac{\partial^2}{\partial a \partial b} P(W_t \geq a, m_t < b).$$

■

Remark 8.1.5. Note that the following integrations.

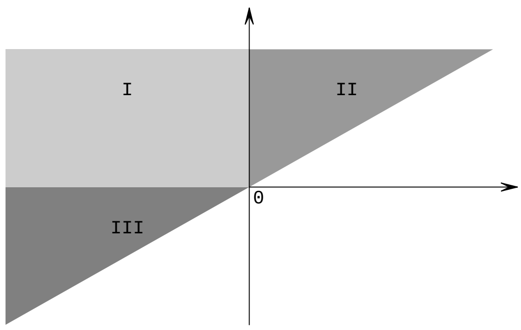


Figure 8.1: Integration region

1. First integration(I):

$$\begin{aligned}
& \int_{-\infty}^{0}\int_{0}^{\infty}\sqrt{\frac{2}{\pi t^3}}(2b-a)\exp\left(-\frac{(2b-a)^2}{2t}\right)db\,da \\
= & \int_{-\infty}^{0}\left[-\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{(2b-a)^2}{2t}\right)\right]_0^{\infty}da \\
= & \int_{-\infty}^{0}\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{a^2}{2t}\right)da \\
= & \frac{1}{2}.
\end{aligned}$$

2. Second integration(II):

$$\begin{aligned}
& \int_{0}^{\infty}\int_{a}^{\infty}\sqrt{\frac{2}{\pi t^3}}(2b-a)\exp\left(-\frac{(2b-a)^2}{2t}\right)db\,da \\
= & \int_{0}^{\infty}\left[-\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{(2b-a)^2}{2t}\right)\right]_a^{\infty}da \\
= & \int_{0}^{\infty}\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{a^2}{2t}\right)da \\
= & \frac{1}{2}.
\end{aligned}$$

3. Third integration(III): In fact, the following integration is not necessary to us in this case.

$$\int_{-\infty}^{0}\int_{a}^{0}\sqrt{\frac{2}{\pi t^3}}(2b-a)\exp\left(-\frac{(2b-a)^2}{2t}\right)db\,da \quad = \quad 0.$$

4. Thus we have the following result.

$$\int_{-\infty}^{\infty}\int_{a}^{\infty}\sqrt{\frac{2}{\pi t^3}}(2b-a)\exp\left(-\frac{(2b-a)^2}{2t}\right)db\,da \quad = \quad \frac{1}{2}+\frac{1}{2}+0 \quad = \quad 1.$$

Theorem 8.1.6. If $Z_T \sim N(\mu T, \sigma^2 T)$ and $G_u(x) = P(Z_T \le x, M_T^Z > b)$ for $x < b$ is the defective distribution function of Z_T on $(-\infty, b)$ when the upper barrier is breached for some $t \in (0, T)$, then

$$\begin{aligned}
G_u(x) &= P(Z_T \le x, M_T^Z > b) = \exp\left(\frac{2\mu b}{\sigma^2}\right)N\left(\frac{x-2b-\mu T}{\sigma\sqrt{T}}\right), \\
G_u'(x) &= \frac{\exp\left(2\mu b/\sigma^2\right)}{\sigma\sqrt{T}}\phi\left(\frac{x-2b-\mu T}{\sigma\sqrt{T}}\right)
\end{aligned}$$

where

$$\phi(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right).$$

Also the PDF of Z_T is given by

$$f_{Z_T}(x) = \frac{1}{\sigma\sqrt{T}}\phi\left(\frac{x-\mu T}{\sigma\sqrt{T}}\right).$$

Proof. To prove the theorem, we first assume $\sigma = 1$. Note that by the Girsanov's Theorem,

$$Z_t = \mu t + W_t$$

is a standard Brownian motion under the probability measure Q, where

$$dQ = \exp\left(-\mu W_t - \frac{1}{2}\mu^2 t\right) dP = \exp\left(-\mu Z_t + \frac{1}{2}\mu^2 t\right) dP.$$

Hence, by Theorem (8.1.4) we can write

$$\begin{aligned} P(Z_T \le x, M_T^Z > b) &= E_Q\Big[\mathbf{1}_{\{Z_T \le x, M_T^Z > b\}} \exp\big(\mu Z_T - \frac{\mu^2}{2}T\big)\Big] \\ &= \int_{-\infty}^{x}\int_{b}^{\infty} \exp\left(\mu z - \frac{\mu^2}{2}T\right)\sqrt{\frac{2}{\pi T}}\frac{1}{T}(2y-z)\exp\left(-\frac{(2y-z)^2}{2T}\right) dy dz \\ &= \int_{-\infty}^{x} \exp\left(\mu z - \frac{\mu^2}{2}T\right)\frac{1}{\sqrt{2\pi T}}\exp\left(-\frac{(2b-z)^2}{2T}\right) dz. \end{aligned}$$

Therefore

$$G_u'(x) = \exp\left(\mu x - \frac{\mu^2}{2}T\right)\frac{1}{\sqrt{2\pi T}}\exp\left(-\frac{(2b-x)^2}{2T}\right).$$

Since

$$\begin{aligned} -\frac{(x-2b)^2 - 2\mu T x + \mu^2 T^2}{2T} &= -\frac{x^2 - 4bx + 4b^2 - 2\mu T x + \mu^2 T^2}{2T} \\ &= -\frac{x^2 - 2(2b+\mu T)x + 4b^2 + \mu^2 T^2}{2T} \\ &= -\frac{x^2 - 2(2b+\mu T)x + (2b+\mu T)^2 - 4\mu b T}{2T} \\ &= -\frac{(x-2b-\mu T)^2}{2T} + 2\mu b, \end{aligned}$$

$$\begin{aligned} G_u'(x) &= \frac{1}{\sqrt{2\pi T}}\exp(2\mu b)\exp\left(-\frac{(x-2b-\mu T)^2}{2T}\right) \\ &= \frac{\exp(2\mu b)}{\sqrt{T}}\phi\left(\frac{x-2b-\mu T}{\sqrt{T}}\right) \end{aligned}$$

Finally, if $\sigma \ne 1$, consider $\widehat{Z}_t = Z_t/\sigma$, and we have

$$\begin{aligned} P(Z_T \le x, M_T^Z > b) &= P\left(\widehat{Z}_T \le \frac{x}{\sigma}, M_T^Z > \frac{b}{\sigma}\right) \\ &= \exp\left(2\frac{\mu}{\sigma}\cdot\frac{b}{\sigma}\right)\frac{1}{\sqrt{T}}\int_{-\infty}^{\frac{x}{\sigma}}\phi\left(\frac{y - 2\frac{b}{\sigma} - \frac{\mu}{\sigma}T}{\sqrt{T}}\right) dy. \end{aligned}$$

By differentiation, we get

$$G_u'(x) = \frac{\exp\left(2\mu b/\sigma^2\right)}{\sigma\sqrt{T}}\phi\left(\frac{x-2b-\mu T}{\sigma\sqrt{T}}\right).$$

When $b = 0$, we have the probability density function of Z_T. ■

Theorem 8.1.7. If $Z_T \sim N(\mu T, \sigma^2 T)$,

$$P(Z_T \geq x, m_T^Z < b) \quad = \quad \exp\left(\frac{2\mu b}{\sigma^2}\right) N\left(\frac{2b - x + \mu T}{\sigma\sqrt{T}}\right).$$

Proof. To prove the theorem, we first assume $\sigma = 1$. By Theorem (8.1.4) we have

$$\begin{aligned}
P(Z_T \geq x, m_T^Z < b) &= E_Q\Big[\mathbf{1}_{\{Z_T \geq x, m_T^Z > b\}} \exp\left(\mu Z_T - \frac{\mu^2}{2}T\right)\Big] \\
&= \int_x^\infty \int_{-\infty}^b \exp\left(\mu z - \frac{\mu^2}{2}T\right)\sqrt{\frac{2}{\pi T}}\frac{1}{T}(z - 2y)\exp\left(-\frac{(2y - z)^2}{2T}\right) dy dz \\
&= \int_x^\infty \exp\left(\mu z - \frac{\mu^2}{2}T\right)\frac{1}{\sqrt{2\pi T}}\exp\left(-\frac{(2b - z)^2}{2T}\right) dz.
\end{aligned}$$

Generally when $\sigma \neq 1$, let $\widehat{Z}_t = Z_t/\sigma$. Then we have

$$\begin{aligned}
P(Z_T \geq x, m_T^Z < b) &= P\left(\widehat{Z}_T \geq \frac{x}{\sigma}, m_T^Z < \frac{b}{\sigma}\right) \\
&= \int_{\frac{x}{\sigma}}^\infty \exp\left(\frac{\mu}{\sigma}z - \frac{\mu^2}{2\sigma^2}T\right)\frac{1}{\sqrt{2\pi T}}\exp\left(-\frac{(2b/\sigma - z)^2}{2T}\right) dz \\
&= \int_x^\infty \exp\left(\frac{\mu z}{\sigma^2} - \frac{\mu^2}{2\sigma^2}T\right)\frac{1}{\sqrt{2\pi T}\,\sigma}\exp\left(-\frac{(2b - z)^2}{2T\sigma^2}\right) dz \\
&= \exp\left(\frac{2\mu b}{\sigma^2}\right) N\left(\frac{2b - x + \mu T}{\sigma\sqrt{T}}\right).
\end{aligned}$$

■

✎ **Conditional Density Functions**

Suppose that

$$\begin{aligned}
Z_t &= \mu t + \sigma W(t), \\
M_t^Z &= \max_{0 \leq s \leq t} Z_s, \quad m_t^Z = \min_{0 \leq s \leq t} Z_s.
\end{aligned}$$

From Theorem (8.1.6), the joint probability function of Z_T and M_T^Z, denoted by $f_{Z_T, M_T^Z}(x, y)$, is given

$$\begin{aligned}
f_{Z_T, M_T^Z}(x, y) &= -\frac{\partial^2}{\partial x \partial y} P(Z_T \leq x, M_T^Z > y) \\
&= -\frac{\partial}{\partial y}\left(\frac{\exp\left(2\mu y/\sigma^2\right)}{\sigma\sqrt{T}}\phi\left(\frac{x - 2y - \mu T}{\sigma\sqrt{T}}\right)\right) \\
&= \frac{2(2y - x)}{\sqrt{2\pi T^3}\,\sigma^3}\exp\left(\frac{4\mu y T - (x - 2y - \mu T)^2}{2\sigma^2 T}\right)
\end{aligned}$$

and the marginal density function of Z_T can be obtained by

$$f_{Z_T}(x) \quad = \quad \frac{1}{\sigma\sqrt{T}}\phi\left(\frac{x - \mu T}{\sigma\sqrt{T}}\right).$$

Thus the conditional probability density function of M_T^Z is given by

$$\frac{f_{Z_T,M_T^Z}(x,y)}{f_{Z_T}(x)} = \frac{2(2y-x)}{\sigma^2 T}\exp\left(-\frac{2y(y-x)}{\sigma^2 T}\right), \quad Z_T = x, y > x.$$

Hence the conditional probability

$$\begin{aligned} P(M_T^Z \geq b \mid Z_T = a) &= \int_b^\infty \frac{2(2y-a)}{\sigma^2 T}\exp\left(-\frac{2y(y-a)}{\sigma^2 T}\right) dy \\ &= \exp\left(-\frac{2b(b-a)}{\sigma^2 T}\right), \quad b > a. \end{aligned}$$

From Theorem (8.1.7), the joint probability function of Z_T and m_T^Z, denoted by $f_{Z_T,m_T^Z}(x,y)$, is given

$$\begin{aligned} f_{Z_T,m_T^Z}(x,y) &= -\frac{\partial^2}{\partial x \partial y} P(Z_T \geq x, m_T^Z < y) \\ &= -\frac{\partial}{\partial y}\left(\frac{\exp\left(2\mu y/\sigma^2\right)}{\sigma\sqrt{T}}\phi\left(\frac{2y-x+\mu T}{\sigma\sqrt{T}}\right)\right) \\ &= \frac{2(x-2y)}{\sqrt{2\pi T^3}\,\sigma^3}\exp\left(\frac{4\mu y T-(x-2y-\mu T)^2}{2\sigma^2 T}\right) \end{aligned}$$

Thus the conditional probability density function of m_T^Z is given by

$$\frac{f_{Z_T,m_T^Z}(x,y)}{f_{Z_T}(x)} = \frac{2(x-2y)}{\sigma^2 T}\exp\left(-\frac{2y(y-x)}{\sigma^2 T}\right), \quad Z_T = x, y < x.$$

Now we have the conditional probability

$$\begin{aligned} P(m_T^Z \leq b \mid Z_T = a) &= \int_{-\infty}^b \frac{2(a-2y)}{\sigma^2 T}\exp\left(-\frac{2y(y-a)}{\sigma^2 T}\right) dy \\ &= \exp\left(-\frac{2b(b-a)}{\sigma^2 T}\right), \quad b < a. \end{aligned} \tag{8.1.1}$$

Example 8.1.8. Suppose that for $t \geq 0$,

$$\begin{aligned} S_t &= S_0 \exp\Big(\mu t + \sigma W_t\Big), \\ M_t^S &= \max_{0\leq s\leq t} S_s, \quad m_t^S = \min_{0\leq s\leq t} S_s. \end{aligned}$$

When $H > S_0$ we can obtain the conditional probability $P(M_T^S \geq H \,|\, S_T)$ by setting

$$a = \log\frac{S_T}{S_0}, \quad b = \log\frac{H}{S_0},$$

i.e.

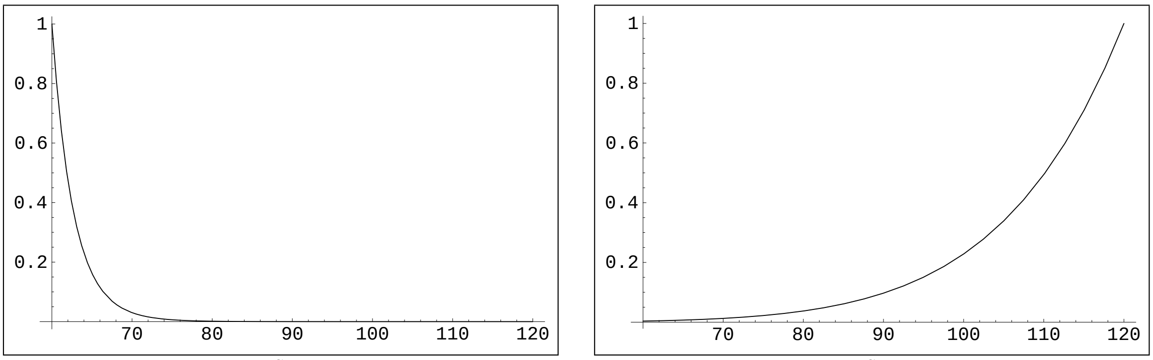
$$\begin{aligned} P(M_T^S \geq H \,|\, S_T) &= \exp\left(-\frac{2b(b-a)}{\sigma^2 T}\right) \\ &= \exp\left(\frac{2\log\frac{S_T}{H}\ \log\frac{H}{S_0}}{\sigma^2 T}\right) \end{aligned}$$

$$= \left(\frac{H}{S_0}\right)^{\frac{2\log\frac{S_T}{H}}{\sigma^2 T}}.$$

When $L < S_0$ we have

$$\begin{aligned} P(m_T^S \leq L \,|\, S_T) &= \exp\left(\frac{2\log\frac{S_T}{L}\ \log\frac{L}{S_0}}{\sigma^2 T}\right) \\ &= \left(\frac{L}{S_0}\right)^{\frac{2\log\frac{S_T}{L}}{\sigma^2 T}}. \end{aligned}$$

We will make use of these probabilities in Section (18.8). □



(a) $P(m_T^S \leq L \,|\, S_T)$ (b) $P(M_T^S \geq H \,|\, S_T)$

Figure 8.2: Conditional Probability: $S_0 = 100, H = 120, L = 60, \sigma = 0.3, T = 0.5$

8.2 Single Barrier Options

Theorem 8.2.1. The premium for Up-and-In call option is

$$P_{\text{call}}^{\text{u,in}} = \left(\frac{H}{S}\right)^{\frac{2\mu}{\sigma^2}} \left[\frac{H^2}{S} e^{-qT}\Big(N(-\bar{k}_1)-N(-k_1)\Big)-e^{-rT}K\Big(N(-\bar{k}_2)-N(-k_2)\Big)\right]+e^{-qT}SN(\bar{d}_1)-e^{-rT}KN(\bar{d}_2).$$

This formula is meaningless when $S \geq H$.

Proof. Note that if $H \leq K$, Up-and-In call option is vanilla option. In this case, $\bar{k}_1 = k_1$, $\bar{k}_2 = k_2$ $\bar{d}_1 = d_1$ and $\bar{d}_2 = d_2$.

Note that $\overline{K} = \max(K, H), k = \log(K/S), \overline{k} = \log(\overline{K}/S), b = \log(H/S)$. In the general case, the arbitrage pricing technique allows us to write

$$\begin{aligned}
P_{\text{call}}^{\text{u,in}} &= e^{-rT} E\Big[(S_T - K)^+ \mathbf{1}_{\{M_T^S > H\}}\Big] \\
&= e^{-rT} E\Big[S_T - K : S_T > K, M_T^S > H\Big] \\
&= e^{-rT} E\Big[Se^{Z_T} - K : Z_T > k, M_T^Z > b\Big] \\
&= e^{-rT} E\Big[Se^{Z_T} - K : k < Z_T < \overline{k}, M_T^Z > b\Big] + e^{-rT} E\Big[Se^{Z_T} - K : \overline{k} < Z_T\Big] \\
&= e^{-rT} \int_k^{\overline{k}} (Se^x - K) g_u(x) dx + e^{-rT} \int_{\overline{k}}^{\infty} (Se^x - K) f_{Z_T} dx.
\end{aligned}$$

where $g_u(x)$ and f_{Z_T} are define in Theorem (8.1.6).

$$\begin{aligned}
P_{\text{call}}^{\text{u,in}} &= e^{-rT} \frac{\exp(2\mu b/\sigma^2)}{\sigma\sqrt{T}} \frac{S}{\sqrt{2\pi}} \int_k^{\overline{k}} e^x \cdot \exp\left(-\frac{(x-2b-\mu T)^2}{2\sigma^2 T}\right) dx \\
&\quad -e^{-rT} \frac{\exp(2\mu b/\sigma^2)}{\sigma\sqrt{T}} \frac{K}{\sqrt{2\pi}} \int_k^{\overline{k}} \exp\left(-\frac{(x-2b-\mu T)^2}{2\sigma^2 T}\right) dx \\
&\quad +e^{-rT} \frac{1}{\sqrt{2\pi T}\sigma} \int_{\overline{k}}^{\infty} (Se^x - K) \exp\left(-\frac{(x-\mu T)^2}{2\sigma^2 T}\right) dx \\
&= e^{-rT} I - e^{-rT} II + III.
\end{aligned}$$

1. I: We have

$$I = \exp\left(\frac{2\mu b}{\sigma^2}\right) \frac{S}{\sigma\sqrt{2\pi T}} \int_k^{\overline{k}} \exp\left(x - \frac{(x-2b-\mu T)^2}{2\sigma^2 T}\right) dx.$$

Note that

$$\begin{aligned}
x - \frac{(x-2b-\mu T)^2}{2\sigma^2 T} &= -\frac{(x-2b-\mu T)^2 - 2\sigma^2 T x}{2\sigma^2 T} \\
&= -\frac{(x-2b-\sigma^2 T-\mu T)^2 + (2b+\mu T)^2 - (2b+\mu T+\sigma^2 T)^2}{2\sigma^2 T} \\
&= -\frac{(x-2b-\sigma^2 T-\mu T)^2 - 2(2b+\mu T)\sigma^2 T - \sigma^4 T^2}{2\sigma^2 T} \\
&= -\frac{(x-2b-\sigma^2 T-\mu T)^2}{2\sigma^2 T} + 2b + \mu T + \frac{\sigma^2}{2} T,
\end{aligned}$$

$$\begin{aligned}
\frac{k-2b-(\sigma^2+\mu)T}{\sigma\sqrt{T}} &= \frac{\log(\frac{K}{S})-2\log(\frac{H}{S})-(\sigma^2+\mu)T}{\sigma\sqrt{T}} \\
&= \frac{\log(\frac{SK}{H^2})-(r-q+\frac{\sigma^2}{2})T}{\sigma\sqrt{T}} \\
&= -k_1, \\
\frac{\overline{k}-2b-(\sigma^2+\mu)T}{\sigma\sqrt{T}} &= -\overline{k}_1. \\
\exp\left(\frac{2\mu b}{\sigma^2}\right) &= \exp\left(\frac{2\mu\log(\frac{H}{S})}{\sigma^2}\right) = \left(\frac{H}{S}\right)^{\frac{2\mu}{\sigma^2}}, \\
\exp\left(2b+\frac{\sigma^2+2\mu}{2}T\right) &= \exp(2b)\cdot\exp\left(\frac{\sigma^2+2\mu}{2}T\right) \\
&= \left(\frac{H}{S}\right)^2\cdot\exp\left(\frac{\sigma^2+2(r-q-\frac{1}{2}\sigma^2)}{2}T\right) \\
&= \left(\frac{H}{S}\right)^2\cdot e^{(r-q)T}.
\end{aligned}$$

Then we have

$$\begin{aligned}
I &= \exp\left(\frac{2\mu b}{\sigma^2}\right)\frac{S}{\sigma\sqrt{2\pi T}}\exp\left(2b+\frac{\sigma^2+2\mu}{2}T\right)\int_k^{\overline{k}}\exp\left(-\frac{(x-2b-\sigma^2T-\mu T)^2}{2\sigma^2T}\right)dx \\
&= \exp\left(\frac{2\mu b}{\sigma^2}\right)\exp\left(2b+\frac{\sigma^2+2\mu}{2}T\right)S\int_{-k_1}^{-\overline{k}_1}\phi(x)dx \\
&= \exp\left(\frac{2\mu b}{\sigma^2}\right)\exp\left(2b+\frac{\sigma^2+2\mu}{2}T\right)S\big[N(-\overline{k}_1)-N(-k_1)\big] \\
&= \left(\frac{H}{S}\right)^{\frac{2\mu}{\sigma^2}}\cdot\left(\frac{H^2}{S}\right)\cdot e^{(r-q)T}\big[N(-\overline{k}_1)-N(-k_1)\big].
\end{aligned}$$

2. II: We have

$$II = \exp\left(\frac{2\mu b}{\sigma^2}\right)\frac{K}{\sigma\sqrt{2\pi T}}\int_k^{\overline{k}}\exp\left(-\frac{(x-2b-\mu T)^2}{2\sigma^2T}\right)dx.$$

$$\begin{aligned}
\frac{k-2b-\mu T}{\sigma\sqrt{T}} &= \frac{\log(\frac{K}{S})-2\log(\frac{H}{S})-\mu T}{\sigma\sqrt{T}} \\
&= \frac{\log(\frac{SK}{H^2})-(r-q-\frac{\sigma^2}{2})T}{\sigma\sqrt{T}} \\
&= \frac{\log(\frac{SK}{H^2})-(r-q+\frac{\sigma^2}{2})T}{\sigma\sqrt{T}}+\sigma\sqrt{T} \\
&= -(k_1-\sigma\sqrt{T}) = -k_2, \\
\frac{\overline{k}-2b-\mu T}{\sigma\sqrt{T}} &= -\overline{k}_2.
\end{aligned}$$

Now,

$$II = \left(\frac{H}{S}\right)^{\frac{2\mu}{\sigma^2}}\cdot K\big[N(-k_2)-N(-\overline{k}_2)\big]$$

$$= \left(\frac{H}{S}\right)^{\frac{2\mu}{\sigma^2}} \cdot K\big[N(\overline{k}_2) - N(k_2)\big].$$

3. III:

   (a) When $H \le K$, the Up-and-In barrier option is the same as the ordinary call option with strike K. Therefore

$$III = e^{-qT}SN(\bar{d}_1) - e^{-rT}KN(\bar{d}_2).$$

   (b) When $H > K$,

$$\begin{aligned} III &= e^{-rT}\int_{\overline{k}}^{\infty}(Se^x - K)f_{Z_T}dx \\ &= e^{-rT}\int_{\overline{k}}^{\infty}(Se^x - H)f_{Z_T}dx + e^{-rT}\int_{\overline{k}}^{\infty}(H-K)f_{Z_T}dx \\ &= (\text{ call option with strike } H\ ) + e^{-rT}\int_{\overline{k}}^{\infty}(H-K)f_{Z_T}dx \\ &= e^{-qT}SN(\bar{d}_1) - e^{-rT}HN(\bar{d}_2) + e^{-rT}(H-K)N(\bar{d}_2) \\ &= e^{-qT}SN(\bar{d}_1) - e^{-rT}KN(\bar{d}_2). \end{aligned}$$

We completed the proof. ■

Theorem 8.2.2. The premium for Down-and-In call option is

$$P_{\text{call}}^{\text{d,in}} = \left(\frac{H}{S}\right)^{\frac{2\mu}{\sigma^2}}\left[\frac{H^2}{S}e^{-qT}N(\overline{k}_1) - e^{-rT}KN(\overline{k}_2)\right] + e^{-qT}S\big[N(d_1) - N(\bar{d}_1)\big] - e^{-rT}K\big[N(d_2) - N(\bar{d}_2)\big].$$

Proof.

$$\begin{aligned} P_{\text{call}}^{\text{d,in}} &= e^{-rT}E\Big[(S_T - K)^+ : m_T^S < H\Big] \\ &= e^{-rT}E\Big[S_T - K : S_T > K, m_T^S < H\Big] \\ &= e^{-rT}E\Big[Se^{Z_T} - K : Z_T > k, m_T^Z < b\Big] \\ &= e^{-rT}E\Big[Se^{Z_T} - K : Z_T > k, Z_T \ge b, m_T^Z < b\Big] + e^{-rT}E\Big[Se^{Z_T} - K : Z_T > k, Z_T < b\Big] \\ &= e^{-rT}\int_{\overline{k}}^{\infty}(Se^x - K)g(x)dx + e^{-rT}\int_{k}^{\overline{k}}(Se^x - K)f(x)dx \\ &= I + II. \end{aligned}$$

Now proceed similarly as in the proof of Up-and-In call option, we can get

$$\begin{aligned} I &= \left(\frac{H}{S}\right)^{\frac{2\mu}{\sigma^2}}\left[\frac{H^2}{S}e^{-qT}N(\overline{k}_1) - e^{-rT}KN(\overline{k}_2)\right] \\ II &= e^{-qT}S\big[N(d_1) - N(\bar{d}_1)\big] - e^{-rT}K\big[N(d_2) - N(\bar{d}_2)\big] \end{aligned}$$

which completes the proof. ■

8.3 Alternative Approach To Barrier Options

In this section we follow Björk(1998). First let us introduce the notations.

✎ **Notation & Basics**

1. Φ_L, Φ^L: For a given function Φ, the function Φ_L and Φ^L are defined by

$$\Phi_L(x) := \begin{cases} \Phi(x) & \text{for } x > L, \\ 0 & \text{for } x \leq L. \end{cases} \qquad \Phi^L(x) := \begin{cases} \Phi(x) & \text{for } x < L, \\ 0 & \text{for } x \geq L. \end{cases} \tag{8.3.1}$$

(8.3.2)

2. S_L: Denote the process S with absorbed at L.
3. Z_{LO}:

$$Z_{LO} = \begin{cases} \Phi(S(T)), & \text{if } S(t) > L \text{ for all } t \in [0,T], \\ 0, & \text{if } S(t) \leq L \text{ for some } t \in [0,T]. \end{cases}$$

4. Z^{LO}:

$$Z^{LO} = \begin{cases} \Phi(S(T)), & \text{if } S(t) < L \text{ for all } t \in [0,T], \\ 0, & \text{if } S(t) \geq L \text{ for some } t \in [0,T]. \end{cases}$$

5. $d_i (i = 1, 2)$ and μ:

$$\begin{aligned} d_1(x) &= \frac{\log x + \left(r - q + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \\ d_2(x) &= \frac{\log x + \left(r - q - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \\ \mu &= r - q - \frac{1}{2}\sigma^2, \end{aligned}$$

where r is the riskless interest rate and q is the continuous dividend rate.

6. K, L, K^* and K_*:

$$\begin{aligned} K &= \text{Strike price}, \\ L &= \text{Upper barrier or lower barrier}, \\ K^* &= \max(K, L), \\ K_* &= \min(K, L). \end{aligned}$$

7. Contingent Claims:

$$\begin{aligned} ST(x) &:= x, \\ BO(x) &:= 1, \forall x, \\ H(x; L) &:= \mathbf{1}_{\{x \geq L\}} = BO_L(x), \\ \widetilde{H}(x; L) &:= \mathbf{1}_{\{x \leq L\}}, \\ C(x; K) &:= (x - K)^+, \\ P(x; K) &:= (K - x)^+. \end{aligned}$$

8. Pricing functions:

$$\begin{aligned}
\mathbf{ST}(t,s) &:= e^{-r(T-t)}E_{t,s}\left[ST(S(T))|\mathscr{F}_t\right] = se^{-q(T-t)}, && (8.3.3)\\
\mathbf{BO}(t,s) &:= e^{-r(T-t)}E_{t,s}\left[BO(S(T))|\mathscr{F}_t\right] = e^{-r(T-t)}, && (8.3.4)\\
\mathbf{H}(t,s;K) &:= e^{-r(T-t)}E_{t,s}\left[H(S(T);K)|\mathscr{F}_t\right]\\
&= e^{-r(T-t)}N\left[\frac{\log\frac{s}{K}+\left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] && (8.3.5)\\
&= e^{-r(T-t)}N\left[d_2\left(\frac{s}{K}\right)\right], && (8.3.6)\\
\widetilde{\mathbf{H}}(t,s;K) &:= e^{-r(T-t)}E_{t,s}\left[\widetilde{H}(S(T);K)|\mathscr{F}_t\right]\\
&= e^{-r(T-t)}-\mathbf{H}(t,s;K)\\
&= e^{-r(T-t)}-e^{-r(T-t)}N\left[\frac{\log\frac{s}{K}+\left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] && (8.3.7)\\
&= e^{-r(T-t)}N\left[-d_2\left(\frac{s}{K}\right)\right], && (8.3.8)\\
\mathbf{C}(t,s;K) &= \text{Black-Scholes formula for call option} && (8.3.9)\\
&= se^{-q(T-t)}N\left[\frac{\log\frac{s}{K}+\left(r-q+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right]\\
&\quad -Ke^{-r(T-t)}N\left[\frac{\log\frac{s}{K}+\left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] && (8.3.10)\\
&= se^{-q(T-t)}N\left[d_1\left(\frac{s}{K}\right)\right]-Ke^{-r(T-t)}N\left[d_2\left(\frac{s}{K}\right)\right], && (8.3.11)\\
\mathbf{P}(t,s;K) &= Ke^{-r(T-t)}N\left[-d_2\left(\frac{s}{K}\right)\right]-se^{-q(T-t)}N\left[-d_1\left(\frac{s}{K}\right)\right]. && (8.3.12)
\end{aligned}$$

$$\mathbf{H}(T,s;L)+\widetilde{\mathbf{H}}(T,s;L)=1, \quad \text{a.s.}$$
$$\mathbf{H}(t,s;L)+\widetilde{\mathbf{H}}(t,s;L)=e^{-r(T-t)}.$$

Example 8.3.1. Let $C(x;K)=(x-K)^+$. When $K<L$, then we have

$$C^L(x;K) = C(x;K)-C(x;L)-(L-K)H(x;L), \quad \text{a.s.}$$

In this case, we need not care the point mass. Now, we have

$$\begin{aligned}
& F\Big(t,s,C^L(*;K)\Big)\\
=\ & e^{-r(T-t)}E_{t,s}\left[C^L(S(T);K)|\mathscr{F}_t\right]\\
=\ & \mathbf{C}(t,s;K)-\mathbf{C}(t,s;L)-(L-K)\mathbf{H}(t,s;L)\\
=\ & \mathbf{C}(t,s;K)-se^{-q(T-t)}N\left[d_1\left(\frac{s}{L}\right)\right]+Le^{-r(T-t)}N\left[d_2\left(\frac{s}{L}\right)\right]-(L-K)e^{-r(T-t)}N\left[d_2\left(\frac{s}{L}\right)\right]\\
=\ & \mathbf{C}(t,s;K)-se^{-q(T-t)}N\left[d_1\left(\frac{s}{L}\right)\right]+Ke^{-r(T-t)}N\left[d_2\left(\frac{s}{L}\right)\right]\\
=\ & se^{-q(T-t)}\left\{N\left[d_1\left(\frac{s}{K}\right)\right]-N\left[d_1\left(\frac{s}{L}\right)\right]\right\}-Ke^{-r(T-t)}\left\{N\left[d_2\left(\frac{s}{K}\right)\right]-N\left[d_2\left(\frac{s}{L}\right)\right]\right\}.
\end{aligned}$$

When $L < K$, $C^L(x;K) = 0$. Hence we have

$$\begin{aligned} & F\Big(t, s, C^L(*;K)\Big) \\ = \;& se^{-q(T-t)}\left\{N\left[d_1\left(\frac{s}{K}\right)\right] - N\left[d_1\left(\frac{s}{K^*}\right)\right]\right\} - Ke^{-r(T-t)}\left\{N\left[d_2\left(\frac{s}{K}\right)\right] - N\left[d_2\left(\frac{s}{K^*}\right)\right]\right\}, \end{aligned}$$

where $K^* = \max(K, L)$.

Note that in general

$$C^L(x;K) \;=\; C(x;K) - C(x;K^*) - (K^* - K)H(x;K^*), \quad \text{a.s.}$$

□

Example 8.3.2. For the put option payoff, $P(x;K)$, we have

$$P_L(x;K) \;=\; P(x;K) - P(x;K_*) - (K - K_*)\widetilde{H}(x;K_*), \quad \text{a.s.}$$

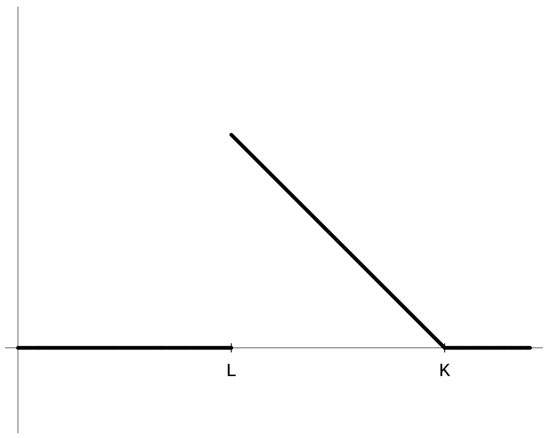


Figure 8.3: $P_L(x;K)$, with $L < K$.

Then we have

$$\begin{aligned} & F\Big(t, s; P_L(*;K)\Big) \\ = \;& F\Big(t, s; P(*;K)\Big) - F\Big(t, s; P(*;K_*)\Big) - (K - K_*)F\Big(t, s; H(x;K_*)\Big) \\ = \;& \mathbf{P}(t, s; K) - \mathbf{P}(t, s; K_*) - (K - K_*)\widetilde{\mathbf{H}}(t, s; K_*) \\ = \;& Ke^{-r(T-t)}N\left[-d_2\left(\frac{s}{K}\right)\right] - se^{-q(T-t)}N\left[-d_1\left(\frac{s}{K}\right)\right] \\ & -K_*e^{-r(T-t)}N\left[-d_2\left(\frac{s}{K_*}\right)\right] + se^{-q(T-t)}N\left[-d_1\left(\frac{s}{K_*}\right)\right] \\ & -(K - K_*)e^{-r(T-t)}N\left[-d_2\left(\frac{s}{K_*}\right)\right] \\ = \;& Ke^{-r(T-t)}\left\{N\left[-d_2\left(\frac{s}{K}\right)\right] - N\left[-d_2\left(\frac{s}{K_*}\right)\right]\right\} - se^{-q(T-t)}\left\{N\left[-d_1\left(\frac{s}{K}\right)\right] - N\left[-d_1\left(\frac{s}{K_*}\right)\right]\right\}. \end{aligned}$$

□

8.3.1 General In-Out Contracts

Theorem 8.3.3 (Down-and-Out Contracts). For a fixed T(maturity) and $T-$ contingent claim $\Phi(S(T))$, the pricing function, denoted by $F_{LO}(t,s;\Phi)$, of the corresponding down-and-out contract Z_{LO}, is given by , for $s > L$,

$$\begin{aligned}
F_{LO}(t,s;\Phi) &= e^{-r(T-t)}E_{t,s}\left[\Phi_L(S_L(T))\right] \\
&= e^{-r(T-t)}E_{t,s}\left[\Phi_L(S(T))\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} E_{t,\frac{L^2}{s}}\left[\Phi_L(S(T))\right] \\
&= F(t,s;\Phi_L) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} F\left(t,\frac{L^2}{s};\Phi_L\right),
\end{aligned}$$

where $\mu = r - q - \frac{1}{2}\sigma^2$.

Proof. Without loss of generality we may set $t = 0$. Under the risk-neutral measure we have

$$\begin{aligned}
F_{LO}(0,s;\Phi) &= e^{-rT}E_{0,s}\left[Z_{LO}\right] \\
&= e^{-rT}E_{0,s}\left[\Phi(S(T)\cdot \mathbf{1}_{\{\inf_{0\le t\le T} S(t)>L\}}\right] \\
&= e^{-rT}E_{0,s}\left[\Phi_L(S_L(T)\cdot \mathbf{1}_{\{\inf_{0\le t\le T} S(t)>L\}}\right] \\
&= e^{-rT}E_{0,s}\left[\Phi_L(S_L(T))\right] \\
&= e^{-rT}\int_L^\infty \Phi_L(s)h(s)ds \\
&= e^{-rT}\int_{\log L}^\infty \Phi_L(e^x)f(x)dx
\end{aligned}$$

where h and f are the density functions for the stochastic variables $S_L(T)$ and $X_{\log L}(T)$, respectively, where

$$X(t) = \log S(t), \quad S_L(t) = \exp\left(X_{\log L}(t)\right).$$

The density f is given by Theorem(6.1.2) as

$$\begin{aligned}
f(x) &= \varphi(x;\mu t + \log s, \sigma^2 t) - \exp\left(-\frac{2\mu(\log s - \log L)}{\sigma^2}\right)\varphi(x;\mu t - \log s + 2\log L, \sigma^2 t) \\
&= \varphi(x;\mu t + \log s, \sigma^2 t) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\varphi(x;\mu t - \log s + 2\log L, \sigma^2 t).
\end{aligned}$$

Thus we have

$$\begin{aligned}
& F_{LO}(0,s;\Phi) \\
=\ & e^{-rT}\int_{\log L}^\infty \Phi_L(e^x)f(x)dx \\
=\ & e^{-rT}\int_{\log L}^\infty \Phi_L(e^x)\varphi(x;\mu t + \log s,\sigma^2 t)dx - e^{-rT}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\int_{\log L}^\infty \Phi_L(e^x)\varphi(x;\mu t - \log s + 2\log L,\sigma^2 t)dx \\
=\ & e^{-rT}\int_{-\infty}^\infty \Phi_L(e^x)\varphi(x;\mu t + \log s,\sigma^2 t)dx - e^{-rT}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\int_{-\infty}^\infty \Phi_L(e^x)\varphi(x;\mu t - \log s + 2\log L,\sigma^2 t)dx
\end{aligned}$$

$$\begin{aligned} &= e^{-rT}E_{0,s}\left[\Phi_L(S(T))\right] - e^{-rT}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} E_{0,\frac{L^2}{s}}\left[\Phi_L(S(T))\right] \\ &= F(0,s;\Phi_L) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} F\left(0,\frac{L^2}{s};\Phi_L\right). \end{aligned}$$

■

Theorem 8.3.4 (Up-and-Out Contracts). For a fixed T(maturity) and $T-$ contingent claim $\Phi(S(T))$, the pricing function, denoted by $F^{LO}(t,s;\Phi)$, of the corresponding up-and-out contract Z^{LO}, is given, for $s < L$, by

$$\begin{aligned} F^{LO}(0,s;\Phi) &= e^{-rT}E_{0,s}\left[\Phi^L(S(T))\right] - e^{-rT}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} E_{0,\frac{L^2}{s}}\left[\Phi^L(S(T))\right] \\ &= F(0,s;\Phi^L) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} F\left(0,\frac{L^2}{s};\Phi^L\right). \end{aligned}$$

The problem of computing the price for a down-and-out claim and a up-and-out claim is reduced to the standard problem of computing the price of ordinary claims without a barrier. Refer to Björk(1998).

8.3.2 In-Out Bond

✎ **Down-and-Out Bond**
By Theorem(8.3.3), the down-and-out bond with barrier L is priced, for $s > L$, by the formula

$$\begin{aligned}
F_{LO}(t,s;BO) &= e^{-r(T-t)}E_{t,s}\left[BO_L(S(T))\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} E_{t,\frac{L^2}{s}}\left[BO_L(S(T))\right] \\
&= e^{-r(T-t)}E_{t,s}\left[H(S(T);L)\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} E_{t,\frac{L^2}{s}}\left[H(S(T);L)\right] \\
&= \mathbf{H}(t,s;L) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \mathbf{H}\left(t,\frac{L^2}{s};L\right),
\end{aligned}$$

where $\mathbf{H}$ is given in Eq(8.3.6). Hence

$$\begin{aligned}
&F_{LO}(t,s;BO) = F_{LO}\Big(t,s;H(*;L)\Big) \\
&= e^{-r(T-t)}N\left[\frac{\log\frac{s}{L} + \left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} N\left[\frac{\log\frac{L}{s} + \left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] \\
&= e^{-r(T-t)}N\left[d_2\left(\frac{s}{L}\right)\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} N\left[d_2\left(\frac{L}{s}\right)\right].
\end{aligned}$$

Note that

$$F_{LO}(t,s;BO) = F_{LO}(t,s;H(*;L)).$$

✎ **Down-and-In Bond**
Let $F_{LI}(t,s;BO)$ be the price of down-and-in Bond with barrier L. By the in-out parity, we have at maturity T

$$\text{Down-and-In} + \text{Down-and-Out} = 1.$$

Thus we have

$$F_{LI}(t,s;BO) = e^{-r(T-t)} - F_{LO}(t,s;BO).$$

Hence

$$\begin{aligned}
&F_{LI}(t,s;BO) \\
&= e^{-r(T-t)}N\left[\frac{\log\frac{L}{s} - \left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] + e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} N\left[\frac{\log\frac{L}{s} + \left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] \\
&= e^{-r(T-t)}N\left[-d_2\left(\frac{s}{L}\right)\right] + e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} N\left[d_2\left(\frac{L}{s}\right)\right].
\end{aligned}$$

✎ **Up-and-Out Bond**

By Theorem(8.3.4), the up-and-out bond with barrier L is priced, for $s < L$, by the formula

$$\begin{aligned}
F^{LO}(t,s;BO) &= e^{-r(T-t)}E_{t,s}\left[BO^L(S(T))\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} E_{t,\frac{L^2}{s}}\left[BO^L(S(T))\right] \\
&= e^{-r(T-t)}E_{t,s}\left[\widetilde{H}(S(T);L)\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} E_{t,\frac{L^2}{s}}\left[\widetilde{H}(S(T);L)\right] \\
&= \widetilde{\mathbf{H}}(t,s;L) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \widetilde{\mathbf{H}}\left(t,\frac{L^2}{s};L\right),
\end{aligned}$$

where $\widetilde{\mathbf{H}}$ is given in Eq(8.3.8). This gives us the following result:

$$\begin{aligned}
&F^{LO}(t,s;BO) = F^{LO}\Big(t,s;\widetilde{H}(*;L)\Big) \\
&= e^{-r(T-t)}N\left[\frac{\log\frac{L}{s} - \left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} N\left[\frac{\log\frac{s}{L} - \left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] \\
&= e^{-r(T-t)}N\left[-d_2\left(\frac{s}{L}\right)\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} N\left[-d_2\left(\frac{L}{s}\right)\right].
\end{aligned}$$

✎ **Up-and-In Bond**

Let $F^{LI}(t,s;BO)$ be the price of up-and-in Bond with barrier L. By the in-out parity, we have at maturity T

$$\text{Up-and-In} \quad + \quad \text{Up-and-Out} \quad = \quad 1.$$

Thus we have

$$F^{LI}(t,s;BO) \quad = \quad e^{-r(T-t)} - F^{LO}(t,s;BO).$$

Hence

$$\begin{aligned}
&F^{LI}(t,s;BO) \\
&= e^{-r(T-t)}N\left[\frac{\log\frac{s}{L} + \left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] + e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} N\left[\frac{\log\frac{s}{L} - \left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] \\
&= e^{-r(T-t)}N\left[d_2\left(\frac{s}{L}\right)\right] + e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} N\left[-d_2\left(\frac{L}{s}\right)\right].
\end{aligned}$$

Refer to Example(9.7.1).

8.3.3 In-Out Calls

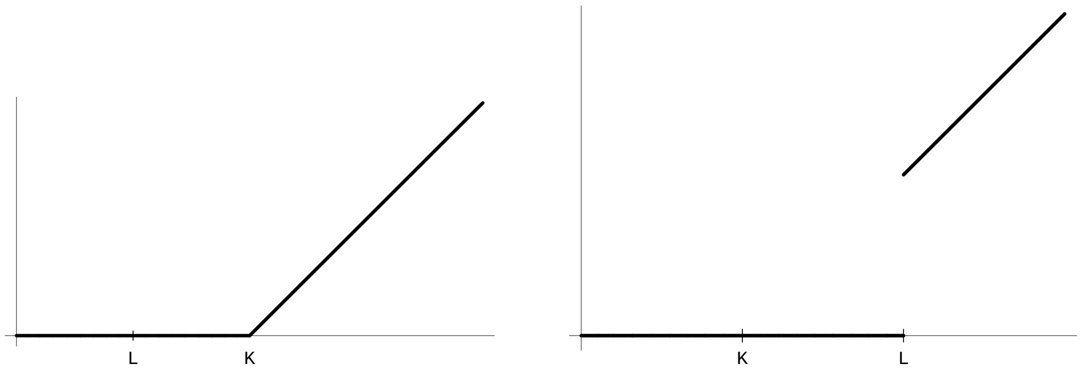


Figure 8.4: payoffs of barrier call option

✎ **Down-and-Out-Call**

1. $L < K$: See Figure(8.4(a)). Since

$$C_L(x;K) = C(x;K),$$

by Theorem(8.3.3), the down-and-out call with barrier L is priced, for $s > L$, by the formula

$$\begin{aligned} F_{LO}(t,s;C(*;K)) &= e^{-r(T-t)}E_{t,s}\left[C_L(S(T);K)\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} E_{t,\frac{L^2}{s}}\left[C_L(S(T);K)\right] \\ &= e^{-r(T-t)}E_{t,s}\left[C(S(T);K)\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} E_{t,\frac{L^2}{s}}\left[C(S(T);K)\right] \\ &= (\text{Black-Scholes formula}) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} (\text{Black-Scholes formula}) \\ &= \mathbf{C}(t,s;K) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \mathbf{C}\left(t,\frac{L^2}{s};K\right), \end{aligned}$$

where Black-Scholes formula, denoted by $\mathbf{C}$, is given in Eq(8.3.10). Hence

$$\begin{aligned} & L<K: \\ & F_{LO}\Big(t,s;C(*;K)\Big) \\ = \;& se^{-q(T-t)}N\left[\frac{\log\frac{s}{K}+\left(r-q+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] - Ke^{-r(T-t)}N\left[\frac{\log\frac{s}{K}+\left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] \\ & -\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left\{\frac{L^2}{s}e^{-q(T-t)}N\left[\frac{\log\frac{L^2}{sK}+\left(r-q+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] - Ke^{-r(T-t)}N\left[\frac{\log\frac{L^2}{sK}+\left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right]\right\}. \end{aligned}$$

2. $L > K$: See Figure(8.4(b)). Since

$$C_L(x;K) \;=\; C(x;L) + (L-K)H(x;L)$$

we have

$$\begin{aligned}
& F_{LO}\Big(t,s;C(*;K)\Big) \\
= \; & F_{LO}(t,s;C(*;L)) + (L-K)F_{LO}(t,s;H(*;L)) \\
= \; & \mathbf{C}(t,s;L) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \mathbf{C}\left(t,\frac{L^2}{s};L\right) + (L-K)\left\{\mathbf{H}(t,s;L) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \mathbf{H}\left(t,\frac{L^2}{s};L\right)\right\} \\
= \; & \mathbf{C}(t,s;L) + (L-K)\mathbf{H}(t,s;L) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \left\{\mathbf{C}\left(t,\frac{L^2}{s};L\right) + (L-K)\mathbf{H}\left(t,\frac{L^2}{s};L\right)\right\}.
\end{aligned}$$

Hence we have

$$\begin{aligned}
& L > K: \\
& F_{LO}(t,s;C(*;K)) \\
= \; & se^{-q(T-t)}N\left[\frac{\log\frac{s}{L} + \left(r-q+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] - Ke^{-r(T-t)}N\left[\frac{\log\frac{s}{L} + \left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] \\
& - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \left\{\frac{L^2}{s}e^{-q(T-t)}N\left[\frac{\log\frac{L}{s} + \left(r-q+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] - Ke^{-r(T-t)}N\left[\frac{\log\frac{L}{s} + \left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right]\right\}.
\end{aligned}$$

3. In general, for the both cases we can write

$$\begin{aligned}
C_L(x;K) &= C(x;K^*) + (K^*-K)H(x;L) \\
&= C(x;K^*) + (K^*-K)H(x;K^*),
\end{aligned}$$

where $K^* = \max(K,L)$. Thus we have

$$\begin{aligned}
& F_{LO}\Big(t,s;C(*;K)\Big) \\
= \; & F_{LO}(t,s;C(*;K^*)) + (K^*-K)F_{LO}(t,s;H(*;K^*)) \\
= \; & \mathbf{C}(t,s;K^*) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \mathbf{C}\left(t,\frac{L^2}{s};K^*\right) + (K^*-K)\left\{\mathbf{H}(t,s;K^*) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \mathbf{H}\left(t,\frac{L^2}{s};K^*\right)\right\} \\
= \; & \mathbf{C}(t,s;K^*) + (K^*-K)\mathbf{H}(t,s;K^*) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \left\{\mathbf{C}\left(t,\frac{L^2}{s};K^*\right) + (K^*-K)\mathbf{H}\left(t,\frac{L^2}{s};K^*\right)\right\}.
\end{aligned}$$

Hence the down-and-out-call price can be written as

$$\begin{aligned}
& F_{LO}\Big(t,s;C(*;K)\Big) \\
= \; & se^{-q(T-t)}N\left[\frac{\log\frac{s}{K^*} + \left(r-q+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] - Ke^{-r(T-t)}N\left[\frac{\log\frac{s}{K^*} + \left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] \\
& - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \left\{\frac{L^2}{s}e^{-q(T-t)}N\left[\frac{\log\frac{L^2}{sK^*} + \left(r-q+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] - Ke^{-r(T-t)}N\left[\frac{\log\frac{L^2}{sK^*} + \left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right]\right\} \\
= \; & se^{-q(T-t)}N\left[d_1\left(\frac{s}{K^*}\right)\right] - Ke^{-r(T-t)}N\left[d_2\left(\frac{s}{K^*}\right)\right] - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \left\{\frac{L^2}{s}e^{-q(T-t)}N\left[d_1\left(\frac{L^2}{sK^*}\right)\right] - Ke^{-r(T-t)}N\left[d_2\left(\frac{L^2}{sK^*}\right)\right]\right\}.
\end{aligned}$$

✎ **Down-and-In-Call**
By in-out parity, we have

$$F_{LI}\Big(t,s;C(*;K)\Big) \quad = \quad \mathbf{C}(t,s;K) - F_{LO}\Big(t,s;C(*;K)\Big).$$

Hence we have

$$\begin{aligned}
& F_{LI}\Big(t,s;C(*;K)\Big) \\
= \quad & se^{-q(T-t)}\left\{N\left[d_1\left(\frac{s}{K}\right)\right] - N\left[d_1\left(\frac{s}{K^*}\right)\right]\right\} - Ke^{-r(T-t)}\left\{N\left[d_2\left(\frac{s}{K}\right)\right] - N\left[d_2\left(\frac{s}{K^*}\right)\right]\right\} \\
& + \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left\{\frac{L^2}{s}e^{-q(T-t)}N\left[d_1\left(\frac{L^2}{sK^*}\right)\right] - Ke^{-r(T-t)}N\left[d_2\left(\frac{L^2}{sK^*}\right)\right]\right\}. \qquad (8.3.13)
\end{aligned}$$

✎ **Up-and-Out-Call**
By Example(8.3.1), we have

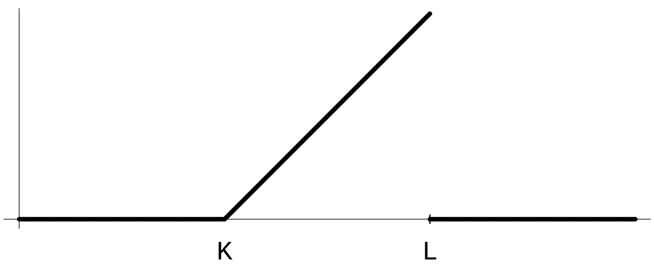


Figure 8.5: $C^L(x;L)$ with $K < L$

$$C^L(x;K) \quad = \quad C(x;K) - C(x;K^*) - (K^*-K)H(x;K^*), \quad \text{a.s.}$$

where $K^* = \max(K,L)$. By Theorem(8.3.4), we get

$$\begin{aligned}
& F^{LO}\Big(t,s;C(*;K)\Big) \\
= \quad & F\Big(t,s;C^L(*;K)\Big) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} F\left(t,\frac{L^2}{s};C^L(*;K)\right) \\
= \quad & F\Big(t,s;C(*;K)\Big) - F\Big(t,s;C(*;K^*)\Big) - (K^*-K)F\Big(t,s;H(*;K^*)\Big) \\
& \quad - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left\{F\left(t,\frac{L^2}{s};C(*;K)\right) - F\left(t,\frac{L^2}{s};C(*;K^*)\right) - (K^*-K)F\left(t,\frac{L^2}{s};H(*;K^*)\right)\right\} \\
= \quad & \mathbf{C}(t,s;K) - \mathbf{C}(t,s;K^*) - (K^*-K)\mathbf{H}(t,s;K^*) \\
& \quad - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left\{\mathbf{C}\left(t,\frac{L^2}{s};K\right) - \mathbf{C}\left(t,\frac{L^2}{s};K^*\right) - (K^*-K)\mathbf{H}\left(t,\frac{L^2}{s};K^*\right)\right\} \\
= \quad & se^{-q(T-t)}\left\{N\left[d_1\left(\frac{s}{K}\right)\right] - N\left[d_1\left(\frac{s}{K^*}\right)\right]\right\} - Ke^{-r(T-t)}\left\{N\left[d_2\left(\frac{s}{K}\right)\right] - N\left[d_2\left(\frac{s}{K^*}\right)\right]\right\}
\end{aligned}$$

$$
\begin{aligned}
& -\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left\{\frac{L^2}{s}e^{-q(T-t)}\left\{N\left[d_1\left(\frac{L^2}{sK}\right)\right]-N\left[d_1\left(\frac{L^2}{sK^*}\right)\right]\right\}\right. \\
& \qquad \left. -Ke^{-r(T-t)}\left\{N\left[d_2\left(\frac{L^2}{sK}\right)\right]-N\left[d_2\left(\frac{L^2}{sK^*}\right)\right]\right\}\right\}.
\end{aligned}
$$

Hence we have

$$
\begin{aligned}
& F^{LO}\Big(t,s;C(*;K)\Big) \\
= \; & se^{-q(T-t)}\left\{N\left[d_1\left(\frac{s}{K}\right)\right]-N\left[d_1\left(\frac{s}{K^*}\right)\right]\right\}-Ke^{-r(T-t)}\left\{N\left[d_2\left(\frac{s}{K}\right)\right]-N\left[d_2\left(\frac{s}{K^*}\right)\right]\right\} \\
& -\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left\{\frac{L^2}{s}e^{-q(T-t)}\left\{N\left[d_1\left(\frac{L^2}{sK}\right)\right]-N\left[d_1\left(\frac{L^2}{sK^*}\right)\right]\right\}-Ke^{-r(T-t)}\left\{N\left[d_2\left(\frac{L^2}{sK}\right)\right]-N\left[d_2\left(\frac{L^2}{sK^*}\right)\right]\right\}\right\}.
\end{aligned}
$$

✎ **Up-and-In-Call**

By the in-out parity, we have

$$
F^{LI}\Big(t,s;C(*;K)\Big) \quad = \quad \mathbf{C}(t,s;K)-F^{LO}\Big(t,s;C(*;K)\Big).
$$

Hence we have

$$
\begin{aligned}
& F^{LI}\Big(t,s;C(*;K)\Big) \\
= \; & se^{-q(T-t)}N\left[d_1\left(\frac{s}{K^*}\right)\right]-Ke^{-r(T-t)}N\left[d_2\left(\frac{s}{K^*}\right)\right] \\
& +\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left\{\frac{L^2}{s}e^{-q(T-t)}\left\{N\left[d_1\left(\frac{L^2}{sK}\right)\right]-N\left[d_1\left(\frac{L^2}{sK^*}\right)\right]\right\}-Ke^{-r(T-t)}\left\{N\left[d_2\left(\frac{L^2}{sK}\right)\right]-N\left[d_2\left(\frac{L^2}{sK^*}\right)\right]\right\}\right\}.
\end{aligned}
$$

8.3.4 In-Out Puts

✎ **Down-and-Out-Put** For the down-and-out-put, we have

$$P_L(x;K) \quad = \quad P(x;K) - P(x;K_*) - (K-K_*)\widetilde{H}(x;K_*), \qquad \text{a.s.}$$

See Figure(8.3). By Theorem(8.3.3) we have

$$F_{LO}\Big(t,s,P(*;K)\Big) \quad = \quad F\Big(t,s,P_L(*;K)\Big) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} F\left(t,\frac{L^2}{s},P_L(*;K)\right).$$

Now we can use Example(8.3.2). Hence we have

$$\begin{aligned}
&F_{LO}\Big(t,s,P(*;K)\Big)\\
=&Ke^{-r(T-t)}\left\{N\left[-d_2\left(\frac{s}{K}\right)\right]-N\left[-d_2\left(\frac{s}{K_*}\right)\right]\right\}-se^{-q(T-t)}\left\{N\left[-d_1\left(\frac{s}{K}\right)\right]-N\left[-d_1\left(\frac{s}{K_*}\right)\right]\right\}\\
&-\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left(Ke^{-r(T-t)}\left\{N\left[-d_2\left(\frac{L^2}{sK}\right)\right]-N\left[-d_2\left(\frac{L^2}{sK_*}\right)\right]\right\}-\frac{L^2}{s}e^{-q(T-t)}\left\{N\left[-d_1\left(\frac{L^2}{sK}\right)\right]-N\left[-d_1\left(\frac{L^2}{sK_*}\right)\right]\right\}\right)
\end{aligned}$$

For the alternative approach, we can use the following equations.

$$\begin{aligned}
P(x;K) &= KBO(x) - ST(x) + C(x;K),\\
ST_L(x) &= LH(x;L) + C(x;L),\\
P_L(x;K) &= KBO_L(x) - ST_L(x) + C_L(x;K)\\
&= KH(x;L) - LH(x;L) - C(x;L) + C_L(x;K)\\
&= (K-L)H(x;L) - C(x;L) + C_L(x;K).
\end{aligned}$$

Hence we have

$$\begin{aligned}
F_{LO}\Big(t,s,P(*;K)\Big) &= (K-L)F_{LO}\Big(t,s,H(*;L)\Big) - F_{LO}\Big(t,s,C(*;L)\Big) + F_{LO}\Big(t,s,C(*;K)\Big)\\
&= (\text{down-and-out-bond}) - (\text{down-and-out-call}) + (\text{down-and-out-call}).
\end{aligned}$$

Straightforward calculation gives us

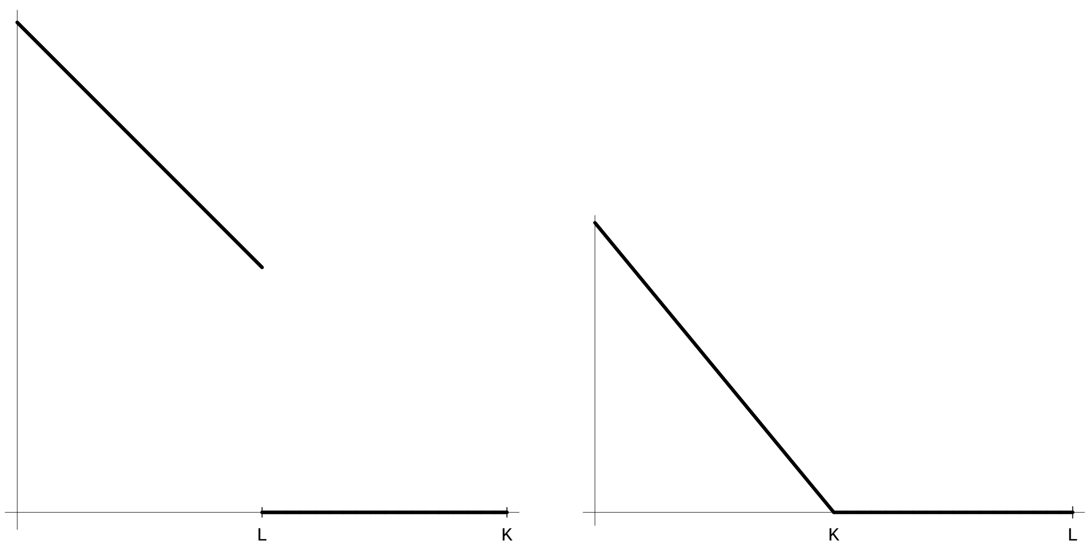
$$\begin{aligned}
&F_{LO}\Big(t,s,P(*;K)\Big)\\
=\;& Ke^{-r(T-t)}\left\{N\left[-d_2\left(\frac{s}{K^*}\right)\right]-N\left[-d_2\left(\frac{s}{L}\right)\right]\right\}-se^{-q(T-t)}\left\{N\left[-d_1\left(\frac{s}{K^*}\right)\right]-N\left[-d_1\left(\frac{s}{L}\right)\right]\right\}\\
&-\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left\{Ke^{-r(T-t)}\left\{N\left[-d_2\left(\frac{L}{s}\right)\right]-N\left[-d_2\left(\frac{L^2}{sK^*}\right)\right]\right\}-\frac{L^2}{s}e^{-q(T-t)}\left\{N\left[-d_1\left(\frac{L}{s}\right)\right]-N\left[-d_1\left(\frac{L^2}{sK^*}\right)\right]\right\}\right\}\\
=\;& Ke^{-r(T-t)}\left\{N\left[-d_2\left(\frac{s}{K}\right)\right]-N\left[-d_2\left(\frac{s}{K_*}\right)\right]\right\}-se^{-q(T-t)}\left\{N\left[-d_1\left(\frac{s}{K}\right)\right]-N\left[-d_1\left(\frac{s}{K_*}\right)\right]\right\}\\
&-\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left\{Ke^{-r(T-t)}\left\{N\left[-d_2\left(\frac{L^2}{sK}\right)\right]-N\left[-d_2\left(\frac{L^2}{sK_*}\right)\right]\right\}-\frac{L^2}{s}e^{-q(T-t)}\left\{N\left[-d_1\left(\frac{L^2}{sK}\right)\right]-N\left[-d_1\left(\frac{L^2}{sK_*}\right)\right]\right\}\right\}.
\end{aligned}$$

✎ **Down-and-In Put**
By the in-out parity, we have

$$F_{LI}\Big(t,s;P(*;K)\Big) \quad = \quad \mathbf{P}(t,s;K) - F_{LO}\Big(t,s;P(*;K)\Big).$$

Hence we have

$$
\begin{aligned}
&F_{LI}\Big(t, s, P(*;K)\Big)\\
=&Ke^{-r(T-t)}N\left[-d_2\left(\frac{s}{K_*}\right)\right] - se^{-q(T-t)}N\left[-d_1\left(\frac{s}{K_*}\right)\right]\\
&+\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left\{Ke^{-r(T-t)}\left\{N\left[-d_2\left(\frac{L^2}{sK}\right)\right] - N\left[-d_2\left(\frac{L^2}{sK_*}\right)\right]\right\} - \frac{L^2}{s}e^{-q(T-t)}\left\{N\left[-d_1\left(\frac{L^2}{sK}\right)\right] - N\left[-d_1\left(\frac{L^2}{sK_*}\right)\right]\right\}\right\}
\end{aligned}
$$



(a) $L < K$ (b) $L > K$

Figure 8.6: $P^L(x;K)$.

✎ **Up-and-Out-Put** For the up-and-out-put, we have

$$P^L(x;K) \quad = \quad P(x;K_*) + (K-K_*)\widetilde{H}(x;K_*).$$

Thus we get

$$
\begin{aligned}
& F^{LO}\Big(t, s; P(*;K)\Big)\\
= \;& F^{LO}\Big(t, s; P(*;K_*)\Big) + (K-K_*)F^{LO}\Big(t, s; \widetilde{H}(*;K)\Big)\\
= \;& F\Big(t, s; P(*;K_*)\Big) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} F\left(t, \frac{L^2}{s}; P(*;K_*)\right)\\
& +(K-K_*)\left\{F\Big(t, s; \widetilde{H}(*,K_*)\Big) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} F\left(t, \frac{L^2}{s}, \widetilde{H}(*;K_*)\right)\right\}\\
= \;& \mathbf{P}(t, s; K_*) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\mathbf{P}\left(t, \frac{L^2}{s}; K_*\right) + (K-K_*)\left\{\widetilde{\mathbf{H}}(t, s; K_*) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\widetilde{\mathbf{H}}\left(t, \frac{L^2}{s}; K_*\right)\right\}
\end{aligned}
$$

$$\begin{aligned}
&= K_* e^{-r(T-t)} N\left[-d_2\left(\frac{s}{K_*}\right)\right] - s e^{-q(T-t)} N\left[-d_1\left(\frac{s}{K_*}\right)\right] \\
&\quad -\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \left\{K_* e^{-r(T-t)} N\left[-d_2\left(\frac{L^2}{sK_*}\right)\right] - \frac{L^2}{s} e^{-q(T-t)} N\left[-d_1\left(\frac{L^2}{sK_*}\right)\right]\right\} \\
&\quad +(K-K_*)\left\{e^{-r(T-t)} N\left[-d_2\left(\frac{s}{K_*}\right)\right] - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} e^{-r(T-t)} N\left[-d_2\left(\frac{L^2}{sK_*}\right)\right]\right\} \\
&= e^{-r(T-t)} K N\left[-d_2\left(\frac{s}{K_*}\right)\right] - s e^{-q(T-t)} N\left[-d_1\left(\frac{s}{K_*}\right)\right] \\
&\quad -\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \left\{K e^{-r(T-t)} N\left[-d_2\left(\frac{L^2}{sK_*}\right)\right] - \frac{L^2}{s} e^{-q(T-t)} N\left[-d_1\left(\frac{L^2}{sK_*}\right)\right]\right\}.
\end{aligned}$$

Hence we have

$$\begin{aligned}
&F^{LO}\Big(t, s; P(*;K)\Big) \\
&= e^{-r(T-t)} K N\left[-d_2\left(\frac{s}{K_*}\right)\right] - s e^{-q(T-t)} N\left[-d_1\left(\frac{s}{K_*}\right)\right] \\
&\qquad -\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \left\{K e^{-r(T-t)} N\left[-d_2\left(\frac{L^2}{sK_*}\right)\right] - \frac{L^2}{s} e^{-q(T-t)} N\left[-d_1\left(\frac{L^2}{sK_*}\right)\right]\right\}.
\end{aligned}$$

✎ **Up-and-In-Put**

By the in-out parity, we have

$$F^{LI}\Big(t, s; P(*;K)\Big) = \mathbf{P}(t, s; K) - F^{LO}\Big(t, s; P(*;K)\Big).$$

Hence we have

$$\begin{aligned}
&F^{LI}\Big(t, s, P(*;K)\Big) \\
&= K e^{-r(T-t)} \left\{N\left[-d_2\left(\frac{s}{K}\right)\right] - N\left[-d_2\left(\frac{s}{K_*}\right)\right]\right\} - s e^{-q(T-t)} \left\{N\left[-d_1\left(\frac{s}{K}\right)\right] - N\left[-d_1\left(\frac{s}{K_*}\right)\right]\right\} \\
&\qquad +\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \left\{K e^{-r(T-t)} N\left[-d_2\left(\frac{L^2}{sK_*}\right)\right] - \frac{L^2}{s} e^{-q(T-t)} N\left[-d_1\left(\frac{L^2}{sK_*}\right)\right]\right\}.
\end{aligned}$$

8.3.5 In-Out Asset

✎ **Down-and-Out-Asset**

We will use the put-call parity. i.e.

$$\mathbf{ST}(x) \quad = \quad C(x;L) - P(x;L) + LBO(x), \qquad \text{for any fixed } L > 0.$$

Hence the down-and-out asset price is given by

$$\begin{aligned}
& F_{LO}(t,s,\mathbf{ST}) \\
= \; & F_{LO}\Big(t,s,C(*;L)\Big) - F_{LO}\Big(t,s,P(*;L)\Big) + LF_{LO}(t,s;BO) \\
= \; & F_{LO}\Big(t,s,C(*;L)\Big) + LF_{LO}(t,s;BO) \\
= \; & se^{-q(T-t)}N\left[d_1\left(\frac{s}{L}\right)\right] - \cancel{Le^{-r(T-t)}N\left[d_2\left(\frac{s}{L}\right)\right]} - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left\{\frac{L^2}{s}e^{-q(T-t)}N\left[d_1\left(\frac{L}{s}\right)\right] - \cancel{Le^{-r(T-t)}N\left[d_2\left(\frac{L}{s}\right)\right]}\right\} \\
& \cancel{+Le^{-r(T-t)}N\left[d_2\left(\frac{s}{L}\right)\right]} - \cancel{Le^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}N\left[d_2\left(\frac{L}{s}\right)\right]} \\
= \; & se^{-q(T-t)}N\left[d_1\left(\frac{s}{L}\right)\right] - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\frac{L^2}{s}e^{-q(T-t)}N\left[d_1\left(\frac{L}{s}\right)\right].
\end{aligned}$$

Hence we have

$$\boxed{F_{LO}(t,s,\mathbf{ST}) = se^{-q(T-t)}N\left[d_1\left(\frac{s}{L}\right)\right] - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\frac{L^2}{s}e^{-q(T-t)}N\left[d_1\left(\frac{L}{s}\right)\right].}$$

✎ **Down-and-In-Asset**

By the in-out parity, we have

$$F_{LI}(t,s,\mathbf{ST}) \quad = \quad se^{-q(T-t)} - F_{LO}(t,s,ST).$$

Hence we have

$$\boxed{F_{LI}(t,s,\mathbf{ST}) = se^{-q(T-t)}N\left[-d_1\left(\frac{s}{L}\right)\right] + \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\frac{L^2}{s}e^{-q(T-t)}N\left[d_1\left(\frac{L}{s}\right)\right].}$$

Note that since asset can be regarded as a call option with exercise price 0. Hence this formula can be obtained from (8.3.13) by putting $K = 0$.

✎ **Up-and-Out-Asset**

By the put-call parity, we have

$$\begin{aligned}
& F^{LO}(t,s,\mathbf{ST}) \\
= \; & F^{LO}\Big(t,s,C(*;L)\Big) - F^{LO}\Big(t,s,P(*;L)\Big) + LF^{LO}(t,s;BO) \\
= \; & -F^{LO}\Big(t,s,P(*;L)\Big) + LF^{LO}(t,s;BO)
\end{aligned}$$

$$
\begin{aligned}
&= \cancel{-e^{-r(T-t)}LN\left[-d_2\left(\frac{s}{L}\right)\right]} + se^{-q(T-t)}N\left[-d_1\left(\frac{s}{L}\right)\right] \\
&\quad + \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left\{\cancel{Le^{-r(T-t)}N\left[-d_2\left(\frac{L}{s}\right)\right]} - \frac{L^2}{s}e^{-q(T-t)}N\left[-d_1\left(\frac{L}{s}\right)\right]\right\} \\
&\quad \cancel{+e^{-r(T-t)}LN\left[-d_2\left(\frac{s}{L}\right)\right]} - \cancel{e^{-r(T-t)}L\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}N\left[-d_2\left(\frac{L}{s}\right)\right]} \\
&= se^{-q(T-t)}N\left[-d_1\left(\frac{s}{L}\right)\right] - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\frac{L^2}{s}e^{-q(T-t)}N\left[-d_1\left(\frac{L}{s}\right)\right].
\end{aligned}
$$

Hence we have

$$
\boxed{F^{LO}(t,s,\mathbf{ST}) = se^{-q(T-t)}N\left[-d_1\left(\frac{s}{L}\right)\right] - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\frac{L^2}{s}e^{-q(T-t)}N\left[-d_1\left(\frac{L}{s}\right)\right].}
$$

✎ **Up-and-In-Asset**

By the in-out parity, we have

$$
F^{LI}(t,s,\mathbf{ST}) = se^{-q(T-t)} - F^{LO}(t,s,ST).
$$

Hence we have

$$
\boxed{F^{LI}(t,s,\mathbf{ST}) = se^{-q(T-t)}N\left[d_1\left(\frac{s}{L}\right)\right] + \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\frac{L^2}{s}e^{-q(T-t)}N\left[-d_1\left(\frac{L}{s}\right)\right].}
$$

8.3.6 In-Out Cash-Or-Nothing-Call

Binary options are options with discontinuous payoffs. A simple example of binary option is a *cash-or-nothing call*(digital call, binary call). It's maturity payoff is given by

$$\begin{cases} 1, & S_T \geq K, \\ 0, & S_T < K. \end{cases}$$

✎ **Down-and-Out-*Cash-Or-Nothing-Call***

$$\begin{aligned} F_{LO}\Big(t,s;\mathbf{H}(*;K^*)\Big) &= e^{-r(T-t)}E_{t,s}\left[H(S(T);K^*)\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} E_{t,\frac{L^2}{s}}\left[H(S(T);K^*)\right] \\ &= \mathbf{H}(t,s;K^*) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \mathbf{H}\left(t,\frac{L^2}{s};K^*\right), \end{aligned}$$

where $\mathbf{H}$ is given in Eq(8.3.6). Hence we have

$$\begin{aligned} &F_{LO}\Big(t,s;\mathbf{H}(*;K)\Big) \\ &= e^{-r(T-t)}N\left[\frac{\log\frac{s}{K^*}+\left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} N\left[\frac{\log\frac{L^2}{sK^*}+\left(r-q-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] \\ &= e^{-r(T-t)}N\left[d_2\left(\frac{s}{K^*}\right)\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} N\left[d_2\left(\frac{L^2}{sK^*}\right)\right]. \end{aligned}$$

✎ **Down-and-In-*Cash-Or-Nothing-Call***

By the in-out parity, we have

$$\begin{aligned} F_{LI}\Big(t,s,\mathbf{H}(*;K)\Big) &= \text{(Digital Call)} - \text{(Down-and-Out-Cash-Or-Nothing-Call)} \\ &= e^{-r(T-t)}N\left(d_2\left(\frac{s}{K}\right)\right) - F_{LO}\Big(t,s;\mathbf{H}(*;K)\Big). \end{aligned}$$

Hence we have

$$\begin{aligned} &F_{LI}\Big(t,s,\mathbf{H}(*;K)\Big) \\ &= e^{-r(T-t)}N\left(d_2\left(\frac{s}{K}\right)\right) - e^{-r(T-t)}N\left[d_2\left(\frac{s}{K^*}\right)\right] + e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} N\left[d_2\left(\frac{L^2}{sK^*}\right)\right]. \end{aligned}$$

✎ **Up-and-Out-*Cash-Or-Nothing-Call***

$$F^{LO}\Big(t,s;\mathbf{H}(*;K)-\mathbf{H}(*;K^*)\Big) = F^{LO}\Big(t,s;\mathbf{H}(*;K)\Big) - F^{LO}\Big(t,s;\mathbf{H}(*;K^*)\Big).$$

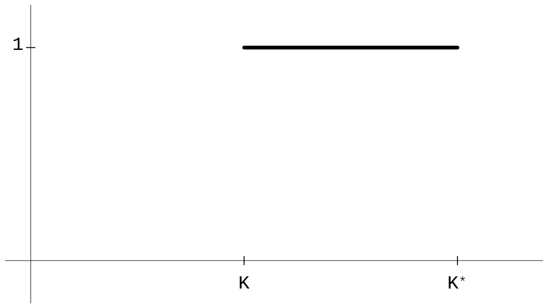


Figure 8.7: Up & out cash or nothing call

Hence we have

$$
\begin{aligned}
& F^{LO}\Big(t, s; \mathbf{H}(*; K) - \mathbf{H}(*; K^*)\Big) \\
= \;& e^{-r(T-t)}\left\{N\left(d_2\left(\frac{s}{K}\right)\right) - N\left[d_2\left(\frac{s}{K^*}\right)\right]\right\} - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left\{N\left[d_2\left(\frac{L^2}{sK}\right)\right] - N\left[d_2\left(\frac{L^2}{sK^*}\right)\right]\right\}.
\end{aligned}
$$

✎ **Up-and-In-*Cash-Or-Nothing-Call***

By the in-out parity, we have

$$
\begin{aligned}
& F^{LI}\Big(t, s; \mathbf{H}(*; K) - \mathbf{H}(*; K^*)\Big) \\
= \;& e^{-r(T-t)}N\left[d_2\left(\frac{s}{K^*}\right)\right] + e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\left\{N\left[d_2\left(\frac{L^2}{sK}\right)\right] - N\left[d_2\left(\frac{L^2}{sK^*}\right)\right]\right\}.
\end{aligned}
$$

8.3.7 In-Out Cash-Or-Nothing-Put

✎ **Down-and-Out-*Cash-Or-Nothing-Put***

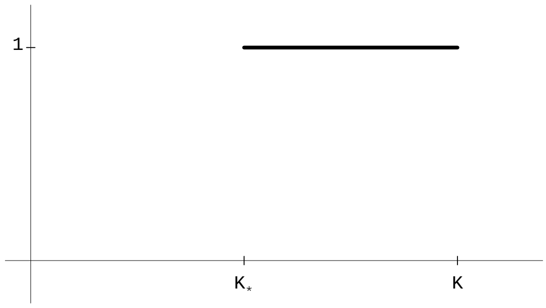


Figure 8.8: Down & out cash or nothing put

$$F_{LO}\Big(t,s;\mathbf{H}(*;K_*)-\mathbf{H}(*;K)\Big) \quad = \quad F_{LO}\Big(t,s;\mathbf{H}(*;K_*)\Big)-F_{LO}\Big(t,s;\mathbf{H}(*;K)\Big).$$

Hence we have

$$\begin{aligned}&F_{LO}\Big(t,s;\mathbf{H}(*;K_*)-\mathbf{H}(*;K)\Big)\\ =\;& e^{-r(T-t)}\Bigg\{N\left[d_2\left(\frac{s}{K_*}\right)\right]-N\left[d_2\left(\frac{s}{K}\right)\right]\Bigg\}-e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\Bigg\{N\left[d_2\left(\frac{L^2}{sK_*}\right)\right]-N\left[d_2\left(\frac{L^2}{sK}\right)\right]\Bigg\}.\end{aligned}$$

✎ **Down-and-In-*Cash-Or-Nothing-Put***

By the in-out parity, we have

$$F_{LI}\Big(t,s;\mathbf{H}(*;K_*)-\mathbf{H}(*;K)\Big) \quad = \quad e^{-r(T-t)}N\left(-d_2\left(\frac{s}{K}\right)\right)-F_{LO}\Big(t,s;\mathbf{H}(*;K_*)-\mathbf{H}(*;K)\Big).$$

Since

$$N\left[d_2\left(\frac{s}{K_*}\right)\right]-N\left[d_2\left(\frac{s}{K}\right)\right] \quad = \quad N\left[-d_2\left(\frac{s}{K}\right)\right]-N\left[-d_2\left(\frac{s}{K_*}\right)\right]$$

we have

$$\begin{aligned}&F_{LI}\Big(t,s;\mathbf{H}(*;K_*)-\mathbf{H}(*;K)\Big)\\ =\;& e^{-r(T-t)}N\left[-d_2\left(\frac{s}{K_*}\right)\right]+e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\Bigg\{N\left[d_2\left(\frac{L^2}{sK_*}\right)\right]-N\left[d_2\left(\frac{L^2}{sK}\right)\right]\Bigg\}.\end{aligned}$$

✎ **Up-and-Out-*Cash-Or-Nothing-Put***

The value of Up & out *cash-or-nothing-put* option is equal to

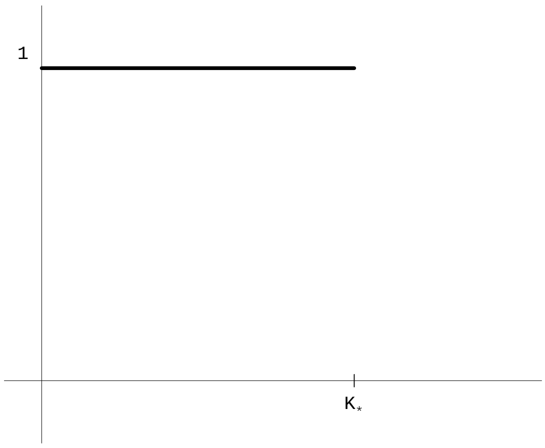


Figure 8.9: Up & out cash or nothing put

$$F^{LO}\Big(t, s; \widetilde{\mathbf{H}}(*; K_*)\Big).$$

Hence we have

$$\begin{aligned} & F^{LO}\Big(t, s; \widetilde{\mathbf{H}}(*; K_*)\Big) \\ = \;& e^{-r(T-t)} N\left[-d_2\left(\frac{s}{K_*}\right)\right] - e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} N\left[-d_2\left(\frac{L^2}{sK_*}\right)\right]. \end{aligned}$$

✎ **Up-and-In-*Cash-Or-Nothing-Put***

By the in-out parity we have

$$\begin{aligned} & F^{LI}\Big(t, s; \widetilde{\mathbf{H}}(*; K_*)\Big) \\ = \;& e^{-r(T-t)}\left\{N\left[-d_2\left(\frac{s}{K}\right)\right] - N\left[-d_2\left(\frac{s}{K_*}\right)\right]\right\} + e^{-r(T-t)}\left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} N\left[-d_2\left(\frac{L^2}{sK_*}\right)\right]. \end{aligned}$$

8.3.8 In-Out Asset-Or-Nothing-Call

✎ **Down-and-Out-*Asset-Or-Nothing-Call***
Note that for $K > 0$,

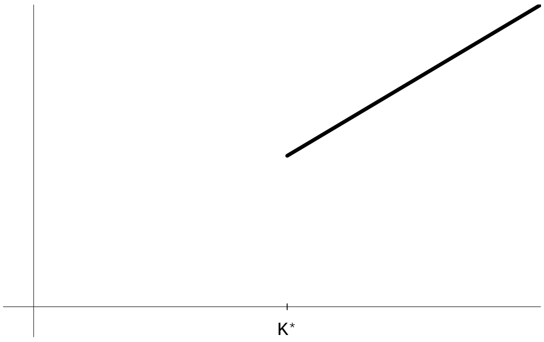


Figure 8.10: Down & out asset or nothing call

$$e^{-r(T-t)}E\Big[S_T\mathbf{1}_{\{S_T\geq K\}}|\mathscr{F}_t\Big] \quad = \quad S_t e^{-q(T-t)}N\left(d_1\left(\frac{S_t}{K}\right)\right).$$

Hence we have

$$\begin{aligned} & F_{LO}\Big(t,s,\mathbf{ST}(*)\mathbf{H}(*;K)\Big) \\ = \quad & se^{-q(T-t)}N\left(d_1\left(\frac{s}{K^*}\right)\right) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\frac{L^2}{s}e^{-q(T-t)}N\left[d_1\left(\frac{L^2}{sK^*}\right)\right]. \end{aligned}$$

✎ **Down-and-In-*Asset-Or-Nothing-Call***
By the in-out parity we have

$$\begin{aligned} & F_{LI}\Big(t,s,\mathbf{ST}(*)\mathbf{H}(*;K)\Big) \\ = \quad & se^{-q(T-t)}\left\{N\left(d_1\left(\frac{s}{K}\right)\right) - N\left(d_1\left(\frac{s}{K^*}\right)\right)\right\} + \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\frac{L^2}{s}e^{-q(T-t)}N\left[d_1\left(\frac{L^2}{sK^*}\right)\right]. \end{aligned}$$

✎ **Up-and-Out-*Asset-Or-Nothing-Call***

Refer to Figure(8.11).

$$\begin{aligned} & F^{LO}\Big(t,s,\mathbf{ST}(*)\Big(\mathbf{H}(*;K) - \mathbf{H}(*;K^*)\Big)\Big) \\ = \quad & se^{-q(T-t)}\left\{N\left(d_1\left(\frac{s}{K}\right)\right) - N\left(d_1\left(\frac{s}{K^*}\right)\right)\right\} - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\frac{L^2}{s}e^{-q(T-t)}\left\{N\left[d_1\left(\frac{L^2}{sK}\right)\right] - N\left[d_1\left(\frac{L^2}{sK^*}\right)\right]\right\}. \end{aligned}$$

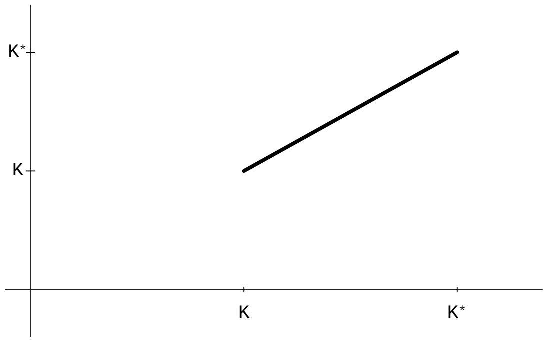


Figure 8.11: Up & out asset or nothing call

✎ **Up-and-In-*Asset-Or-Nothing-Call***
By the in-out parity we have

$$
\begin{aligned}
& F^{LI}\Big(t, s, \mathbf{ST}(*)\Big(\mathbf{H}(*;K) - \mathbf{H}(*;K^*)\Big)\Big) \\
= \quad & se^{-q(T-t)} N\left(d_1\left(\frac{s}{K^*}\right)\right) + \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \frac{L^2}{s} e^{-q(T-t)} \left\{ N\left[d_1\left(\frac{L^2}{sK}\right)\right] - N\left[d_1\left(\frac{L^2}{sK^*}\right)\right]\right\}.
\end{aligned}
$$

8.3.9 In-Out Asset-Or-Nothing-Put

✎ **Down-and-Out-*Asset-Or-Nothing-Put***

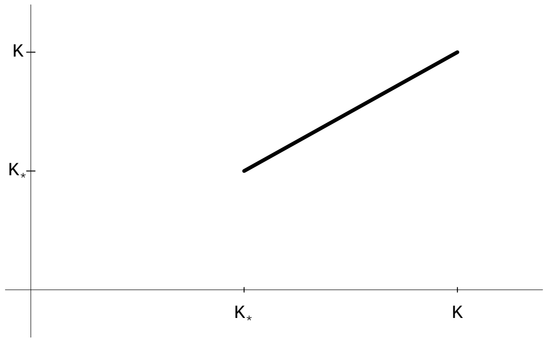


Figure 8.12: Down & out asset or nothing put

Refer to Figure(8.12).

$$\begin{aligned}
& F_{LO}\Big(t, s, \mathbf{ST}(*)\Big(\mathbf{H}(*;K_*) - \mathbf{H}(*;K)\Big)\Big) \\
= \;& s e^{-q(T-t)}\left\{N\left(d_1\left(\frac{s}{K_*}\right)\right) - N\left(d_1\left(\frac{s}{K}\right)\right)\right\} - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \frac{L^2}{s} e^{-q(T-t)}\left\{N\left[d_1\left(\frac{L^2}{sK_*}\right)\right] - N\left[d_1\left(\frac{L^2}{sK}\right)\right]\right\}.
\end{aligned}$$

✎ **Down-and-In-*Asset-Or-Nothing-Put***
By the in-out parity we have

$$\begin{aligned}
& F_{LI}\Big(t, s, \mathbf{ST}(*)\Big(\mathbf{H}(*;K_*) - \mathbf{H}(*;K)\Big)\Big) \\
= \;& s e^{q(T-t)} N\left(-d_1\left(\frac{s}{K}\right)\right) - F_{LO}\Big(t, s, \mathbf{ST}(*)\Big(\mathbf{H}(*;K_*) - \mathbf{H}(*;K)\Big)\Big)
\end{aligned}$$

Thus we have

$$\begin{aligned}
& F_{LI}\Big(t, s, \mathbf{ST}(*)\Big(\mathbf{H}(*;K_*) - \mathbf{H}(*;K)\Big)\Big) \\
= \;& s e^{-q(T-t)} N\left(-d_1\left(\frac{s}{K_*}\right)\right) + \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}} \frac{L^2}{s} e^{-q(T-t)}\left\{N\left[d_1\left(\frac{L^2}{sK_*}\right)\right] - N\left[d_1\left(\frac{L^2}{sK}\right)\right]\right\}.
\end{aligned}$$

✎ **Up-and-Out-*Asset-Or-Nothing-Put***
Refer to Figure(8.13). Note that for $K > 0$,

$$e^{-r(T-t)} E\Big[S_T \mathbf{1}_{\{S_T \le K\}} | \mathscr{F}_t\Big] \quad = \quad S_t e^{-q(T-t)} N\left(-d_1\left(\frac{S_t}{K}\right)\right).$$

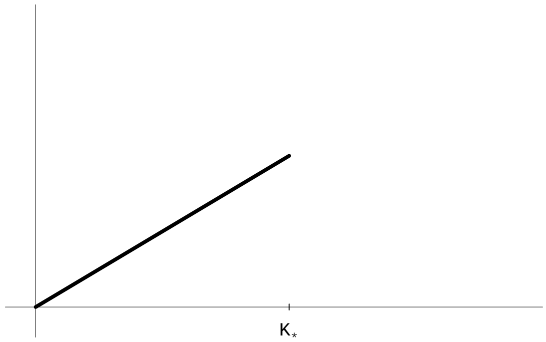


Figure 8.13: Up & out asset or nothing put

Hence we have

$$
\begin{aligned}
& F^{LO}\Big(t, s, \mathbf{ST}(*)\widetilde{\mathbf{H}}(*;K)\Big) \\
= \quad & se^{-q(T-t)}N\left(-d_1\left(\frac{s}{K_*}\right)\right) - \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\frac{L^2}{s}e^{-q(T-t)}N\left[-d_1\left(\frac{L^2}{sK_*}\right)\right].
\end{aligned}
$$

✎ **Up-and-In-*Asset-Or-Nothing-Put***

By the in-out parity we have

$$
\begin{aligned}
& F^{LI}\Big(t, s, \mathbf{ST}(*)\widetilde{\mathbf{H}}(*;K)\Big) \\
= \quad & se^{-q(T-t)}\left\{N\left(-d_1\left(\frac{s}{K}\right)\right) - N\left(-d_1\left(\frac{s}{K_*}\right)\right)\right\} + \left(\frac{L}{s}\right)^{\frac{2\mu}{\sigma^2}}\frac{L^2}{s}e^{-q(T-t)}N\left[-d_1\left(\frac{L^2}{sK_*}\right)\right].
\end{aligned}
$$

8.4 Double Barrier Options

8.4.1 Double Barrier Knock Out Call and Put

We assume that asset process S_t is given by

$$S_t = S_0 \exp(\mu t + \sigma W_t),$$

where $\mu = r - q - \frac{1}{2}\sigma^2$ and W_t is a Brownian motion under risk-neutral measure.

The double barrier options are expressed as infinite series of weighted normal distribution functions. However, numerical studies show that the convergence of the formula is quite rapid. Ikeda and Kunitomo(1992) suggest that it suffices to calculate the leading two or three terms in most cases.

Theorem 8.4.1 (Double Barrier Knock Out Call). The price of a knock out call with barriers H, $L(L < S_0 < H)$, and with strike $K(< H)$, denoted by $\Pi_c(0, S_0, H, L)$, is given by

$$\begin{aligned}
&\Pi_c(0, S_0, H, L)\\
= &\sum_{n=-\infty}^{\infty}\left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}}\left[S_0 e^{-qT}\left(\frac{L}{H}\right)^{2n}\left\{N\left(\frac{\log\frac{H^{2n+1}}{S_0 L^{2n}} - (\mu+\sigma^2)T}{\sigma\sqrt{T}}\right) - N\left(\frac{\log\frac{\overline{K}H^{2n}}{S_0 L^{2n}} - (\mu+\sigma^2)T}{\sigma\sqrt{T}}\right)\right\}\right.\\
&\quad -Ke^{-rT}\left\{N\left(\frac{\log\frac{H^{2n+1}}{S_0 L^{2n}} - \mu T}{\sigma\sqrt{T}}\right) - N\left(\frac{\log\frac{\overline{K}H^{2n}}{S_0 L^{2n}} - \mu T}{\sigma\sqrt{T}}\right)\right\}\\
&\quad -S_0 e^{-qT}\left(\frac{H}{S_0}\right)^{\frac{2\mu}{\sigma^2}}\left(\frac{L}{H}\right)^{2n}\frac{H^2}{S_0^2}\left\{N\left(\frac{\log\frac{S_0 H^{2n-1}}{L^{2n}} - (\mu+\sigma^2)T}{\sigma\sqrt{T}}\right) - N\left(\frac{\log\frac{\overline{K}S_0 H^{2n-2}}{L^{2n}} - (\mu+\sigma^2)T}{\sigma\sqrt{T}}\right)\right\}\\
&\quad \left.+Ke^{-rT}\left(\frac{H}{S_0}\right)^{\frac{2\mu}{\sigma^2}}\left\{N\left(\frac{\log\frac{S_0 H^{2n-1}}{L^{2n}} - \mu T}{\sigma\sqrt{T}}\right) - N\left(\frac{\log\frac{\overline{K}S_0 H^{2n-2}}{L^{2n}} - \mu T}{\sigma\sqrt{T}}\right)\right\}\right],
\end{aligned}$$

where $\overline{K} = \max(L, K)$.

Proof. Let $X_t = \log\frac{S_t}{S_0}$

$$m_S(t) = \inf_{0\le s\le t} S_s, \quad M_S(t) = \sup_{0\le s\le t} S_s$$

and

$$m_X(t) = \inf_{0\le s\le t} X_s, \quad M_X(t) = \sup_{0\le s\le t} X_s.$$

We have

$$\begin{aligned}
\Pi_c(0, S_0, H, L) &= e^{-rT}E\left[(S_T - K)^+ \middle| L < m_S(T) < M_S(T) < H\right]\\
&= e^{-rT}E\left[(S_0 e^{X_T} - K)^+ \middle| \log\frac{L}{S_0} < m_X(T) < M_X(T) < \log\frac{H}{S_0}\right]\\
&= e^{-rT}\int_{\overline{k}}^{b}(S_0 e^x - K)^+ h_{ab}(x)dx,
\end{aligned}$$

where $\overline{k} = \max\left(\log\frac{L}{S_0}, \log\frac{K}{S_0}\right)$, $a = \log\frac{L}{S_0}$, $b = \log\frac{H}{S_0}$ and $h_{ab}(x)$ is given in Theorem (6.4.2), i.e.

$$h_{a,b} = \frac{1}{\sigma\sqrt{2\pi T}} \sum_{n=-\infty}^{\infty} \left[\exp\left(\frac{-2\mu n(b-a)}{\sigma^2}\right) \exp\left(-\frac{\big(x+2n(b-a)-\mu T\big)^2}{2\sigma^2 T}\right) - \exp\left(\frac{-2\mu\big(n(b-a)-b\big)}{\sigma^2}\right) \exp\left(-\frac{\big(x-2b+2n(b-a)-\mu T\big)^2}{2\sigma^2 T}\right) \right].$$

Thus we could have

$$\Pi_c(0, S_0, H, L) = \frac{e^{-rT}}{\sigma\sqrt{2\pi T}} \sum_{n=-\infty}^{\infty} \left[\exp\left(\frac{-2\mu n(b-a)}{\sigma^2}\right) \int_{\overline{k}}^{b} (S_0 e^x - K) \exp\left(-\frac{\big(x+2n(b-a)-\mu T\big)^2}{2\sigma^2 T}\right) dx - \exp\left(\frac{-2\mu\big(n(b-a)-b\big)}{\sigma^2}\right) \int_{\overline{k}}^{b} (S_0 e^x - K) \exp\left(-\frac{\big(x-2b+2n(b-a)-\mu T\big)^2}{2\sigma^2 T}\right) dx \right].$$

It follows from (7.1.1) that

1.

$$\begin{aligned}
&\frac{S_0 e^{-rT}}{\sigma\sqrt{2\pi T}} \exp\left(\frac{-2\mu n(b-a)}{\sigma^2}\right) \int_{\overline{k}}^{b} e^x \exp\left(-\frac{\big(x+2n(b-a)-\mu T\big)^2}{2\sigma^2 T}\right) dx \\
= \;& S_0 e^{-rT} \left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}} \exp\left(\mu T - 2n(b-a) + \frac{1}{2}\sigma^2 T\right) \\
& \left\{ N\left(\frac{b-\mu T+2n(b-a)-\sigma^2 T}{\sigma\sqrt{T}}\right) - N\left(\frac{\overline{k}-\mu T+2n(b-a)-\sigma^2 T}{\sigma\sqrt{T}}\right) \right\} \\
= \;& S_0 e^{-qT} \left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}} \left(\frac{L}{H}\right)^{2n} \left\{ N\left(\frac{\log\frac{H^{2n+1}}{S_0 L^{2n}} - (\mu+\sigma^2)T}{\sigma\sqrt{T}}\right) - N\left(\frac{\log\frac{\overline{K} H^{2n}}{S_0 L^{2n}} - (\mu+\sigma^2)T}{\sigma\sqrt{T}}\right) \right\},
\end{aligned}$$

where $\overline{K} = \max(L, K)$.

2.

$$\begin{aligned}
&\frac{K e^{-rT}}{\sigma\sqrt{2\pi T}} \exp\left(\frac{-2\mu n(b-a)}{\sigma^2}\right) \int_{\overline{k}}^{b} \exp\left(-\frac{\big(x+2n(b-a)-\mu T\big)^2}{2\sigma^2 T}\right) dx \\
= \;& K e^{-rT} \left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}} \left\{ N\left(\frac{b-\mu T+2n(b-a)}{\sigma\sqrt{T}}\right) - N\left(\frac{\overline{k}-\mu T+2n(b-a)}{\sigma\sqrt{T}}\right) \right\} \\
= \;& K e^{-rT} \left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}} \left\{ N\left(\frac{\log\frac{H^{2n+1}}{S_0 L^{2n}} - \mu T}{\sigma\sqrt{T}}\right) - N\left(\frac{\log\frac{\overline{K} H^{2n}}{S_0 L^{2n}} - \mu T}{\sigma\sqrt{T}}\right) \right\}.
\end{aligned}$$

3.

$$\begin{aligned}
&\frac{S_0 e^{-rT}}{\sigma\sqrt{2\pi T}} \exp\left(\frac{-2\mu\big(n(b-a)-b\big)}{\sigma^2}\right) \int_{\overline{k}}^{b} e^x \exp\left(-\frac{\big(x-2b+2n(b-a)-\mu T\big)^2}{2\sigma^2 T}\right) dx \\
= \;& S_0 e^{-rT} \left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}} \left(\frac{H}{S_0}\right)^{\frac{2\mu}{\sigma^2}} \exp\left(\mu T + 2b - 2n(b-a) + \frac{1}{2}\sigma^2 T\right)
\end{aligned}$$

$$\left\{N\left(\frac{b-\mu T-2b+2n(b-a)-\sigma^2 T}{\sigma\sqrt{T}}\right)-N\left(\frac{\overline{k}-\mu T-2b+2n(b-a)-\sigma^2 T}{\sigma\sqrt{T}}\right)\right\}$$

$$= S_0 e^{-qT}\left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}}\left(\frac{H}{S_0}\right)^{\frac{2\mu}{\sigma^2}}\left(\frac{L}{H}\right)^{2n}\frac{H^2}{S_0^2}$$

$$\left\{N\left(\frac{\log\frac{S_0H^{2n-1}}{L^{2n}}-(\mu+\sigma^2)T}{\sigma\sqrt{T}}\right)-N\left(\frac{\log\frac{\overline{K}S_0H^{2n-2}}{L^{2n}}-(\mu+\sigma^2)T}{\sigma\sqrt{T}}\right)\right\}.$$

4.

$$\frac{Ke^{-rT}}{\sigma\sqrt{2\pi T}}-\exp\left(\frac{-2\mu\big(n(b-a)-b\big)}{\sigma^2}\right)\int_{\overline{k}}^{b}\exp\left(-\frac{\big(x-2b+2n(b-a)-\mu T\big)^2}{2\sigma^2 T}\right)dx$$

$$= Ke^{-rT}\left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}}\left(\frac{H}{S_0}\right)^{\frac{2\mu}{\sigma^2}}$$

$$\left\{N\left(\frac{b-\mu T-2b+2n(b-a)}{\sigma\sqrt{T}}\right)-N\left(\frac{\overline{k}-\mu T-2b+2n(b-a)}{\sigma\sqrt{T}}\right)\right\}$$

$$= Ke^{-rT}\left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}}\left(\frac{H}{S_0}\right)^{\frac{2\mu}{\sigma^2}}\left\{N\left(\frac{\log\frac{S_0H^{2n-1}}{L^{2n}}-\mu T}{\sigma\sqrt{T}}\right)-N\left(\frac{\log\frac{\overline{K}S_0H^{2n-2}}{L^{2n}}-\mu T}{\sigma\sqrt{T}}\right)\right\}.$$

These complete the proof. ■

Theorem 8.4.2 (Double Barrier Knock Out Put). The price of a knock out put with barriers H, $L(L<S_0<H)$, and with strike $K(>L)$, denoted by $\Pi_p(0,S_0,H,L)$, is given by

$$\begin{aligned}
&\Pi_p(0,S_0,H,L)\\
=&\sum_{n=-\infty}^{\infty}\left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}}\left[Ke^{-rT}\left\{N\left(\frac{\log\frac{\underline{K}H^{2n}}{S_0L^{2n}}-\mu T}{\sigma\sqrt{T}}\right)-N\left(\frac{\log\frac{H^{2n}}{S_0L^{2n-1}}-\mu T}{\sigma\sqrt{T}}\right)\right\}\right.\\
&-S_0e^{-qT}\left(\frac{L}{H}\right)^{2n}\left\{N\left(\frac{\log\frac{\underline{K}H^{2n}}{S_0L^{2n}}-(\mu+\sigma^2)T}{\sigma\sqrt{T}}\right)-N\left(\frac{\log\frac{H^{2n}}{S_0L^{2n-1}}-(\mu+\sigma^2)T}{\sigma\sqrt{T}}\right)\right\}\\
&-Ke^{-rT}\left(\frac{H}{S_0}\right)^{\frac{2\mu}{\sigma^2}}\left\{N\left(\frac{\log\frac{\underline{K}S_0H^{2n-2}}{L^{2n}}-\mu T}{\sigma\sqrt{T}}\right)-N\left(\frac{\log\frac{S_0H^{2n-2}}{L^{2n-1}}-\mu T}{\sigma\sqrt{T}}\right)\right\}\\
&\left.+S_0e^{-qT}\left(\frac{H}{S_0}\right)^{\frac{2\mu}{\sigma^2}}\left(\frac{L}{H}\right)^{2n}\frac{H^2}{S_0^2}\left\{N\left(\frac{\log\frac{\underline{K}S_0H^{2n-2}}{L^{2n}}-(\mu+\sigma^2)T}{\sigma\sqrt{T}}\right)-N\left(\frac{\log\frac{S_0H^{2n-2}}{L^{2n-1}}-(\mu+\sigma^2)T}{\sigma\sqrt{T}}\right)\right\}\right],
\end{aligned}$$

where $\underline{K}=\min(H,K)$.

Proof. We follow the same notation in Theorem (8.4.1). We have

$$\begin{aligned}
\Pi_p(0,S_0,H,L) &= e^{-rT}E\Big[(K-S_T)^+\Big|L<m_S(T)<M_S(T)<H\Big]\\
&= e^{-rT}E\left[(K-S_0e^{X_T})^+\middle|\log\frac{L}{S_0}<m_X(T)<M_X(T)<\log\frac{H}{S_0}\right]
\end{aligned}$$

$$= \quad e^{-rT}\int_a^{\underline{k}}(K-S_0e^x)^+h_{ab}(x)dx,$$

where $\underline{k}=\min\left(\log\frac{H}{S_0},\log\frac{K}{S_0}\right)$, $a=\log\frac{L}{S_0}$, $b=\log\frac{H}{S_0}$ and $h_{ab}(x)$ is given in Theorem (8.4.1).

Thus we could have

$$\begin{aligned}\Pi_p(0,S_0,H,L) \quad &= \quad \frac{e^{-rT}}{\sigma\sqrt{2\pi T}}\sum_{n=-\infty}^{\infty}\left[\exp\left(\frac{-2\mu n(b-a)}{\sigma^2}\right)\int_a^{\underline{k}}(K-S_0e^x)\exp\left(-\frac{\big(x+2n(b-a)-\mu T\big)^2}{2\sigma^2T}\right)dx\right.\\ &\qquad \left.-\exp\left(\frac{-2\mu\big(n(b-a)-b\big)}{\sigma^2}\right)\int_a^{\overline{k}}(K-S_0e^x)\exp\left(-\frac{\big(x-2b+2n(b-a)-\mu T\big)^2}{2\sigma^2T}\right)dx\right].\end{aligned}$$

It follows from (7.1.1) that

1.
$$\begin{aligned}&\frac{S_0e^{-rT}}{\sigma\sqrt{2\pi T}}\exp\left(\frac{-2\mu n(b-a)}{\sigma^2}\right)\int_a^{\underline{k}}e^x\exp\left(-\frac{\big(x+2n(b-a)-\mu T\big)^2}{2\sigma^2T}\right)dx\\ =\quad & S_0e^{-rT}\left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}}\exp\left(\mu T-2n(b-a)+\frac{1}{2}\sigma^2T\right)\\ &\left\{N\left(\frac{\overline{k}-\mu T+2n(b-a)-\sigma^2T}{\sigma\sqrt{T}}\right)-N\left(\frac{a-\mu T+2n(b-a)-\sigma^2T}{\sigma\sqrt{T}}\right)\right\}\\ =\quad & S_0e^{-qT}\left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}}\left(\frac{L}{H}\right)^{2n}\left\{N\left(\frac{\log\frac{\underline{K}H^{2n}}{S_0L^{2n}}-(\mu+\sigma^2)T}{\sigma\sqrt{T}}\right)-N\left(\frac{\log\frac{H^{2n}}{S_0L^{2n-1}}-(\mu+\sigma^2)T}{\sigma\sqrt{T}}\right)\right\},\end{aligned}$$

where $\underline{K}=\min(H,K)$.

2.
$$\begin{aligned}&\frac{Ke^{-rT}}{\sigma\sqrt{2\pi T}}\exp\left(\frac{-2\mu n(b-a)}{\sigma^2}\right)\int_a^{\underline{k}}\exp\left(-\frac{\big(x+2n(b-a)-\mu T\big)^2}{2\sigma^2T}\right)dx\\ =\quad & Ke^{-rT}\left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}}\left\{N\left(\frac{\underline{k}-\mu T+2n(b-a)}{\sigma\sqrt{T}}\right)-N\left(\frac{a-\mu T+2n(b-a)}{\sigma\sqrt{T}}\right)\right\}\\ =\quad & Ke^{-rT}\left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}}\left\{N\left(\frac{\log\frac{\underline{K}H^{2n}}{S_0L^{2n}}-\mu T}{\sigma\sqrt{T}}\right)-N\left(\frac{\log\frac{H^{2n}}{S_0L^{2n-1}}-\mu T}{\sigma\sqrt{T}}\right)\right\}.\end{aligned}$$

3.
$$\begin{aligned}&\frac{S_0e^{-rT}}{\sigma\sqrt{2\pi T}}\exp\left(\frac{-2\mu\big(n(b-a)-b\big)}{\sigma^2}\right)\int_a^{\underline{k}}e^x\exp\left(-\frac{\big(x-2b+2n(b-a)-\mu T\big)^2}{2\sigma^2T}\right)dx\\ =\quad & S_0e^{-rT}\left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}}\left(\frac{H}{S_0}\right)^{\frac{2\mu}{\sigma^2}}\exp\left(\mu T+2b-2n(b-a)+\frac{1}{2}\sigma^2T\right)\\ &\left\{N\left(\frac{\underline{k}-\mu T-2b+2n(b-a)-\sigma^2T}{\sigma\sqrt{T}}\right)-N\left(\frac{a-\mu T-2b+2n(b-a)-\sigma^2T}{\sigma\sqrt{T}}\right)\right\}\end{aligned}$$

$$= S_0e^{-qT}\left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}}\left(\frac{H}{S_0}\right)^{\frac{2\mu}{\sigma^2}}\left(\frac{L}{H}\right)^{2n}\frac{H^2}{S_0^2}$$
$$\left\{N\left(\frac{\log\frac{KS_0H^{2n-2}}{L^{2n}}-(\mu+\sigma^2)T}{\sigma\sqrt{T}}\right)-N\left(\frac{\log\frac{S_0H^{2n-2}}{L^{2n-1}}-(\mu+\sigma^2)T}{\sigma\sqrt{T}}\right)\right\}.$$

4.

$$\frac{Ke^{-rT}}{\sigma\sqrt{2\pi T}}-\exp\left(\frac{-2\mu\big(n(b-a)-b\big)}{\sigma^2}\right)\int_a^{\underline{k}}\exp\left(-\frac{\big(x-2b+2n(b-a)-\mu T\big)^2}{2\sigma^2T}\right)dx$$
$$= Ke^{-rT}\left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}}\left(\frac{H}{S_0}\right)^{\frac{2\mu}{\sigma^2}}$$
$$\left\{N\left(\frac{\underline{k}-\mu T-2b+2n(b-a)}{\sigma\sqrt{T}}\right)-N\left(\frac{a-\mu T-2b+2n(b-a)}{\sigma\sqrt{T}}\right)\right\}$$
$$= Ke^{-rT}\left(\frac{L}{H}\right)^{\frac{2\mu n}{\sigma^2}}\left(\frac{H}{S_0}\right)^{\frac{2\mu}{\sigma^2}}\left\{N\left(\frac{\log\frac{\underline{K}S_0H^{2n-2}}{L^{2n}}-\mu T}{\sigma\sqrt{T}}\right)-N\left(\frac{\log\frac{S_0H^{2n-2}}{L^{2n-1}}-\mu T}{\sigma\sqrt{T}}\right)\right\}.$$

These complete the proof of the theorem. ■

Example 8.4.3. Consider an *up-out-and-down-in call* option. If stock price S reaches the upper barrier H, the the option is worthless. Also an *up-out-and-down-in call* option is a up-and-out call option that comes into existence only if the lower barrier L is reached. By in-out parity we can see that

$$\text{up-out-and-down-in call} + \text{double out call} = \text{up-and-out call}.$$

Figure 8.14 gives the boundary conditions for an *up-out-and-down-in call* option. □

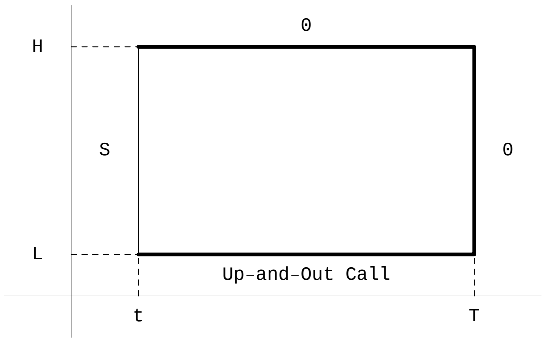


Figure 8.14: Up-Out-Down-In Call Option

8.4.2 Double Knock Out Binary

Hui(1996) have suggested double barrier option pricing method.

We will use the technique introduced in Section 1.16. Consider a Black-Scholes PDE

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r-\delta) S \frac{\partial V}{\partial S} - rV = 0.$$

with $0 < L < H$,

$$V(H,t) = 0, \quad V(L,t) = 0, \quad \text{for all } t < T,$$

and

$$V(S,T) \;=\; 1, \quad L < S < H.$$

If we write

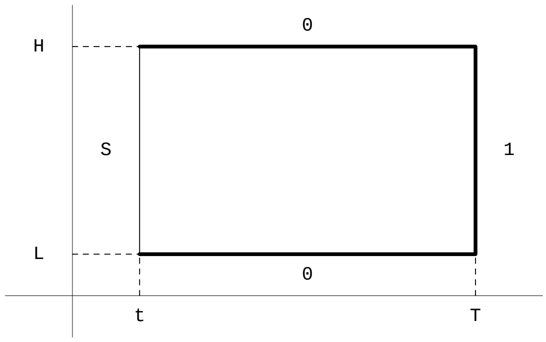


Figure 8.15: Double Knock-Out Binary

$$x \;=\; \log\frac{S}{L}, \quad \tau \;=\; \frac{\sigma^2}{2}(T-t), \quad V(S,t) \;=\; L\, e^{\alpha x + \beta \tau} U(x,\tau),$$

where

$$\alpha \;=\; -\frac{1}{2}\left(\frac{2(r-\delta)}{\sigma^2} - 1\right), \quad \beta \;=\; -\frac{1}{4}\left(\frac{2(r-\delta)}{\sigma^2} - 1\right)^2 - \frac{2r}{\sigma^2},$$

then the Black-Scholes equation could be transformed into a heat equation

$$\frac{\partial U}{\partial \tau} \;=\; \frac{\partial^2 U}{\partial x^2}, \quad 0 < x < \log\frac{H}{L}$$

in which the boundary conditions are given by

$$U(0,\tau) \;=\; 0, \quad U\left(\log\frac{H}{L}, \tau\right) \;=\; 0, \quad \text{for } \tau > 0 \tag{8.4.1}$$

and the initial condition is given by

$$U(x,0) \;=\; \frac{e^{-\alpha x}}{L}, \quad 0 < x < \log\frac{H}{L}. \tag{8.4.2}$$

Let $\omega = \frac{\pi}{\log\frac{H}{L}}$. Then the equation

$$U(x,\tau) \;=\; \sum_{n=0}^{\infty} C_n e^{-n^2\omega^2\tau} \sin(n\omega x)$$

is a solution to the heat equation satisfying boundary condition (8.4.1). All that remains is to choose C_n to satisfy the initial condition (8.4.2):

$$\sum_{n=0}^{\infty} C_n \sin(n\omega x) \quad = \quad \frac{e^{-\alpha x}}{L}.$$

Multiplying both side by

$$\sin(m\omega x)$$

and integrating between 0 and $\log \frac{H}{L} =: A$ we find that

$$\sum_{n=0}^{\infty} C_n \int_0^A \sin(n\omega x)\sin(m\omega x)\,dx \quad = \quad \int_0^A \frac{e^{-\alpha x}}{L}\sin(m\omega x)\,dx.$$

Since on LHS

$$\sum_{n=0}^{\infty} C_n \int_0^A \sin(n\omega x)\sin(m\omega x)\,dx \quad = \quad C_m \frac{A}{2},$$

we have

$$\begin{aligned} C_m &= \frac{2}{A}\int_0^A \frac{e^{-\alpha x}}{L}\sin(m\omega x)\,dx \\ &= \frac{2\pi m}{LA^2}\left(\frac{1-(-1)^n e^{-\alpha L}}{\alpha^2 + \left(\frac{n\pi}{L}\right)^2}\right) \\ &= \frac{2\pi m}{LA^2}\left(\frac{1-(-1)^n \left(\frac{H}{L}\right)^\alpha}{\alpha^2 + \left(\frac{n\pi}{L}\right)^2}\right). \end{aligned}$$

Hence we obtain the $V(S,t)$ as

$$\begin{aligned} V(S,t) &= L\,e^{\alpha x + \beta\tau}\sum_{n=0}^{\infty} C_n e^{-n^2\omega^2\tau}\sin(n\omega x) \\ &= \sum_{n=1}^{\infty}\frac{2\pi n}{A^2}\left(\frac{\left(\frac{S}{L}\right)^\alpha - (-1)^n\left(\frac{S}{H}\right)^\alpha}{\alpha^2 + \left(\frac{n\pi}{L}\right)^2}\right)\sin\left(\frac{n\pi}{A}\log\frac{S}{L}\right)\exp\left[-\frac{1}{2}\left(\left(\frac{n\pi}{A}\right)^2 - \beta\right)\sigma^2(T-t)\right]. \end{aligned}$$

8.4.3 American Binary Knock Out Option

In this option, if S ever reach on barrier, H, then the option is worthless; thus on the line H the option value is zero. If S ever reach another barrier, L, the payment is 1 at the time of touching the payment barrier L. Therefore the boundary conditions for $t < T$ of an American binary knock out option are

$$V(L,t) \quad = \quad 1, \quad V(H,t) \quad = \quad 0.$$

The option is worthless at maturity if the barrier L is never touched, thus the final condition of Black-Scholes PDE is

$$V(S,T) \quad = \quad 0, \quad L < S < H$$

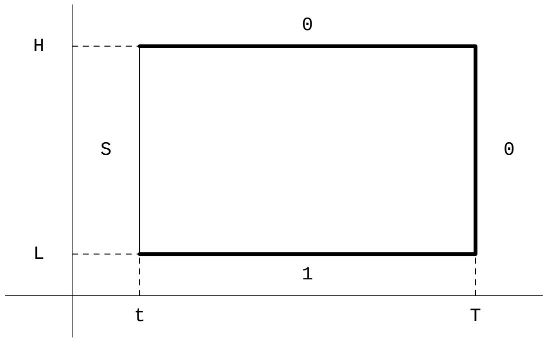


Figure 8.16: American Binary Knock-Out Option

where we assume $L < S < H$. By the same transformation in the previous section, we get a standard heat equation with boundary condition

$$U(0,\tau) = \frac{e^{-\beta\tau}}{L}, \quad U\left(\log\frac{H}{L},\tau\right) = 0$$

and the initial condition is

$$U(x,0) = 0, \quad 0 < x < \log\frac{H}{L}.$$

Let $A := \log\frac{H}{L}$ and

$$W(x,\tau) = U(x,\tau) - y(x,\tau)$$

where

$$y(x,\tau) = \frac{e^{-\beta\tau}}{L}\left(1-\frac{x}{A}\right).$$

We thus find that the boundary condition of W become

$$W(0,\tau) = 0, \quad W(A,\tau) = 0$$

and the initial condition is

$$W(x,0) = \frac{1}{L}\left(1-\frac{x}{A}\right). \tag{8.4.3}$$

Therefore the heat equation is modified as

$$\frac{\partial W}{\partial \tau} = \frac{\partial^2 W}{\partial x^2} + \frac{\beta}{L}\left(1-\frac{x}{A}\right)e^{-\beta\tau}. \tag{8.4.4}$$

Assume W can be expressed by

$$W(x,\tau) = \sum_{n=1}^{\infty} b_n(\tau)\sin\left(\frac{n\pi x}{A}\right) \tag{8.4.5}$$

Substituting (8.4.5) into (8.4.4) gives

$$\sum_{n=1}^{\infty} b_n'(\tau)\sin\left(\frac{n\pi x}{A}\right) = -\sum_{n=1}^{\infty}\left(\frac{n\pi}{A}\right)^2 b_n(\tau)\sin\left(\frac{n\pi x}{A}\right) + \frac{\beta}{L}\left(1-\frac{x}{A}\right)e^{-\beta\tau}.$$

Multiplying both sides by

$$\sin\left(\frac{m\pi x}{A}\right)$$

and integrating between 0 and A we have that

$$b_m'(\tau) + \left(\frac{n\pi}{A}\right)^2 b_m(\tau) = \gamma_m(\tau) \tag{8.4.6}$$

where

$$\begin{aligned}\gamma_m(\tau) &= \frac{2}{A}\int_0^A \frac{\beta}{L}\left(1-\frac{x}{A}\right)e^{-\beta\tau}\sin\left(\frac{m\pi x}{A}\right)dx \\ &= \frac{2\beta}{m\pi L}e^{-\beta\tau}.\end{aligned} \tag{8.4.7}$$

Similarly, from (8.4.3) and (8.4.5), the initial condition is expanded as

$$\begin{aligned}b_m(0) &= \frac{2}{A}\int_0^A \frac{1}{L}\left(\frac{x}{A}-1\right)\sin\left(\frac{m\pi x}{A}\right)dx \\ &= -\frac{2}{m\pi L}.\end{aligned} \tag{8.4.8}$$

By combining (8.4.7) and (8.4.8), the solution of (8.4.6) is

$$\begin{aligned}b_m(\tau) &= e^{-\left(\frac{m\pi}{A}\right)^2\tau}\left\{b_m(0) + \int_0^\tau e^{\left(\frac{m\pi}{A}\right)^2 s}\gamma_m(s)ds\right\} \\ &= -\frac{2}{m\pi L}e^{-\left(\frac{m\pi}{A}\right)^2\tau} + \frac{2\beta}{m\pi L}\int_0^\tau e^{-\left(\frac{m\pi}{A}\right)^2(\tau-s)}e^{-\beta s}ds \\ &= -\frac{2}{m\pi L}e^{-\left(\frac{m\pi}{A}\right)^2\tau} + \frac{2\beta}{m\pi L}\frac{e^{-\beta\tau}-e^{-\left(\frac{m\pi}{A}\right)^2\tau}}{\left(\frac{m\pi}{A}\right)^2-\beta} \\ &= \frac{2}{m\pi L}\left[\frac{\beta e^{-\beta\tau}-\left(\frac{m\pi}{A}\right)^2 e^{-\left(\frac{m\pi}{A}\right)^2\tau}}{\left(\frac{m\pi}{A}\right)^2-\beta}\right]\end{aligned}$$

Thus the solution of (8.4.4) is

$$W(x,\tau) = \sum_{n=1}^{\infty}\frac{2}{n\pi L}\left[\frac{\beta e^{-\beta\tau}-\left(\frac{n\pi}{A}\right)^2 e^{-\left(\frac{n\pi}{A}\right)^2\tau}}{\left(\frac{n\pi}{A}\right)^2-\beta}\right]\sin\left(\frac{n\pi x}{A}\right).$$

We obtain the solution of non-homogeneous heat equation with boundary value from

$$U(x,\tau) = y(x,\tau) + W(x,\tau)$$

$$= \frac{e^{-\beta\tau}}{L}\left(1-\frac{x}{A}\right)+\sum_{n=1}^{\infty}\frac{2}{n\pi L}\left[\frac{\beta e^{-\beta\tau}-\left(\frac{n\pi}{A}\right)^2 e^{-\left(\frac{n\pi}{A}\right)^2\tau}}{\left(\frac{n\pi}{A}\right)^2-\beta}\right]\sin\left(\frac{n\pi x}{A}\right).$$

After putting back the changing variables and transformations, the value of the American binary knock-out option is

$$\begin{aligned}
V(S,t) &= Le^{\alpha x+\beta\tau}U(x,\tau)\\
&= e^{\alpha x}\left(1-\frac{x}{A}\right)+e^{\alpha x}\sum_{n=1}^{\infty}\frac{2}{n\pi}\left[\frac{\beta-\left(\frac{n\pi}{A}\right)^2 e^{-\left[\left(\frac{n\pi}{A}\right)^2-\beta\right]\tau}}{\left(\frac{n\pi}{A}\right)^2-\beta}\right]\sin\left(\frac{n\pi x}{A}\right)\\
&= e^{\alpha x}\left\{\left(1-\frac{x}{A}\right)+\sum_{n=1}^{\infty}\frac{2}{n\pi}\left[\frac{\beta-\left(\frac{n\pi}{A}\right)^2 e^{-\left[\left(\frac{n\pi}{A}\right)^2-\beta\right]\tau}}{\left(\frac{n\pi}{A}\right)^2-\beta}\right]\sin\left(\frac{n\pi x}{A}\right)\right\}\\
&= \left(\frac{S}{L}\right)^{\alpha}\left\{\left(1-\frac{\log\frac{S}{L}}{A}\right)+\sum_{n=1}^{\infty}\frac{2}{n\pi}\left[\frac{\beta-\left(\frac{n\pi}{A}\right)^2 e^{-\frac{1}{2}\left[\left(\frac{n\pi}{A}\right)^2-\beta\right]\sigma^2(T-t)}}{\left(\frac{n\pi}{A}\right)^2-\beta}\right]\sin\left(\frac{n\pi}{A}\log\frac{S}{L}\right)\right\}.
\end{aligned}$$

From the derivation, we find that the assumption $L < S < H$ does not affect the result. Thus the same solution can be applied to the condition

$$L > S > H.$$

Chapter 9

Miscellaneous Exotic Options

9.1 Compound Option

The compound option gives the holer the right to buy or sell another option. The payoff for the compound option depends on the market value of the underlying option, and not on the theoretical price.

✎ **Call on Call**
Let us consider, for instance, the case of a call on a call compound option. For two future dates T_1 and T_2, with $T_1 < T_2$, and two exercise prices K_1 and K_2, consider a call option with exercise price K_1 and expiry date T_1 on a call option with strike price K_2 and maturity T_2. It is clear that the payoff the compound option at time T_1 is

$$\Big(C(S_{T_1}, \tau, K_2) - K_1\Big)^+$$

where $C(S_{T_1}, \tau, K_2)$ stands for the value at time T_1 of a standard call option with strike price K_2 and expiry date $T = T_1 + \tau$. In the Black-Scholes framework, we obtain the following equality

$$C(s, \tau, K_2) \quad = \quad se^{-q\tau} N\big(d_1(s, \tau, K_2)\big) - K_2 e^{-r\tau} N\big(d_2(s, \tau, K_2)\big)$$

The price of the compound option(call on call) at time 0, CC_0, equals

$$CC_0 \quad = \quad e^{-rT_1} \int_{x_0}^{\infty} \Big(g(x)e^{-q\tau} N(\hat{d}_1) - K_2 e^{-r\tau} N(\hat{d}_2) - K_1\Big) n(x) dx$$

where $\hat{d}_i = d_i(g(x), \tau, K)$ for $i = 1, 2$, and the function g and n are given by

$$\begin{aligned} g(x) &= S_0 \exp\Big(\sigma\sqrt{T_1}x + (r - q - \frac{1}{2}\sigma^2)T_1\Big), \\ n(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \end{aligned}$$

and, finally, the constant x_0 is defined implicitly by the equation

$$x_0 \quad := \quad \inf\{x | C(g(x), \tau, K_2) \geq K_1\}.$$

Straightforward calculations yield

$$\begin{aligned}
d_1(g(x),\tau,K_2) &= \frac{\log\frac{S_0}{K_2}+\sigma\sqrt{T_1}x+(r-q-\frac{1}{2}\sigma^2)T_1+(r-q+\frac{1}{2}\sigma^2)(T_2-T_1)}{\sigma\sqrt{T_2-T_1}}\\
&= \frac{\log\frac{S_0}{K_2}+\sigma\sqrt{T_1}x+(r-q)T_2-\sigma^2T_1+\frac{1}{2}\sigma^2T_2}{\sigma\sqrt{T_2-T_1}},\\
d_2(g(x),\tau,K_2) &= \frac{\log\frac{S_0}{K_2}+\sigma\sqrt{T_1}x+(r-q-\frac{1}{2}\sigma^2)T_1+(r-q-\frac{1}{2}\sigma^2)(T_2-T_1)}{\sigma\sqrt{T_2-T_1}}\\
&= \frac{\log\frac{S_0}{K_2}+\sigma\sqrt{T_1}x+(r-q)T_2-\frac{1}{2}\sigma^2T_2}{\sigma\sqrt{T_2-T_1}}.
\end{aligned}$$

Put

$$S^* := S_0\exp\left(\sigma\sqrt{T_1}x_0+\left(r-q-\frac{1}{2}\sigma^2\right)T_1\right),$$

then, we have

$$x_0 = \frac{\log\frac{S^*}{S_0}-(r-q-\frac{1}{2}\sigma^2)T_1}{\sigma\sqrt{T_1}}.$$

Now, let us decompose CO_0 into I, II, and III, as follows.

$$\begin{aligned}
CC_0 &= I-II-III,\\
I &= e^{-rT_1}\int_{x_0}^{\infty} g(x)e^{-q\tau}N(\hat{d}_1)n(x)dx,\\
II &= e^{-rT_1}\int_{x_0}^{\infty} K_2e^{-r\tau}N(\hat{d}_2)n(x)dx,\\
III &= e^{-rT_1}\int_{x_0}^{\infty} K_1n(x)dx.
\end{aligned}$$

We have that

$$\begin{aligned}
I &= e^{-rT_1}e^{-q\tau}\frac{1}{2\pi}\int_{x_0}^{\infty} S_0\exp\Big(\sigma\sqrt{T_1}x+(r-q-\frac{1}{2}\sigma^2)T_1\Big)\left\{\int_{-\infty}^{\hat{d}_1}\exp\Big(-\frac{1}{2}y^2\Big)dy\right\}\exp\Big(-\frac{1}{2}x^2\Big)dx\\
&\quad\downarrow\left\{\begin{array}{lll} -x & =: & \hat{x}\\ y+\frac{\sqrt{T_1}}{\sqrt{T_2-T_1}}\hat{x} & =: & \hat{y}\end{array}\right.\\
&= \frac{1}{2\pi}\int_{-\infty}^{-x_0} S_0e^{-qT_2}\exp\left(-\sigma\sqrt{T_1}\hat{x}-\frac{1}{2}\sigma^2T_1\right)\\
&\qquad\left\{\int_{-\infty}^{\frac{\log\frac{S_0}{K_2}+(r-q)T_2-\sigma^2T_1+\frac{1}{2}\sigma^2T_2}{\sigma\sqrt{T_2-T_1}}}\exp\left(-\frac{1}{2}\left(\hat{y}-\frac{\sqrt{T_1}}{\sqrt{T_2-T_1}}\hat{x}\right)^2\right)d\hat{y}\right\}\exp\left(-\frac{1}{2}\hat{x}^2\right)d\hat{x}\\
&\quad\downarrow\left\{\begin{array}{lll} \hat{x}+\sigma\sqrt{T_1} & =: & x'\\ \hat{y}+\frac{\sigma T_1}{\sqrt{T_2-T_1}} & =: & y'\end{array}\right.\\
&= \frac{S_0e^{-qT_2}}{2\pi}\int_{-\infty}^{\frac{\log\frac{S_0}{S^*}+(r-q+\frac{1}{2}\sigma^2)T_1}{\sigma\sqrt{T_1}}}\int_{-\infty}^{\frac{\log\frac{S_0}{K_2}+(r-q)T_2+\frac{1}{2}\sigma^2T_2}{\sigma\sqrt{T_2-T_1}}}
\end{aligned}$$

$$
\begin{aligned}
& \exp\left(-\frac{1}{2(T_2-T_1)}\left(T_2(x'^2+y'^2)-2\sqrt{(T_2-T_1)T_1}x'y'-T_1y'^2\right)\right)dx'dy' \\
& \downarrow\begin{cases} x' & =: \ \tilde{x} \\ y'\frac{\sqrt{T_2-T_1}}{\sqrt{T_2}} & =: \ \tilde{y} \end{cases} \\
= \ & S_0e^{-qT_2}\ \frac{1}{2\pi}\frac{\sqrt{T_2}}{\sqrt{T_2-T_1}}\int_{-\infty}^{\frac{\log\frac{S_0}{S^*}+(r-q+\frac{1}{2}\sigma^2)T_1}{\sigma\sqrt{T_1}}}\int_{-\infty}^{\frac{\log\frac{S_0}{K_2}+(r-q+\frac{1}{2}\sigma^2)T_2}{\sigma\sqrt{T_2}}}\exp\left(-\frac{1}{2}\frac{T_2}{T_2-T_1}\left(\tilde{x}^2+\tilde{y}^2-2\frac{\sqrt{T_1}}{\sqrt{T_2}}\tilde{x}\tilde{y}\right)\right)d\tilde{y}d\tilde{x} \\
= \ & S_0e^{-qT_2}M\left(\frac{\log\frac{S_0}{S^*}+(r-q+\frac{1}{2}\sigma^2)T_1}{\sigma\sqrt{T_1}},\frac{\log\frac{S_0}{K_2}+(r-q+\frac{1}{2}\sigma^2)T_2}{\sigma\sqrt{T_2}};\sqrt{\frac{T_1}{T_2}}\right),
\end{aligned}
$$

where $M(a,b;\rho)$ is the cumulative probability in a standardized bivariate normal distribution that the first variable is less than a and the second variable is less than b, when the coefficient of correlation between the variables is ρ [1].

Similarly we get

$$
\begin{aligned}
II &= K_2e^{-rT_2}M\left(\frac{\log\frac{S_0}{S^*}+(r-q-\frac{1}{2}\sigma^2)T_1}{\sigma\sqrt{T_1}},\frac{\log\frac{S_0}{K_2}+(r-q-\frac{1}{2}\sigma^2)T_2}{\sigma\sqrt{T_2}};\sqrt{\frac{T_1}{T_2}}\right), \quad \text{and} \\
III &= K_1e^{-rT_1}N\left(\frac{\log\frac{S_0}{S^*}+(r-q-\frac{1}{2}\sigma^2)T_1}{\sigma\sqrt{T_1}}\right).
\end{aligned}
$$

Put

$$
\begin{aligned}
a_1 &:= \frac{\log\frac{S_0}{S^*}+\left(r-q+\frac{\sigma^2}{2}\right)T_1}{\sigma\sqrt{T_1}}, \qquad a_2 := a_1-\sigma\sqrt{T_1}, \\
b_1 &:= \frac{\log\frac{S_0}{K_2}+\left(r-q+\frac{\sigma^2}{2}\right)T_2}{\sigma\sqrt{T_2}}, \qquad b_2 := b_1-\sigma\sqrt{T_2}.
\end{aligned}
$$

Then we have

$$
CC0 = S_0e^{-qT_2}M\left(a_1,b_1;\sqrt{\frac{T_1}{T_2}}\right)-K_2e^{-rT_2}M\left(a_2,b_2;\sqrt{\frac{T_1}{T_2}}\right)-e^{-rT_1}K_1N(a_2).
$$

✎ **Put on Call**

The price of the compound option(call on put) at time 0, PC_0, equals

$$
PC_0 = CC_0 - C(S_0,T_2,K_2)+K_1e^{-rT_1}
$$

by the put-call parity.

Hence we have

$$
PC_0 = S_0e^{-qT_2}M\left(a_1,b_1;\sqrt{\frac{T_1}{T_2}}\right)-K_2e^{-rT_2}M\left(a_2,b_2;\sqrt{\frac{T_1}{T_2}}\right)-e^{-rT_1}K_1N(a_2)
$$

[1] The bivariate normal density involving the individual parameters $\mu_1,\mu_2,\sigma_1,\sigma_2$ and ρ is given by

$$
f(x_1,x_2) = \frac{1}{2\pi\sqrt{\sigma_1\sigma_2(1-\rho^2)}}\times\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_1}}\right)^2+\left(\frac{x_2-\mu_2}{\sqrt{\sigma_2}}\right)^2-2\rho\left(\frac{x_1-\mu_1}{\sqrt{\sigma_1}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_2}}\right)\right]\right\}
$$

$$
\begin{aligned}
& -S_0 e^{-qT_2} N(b_1) + K_2 e^{-rT_2} N(b_2) + K_1 e^{-rT_1} \\
= \quad & K_2 e^{-rT_2} \left\{ N(b_2) - M\left(a_2, b_2; \sqrt{\frac{T_1}{T_2}}\right) \right\} - S_0 e^{-qT_2} \left\{ N(b_1) - M\left(a_1, b_1; \sqrt{\frac{T_1}{T_2}}\right) \right\} + K_1 e^{-rT_1} \{1 - N(a_2)\} \\
= \quad & K_2 e^{-rT_2} M\left(-a_2, b_2; -\sqrt{\frac{T_1}{T_2}}\right) - S_0 e^{-qT_2} M\left(-a_1, b_1; -\sqrt{\frac{T_1}{T_2}}\right) + K_1 e^{-rT_1} N(-a_2).
\end{aligned}
$$

Note that

$$
N(y) - M(x, y; \rho) \quad = \quad M(-x, y; -\rho).
$$

✎ **Call on Put**

In the Black-Scholes framework, we obtain the following equality

$$
P(s, \tau, K_2) \quad = \quad K_2 e^{-r\tau} N\big(-d_2(s, \tau, K_2)\big) - s e^{-q\tau} N\big(-d_1(s, \tau, K_2)\big)
$$

The price of the compound option(call on put) at time 0, CP_0, equals

$$
CP_0 \quad = \quad e^{-rT_1} \int_{-\infty}^{x_0} \Big(K_2 e^{-r\tau} N(-\hat{d}_2) - g(x) e^{-q\tau} N(-\hat{d}_1) - K_1 \Big) n(x) dx
$$

The constant x_0 is defined implicitly by the equation

$$
x_0 \quad = \quad \sup\{x | P(g(x), \tau, K_2) \geq K_1\}.
$$

Now let us decompose CP_0 as follows:

$$
\begin{aligned}
CP_0 \quad &= \quad I - II - III, \\
I \quad &= \quad K_2 e^{-rT_2} \int_{-\infty}^{x_0} N(-\hat{d}_2) n(x) dx, \\
II \quad &= \quad e^{-rT_1} e^{-q\tau} \int_{-\infty}^{x_0} g(x) N(-\hat{d}_1) n(x) dx, \\
III \quad &= \quad K_1 e^{-rT_1} \int_{-\infty}^{x_0} n(x) dx.
\end{aligned}
$$

Firstly compute I:

$$
\begin{aligned}
I \quad = \quad & \frac{K_2 e^{-rT_2}}{2\pi} \int_{-\infty}^{x_0} \int_{-\infty}^{-\hat{d}_2} \exp\left(-\frac{1}{2} y^2\right) \exp\left(-\frac{1}{2} x^2\right) \; dy dx \\
= \quad & \frac{K_2 e^{-rT_2}}{2\pi} \int_{-\infty}^{-a_2} \int_{-\infty}^{-\frac{\log \frac{S_0}{K_2} + \sigma\sqrt{T_1} x + \left(r - q - \frac{1}{2}\sigma^2\right) T_2}{\sigma\sqrt{T_2 - T_1}}} \exp\left(-\frac{1}{2} y^2\right) \exp\left(-\frac{1}{2} x^2\right) \; dy dx \\
& \downarrow \left\{ \begin{array}{lll} x & =: & x' \\ y + \frac{\sqrt{T_1}}{\sqrt{T_2 - T_1}} x & =: & y' \end{array} \right. \\
= \quad & \frac{K_2 e^{-rT_2}}{2\pi} \int_{-\infty}^{-a_2} \int_{-\infty}^{-\frac{\log \frac{S_0}{K_2} + (r-q)T_2 - \frac{1}{2}\sigma^2 T_2}{\sigma\sqrt{T_2 - T_1}}} \exp\left(-\frac{1}{2}\left(y' - \frac{\sqrt{T_1}}{\sqrt{T_2 - T_1}} x'\right)^2\right) \exp\left(-\frac{1}{2} x'^2\right) \; dy' dx'
\end{aligned}
$$

$$\downarrow \left\{ \begin{array}{lcl} x' & =: & \tilde{x} \\ y' \frac{\sqrt{T_2 - T_1}}{\sqrt{T_2}} & =: & \tilde{y} \end{array} \right.$$

$$\begin{aligned}
&= \frac{K_2 e^{-rT_2}}{2\pi} \frac{\sqrt{T_2}}{\sqrt{T_2 - T_1}} \int_{-\infty}^{-a_2} \int_{-\infty}^{-\frac{\log \frac{S_0}{K_2} + \left(r - q - \frac{1}{2}\sigma^2\right) T_2}{\sigma\sqrt{T_2}}} \exp\left(-\frac{1}{2} \left(\frac{\sqrt{T_2}}{\sqrt{T_2 - T_1}} \tilde{y} - \frac{\sqrt{T_1}}{\sqrt{T_2 - T_1}} \tilde{x} \right)^2 \right) \exp\left(-\frac{1}{2} \tilde{x}^2 \right) \, d\tilde{y} d\tilde{x} \\
&= \frac{K_2 e^{-rT_2}}{2\pi} \frac{\sqrt{T_2}}{\sqrt{T_2 - T_1}} \int_{-\infty}^{-a_2} \int_{-\infty}^{-b_2} \exp\left(-\frac{1}{2} \frac{T_2}{T_2 - T_1} \left(\tilde{x}^2 + \tilde{y}^2 - 2 \frac{\sqrt{T_1}}{\sqrt{T_2}} \tilde{x}\tilde{y} \right) \right) d\tilde{y} d\tilde{x} \\
&= K_2 e^{-rT_2} M\left(-a_2, -b_2; \sqrt{\frac{T_1}{T_2}} \right).
\end{aligned}$$

Secondly compute II:

$$\begin{aligned}
II &= e^{-rT_1} e^{-q\tau} \frac{1}{2\pi} \int_{-\infty}^{x_0} S_0 \exp\left(\sigma\sqrt{T_1} x + (r - q - \frac{1}{2}\sigma^2) T_1 \right) \left\{ \int_{-\infty}^{-\hat{d}_1} \exp\Big(-\frac{1}{2} y^2 \Big) dy \right\} \exp\Big(-\frac{1}{2} x^2 \Big) dx \\
&\quad \downarrow \left\{ \begin{array}{lcl} x & =: & \hat{x} \\ y + \frac{\sqrt{T_1}}{\sqrt{T_2 - T_1}} \hat{x} & =: & \hat{y} \end{array} \right. \\
&= \frac{S_0 e^{-qT_2}}{2\pi} \int_{-\infty}^{-a_2} \exp\left(\sigma\sqrt{T_1} \hat{x} - \frac{1}{2}\sigma^2 T_1 \right) \\
&\qquad \left\{ \int_{-\infty}^{-\frac{\log \frac{S_0}{K_2} + \left(r - q + \frac{1}{2}\sigma^2\right) T_2 - \sigma^2 T_1}{\sigma\sqrt{T_2 - T_1}}} \exp\left(-\frac{1}{2} \left(\hat{y} - \frac{\sqrt{T_1}}{\sqrt{T_2 - T_1}} \hat{x} \right)^2 \right) d\hat{y} \right\} \exp\left(-\frac{1}{2} \hat{x}^2 \right) d\hat{x} \\
&\quad \downarrow \left\{ \begin{array}{lcl} \hat{x} - \sigma\sqrt{T_1} & =: & x' \\ \hat{y} + \frac{\sigma T_1}{\sqrt{T_2 - T_1}} & =: & y' \end{array} \right. \\
&= \frac{S_0 e^{-qT_2}}{2\pi} \int_{-\infty}^{-a_1} \int_{-\infty}^{-b_1} \exp\left(-\frac{1}{2(T_2 - T_1)} \left(T_2 (x'^2 + y'^2) - 2\sqrt{(T_2 - T_1) T_1} x' y' - T_1 y'^2 \right) \right) dx' dy' \\
&\quad \downarrow \left\{ \begin{array}{lcl} x' & =: & \tilde{x} \\ y' \frac{\sqrt{T_2 - T_1}}{\sqrt{T_2}} & =: & \tilde{y} \end{array} \right. \\
&= \frac{S_0 e^{-qT_2}}{2\pi} \frac{\sqrt{T_2}}{\sqrt{T_2 - T_1}} \int_{-\infty}^{-a_1} \int_{-\infty}^{-b_1} \exp\left(-\frac{1}{2} \frac{T_2}{T_2 - T_1} \left(\tilde{x}^2 + \tilde{y}^2 - 2 \frac{\sqrt{T_1}}{\sqrt{T_2}} \tilde{x}\tilde{y} \right) \right) d\tilde{y} d\tilde{x} \\
&= S_0 e^{-qT_2} M\left(-a_1, -b_1; \sqrt{\frac{T_1}{T_2}} \right).
\end{aligned}$$

Finally, III can be computed as follows:

$$\begin{aligned}
III &= K_1 e^{-rT_1} \int_{-\infty}^{x_0} n(x) dx \\
&= K_1 e^{-rT_1} N(-a_2).
\end{aligned}$$

Hence we have

$$CP_0 = K_2 e^{-rT_2} M\left(-a_2, -b_2; \sqrt{\frac{T_1}{T_2}} \right) - S_0 e^{-qT_2} M\left(-a_1, -b_1; \sqrt{\frac{T_1}{T_2}} \right) - K_1 e^{-rT_1} N(-a_2).$$

✎ **Put on Put**

The price of the compound option(call on put) at time 0, PP_0, equals

$$PP_0 = CP_0 - P(S_0, T_2, K_2) + K_1 e^{-rT_1}$$

by the put-call parity.

Hence we have

$$\begin{aligned}
PC_0 &= K_2 e^{-rT_2} M\left(-a_2, -b_2; \sqrt{\frac{T_1}{T_2}}\right) - S_0 e^{-qT_2} M\left(-a_1, -b_1; \sqrt{\frac{T_1}{T_2}}\right) - K_1 e^{-rT_1} N(-a_2) \\
&\quad - K_2 e^{-rT_2} N(-b_2) + S_0 e^{-qT_2} N(-b_1) + K_1 e^{-rT_1} \\
&= S_0 e^{-qT_2}\left\{N(-b_1) - M\left(-a_1, -b_1; \sqrt{\frac{T_1}{T_2}}\right)\right\} - K_2 e^{-rT_2}\left\{N(-b_2) - M\left(-a_2, -b_2; \sqrt{\frac{T_1}{T_2}}\right)\right\} \\
&\qquad + K_1 e^{-rT_1}\left\{1 - N(-a_2)\right\} \\
&= S_0 e^{-qT_2} M\left(a_1, -b_1; -\sqrt{\frac{T_1}{T_2}}\right) - K_2 e^{-rT_2} M\left(a_2, -b_2; -\sqrt{\frac{T_1}{T_2}}\right) + K_1 e^{-rT_1} N(a_2).
\end{aligned}$$

9.2 Chooser Option

A chooser option is an agreement in which one party has the right to choose at some future time T_1 whether the option is to be a call or put option with a common exercise price K and remaining time to expiry $T_2 - T_1$. The value of the choose option at time T_1, CH_{T_1}, is

$$\begin{aligned}
CH_{T_1} &= \max\Big(C_{T_1}, P_{T_1}\Big) \\
&= \max\Big(C_{T_1}, C_{T_1} + Ke^{-r(T_2-T_1)} - S_{T_1} e^{-q(T_2-T_1)}\Big) \\
&= C_{T_1} + e^{-q(T_2-T_1)} \max\Big(0, Ke^{-(r-q)(T_2-T_1)} - S_{T_1}\Big),
\end{aligned}$$

where C_{T_1} and P_{T_1} are call option value and put option value at time T_1, respectively. This show that the chooser option is a package consisting of a call option with strike price K and maturity T_2 and $e^{-q(T_2-T_1)}$ put option with strike price $Ke^{-(r-q)(T_2-T_1)}$ and maturity T_1.

$$\begin{aligned}
CH_0 &= C_0(S_0, T_2, K) + e^{-q(T_2-T_1)} P_0(S_0, T_1, Ke^{-r(T_2-T_1)}) \\
&= S_0 e^{-qT_2} N(d_1) - Ke^{-rT_2} N(d_2) + e^{-q(T_2-T_1)}\Big\{Ke^{-(r-q)(T_2-T_1)} e^{-rT_1} N(-\hat{d}_2) - S_0 e^{-qT_1} N(-\hat{d}_1)\Big\} \\
&= S_0 e^{-qT_2} N(d_1) - Ke^{-rT_2} N(d_2) + Ke^{-rT_2} N(-\hat{d}_2) - S_0 e^{-qT_2} N(-\hat{d}_1) \\
&= S_0 e^{-qT_2}\Big\{N(d_1) - N(-\hat{d}_1)\Big\} - Ke^{-rT_2}\Big\{N(d_2) - N(-\hat{d}_2)\Big\},
\end{aligned}$$

where

$$\begin{aligned}
d_{1,2} &= \frac{\log\frac{S_0}{K} + \left(r - q \pm \frac{1}{2}\sigma^2\right) T_2}{\sigma\sqrt{T_2}}, \\
\hat{d}_{1,2} &= \frac{\log\frac{S_0}{K} + (r-q)(T_2 - T_1) + (r - q \pm \frac{1}{2}\sigma^2) T_1}{\sigma\sqrt{T_1}}, \\
&= \frac{\log\frac{S_0}{K} + (r-q)T_2 \pm \frac{1}{2}\sigma^2 T_1}{\sigma\sqrt{T_1}}.
\end{aligned}$$

9.3 Binary(Digital) Option

The binary(digital) call option's payoff is given by Figure (9.1).

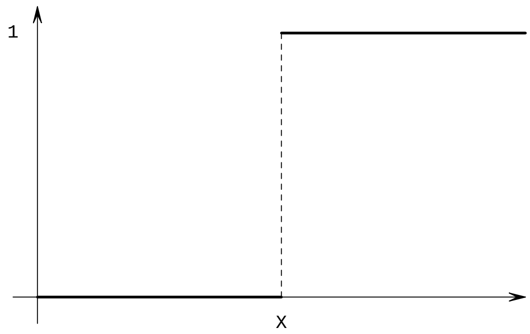


Figure 9.1: The payoff of binary option

With our usual notation, in a risk-neutral world, the probability of the stock price being above the strike price at the maturity of the option is given by

$$\text{Prob}(S_T \geq X) = N(d_2).$$

Hence we have the value of binary call option at time 0, $\Pi(0)$, as follows.

$$\begin{aligned}\Pi(0) &= e^{-rT}E[\mathbf{1}_{\{S_T \geq X\}}] \\ &= e^{-rT}\text{Prob}(S_T > X) \\ &= e^{-rT}N(d_2).\end{aligned}$$

Similarly, the value of a binary put option is

$$e^{-rT}N(-d_2).$$

9.4 Asset or Nothing Option

The *asset-or-nothing call* option pays off nothing if the underlying stock price ends up below the strike price and pays an amount equal to the stock price itself if it ends up above the strike price. With our usual notation, the value of an *asset-or-nothing call* is

$$\begin{aligned}\Pi(0) &= e^{-rT}E[S_T \cdot \mathbf{1}_{\{S_T \geq X\}}] \\ &= e^{-qT}S_0N(d_1).\end{aligned}$$

The value of an *asset-or-nothing put* is

$$e^{-qT}S_0N(-d_1).$$

9.5 Forward Start Options

Forward start option is designed so that it will be at-the-money at some time in the future. Let $c(s, k, \tau)$ be the Black-Scholes call option price where exercise price is k, time to maturity is τ and current stock price s. Consider a froward start European call option that will start at time T_1 and mature at time T_2. The value of the forward start option is given by

$$\begin{aligned}
& e^{-rT_1} E\Big[c(S_{T_1}, S_{T_1}, T_2 - T_1)\Big] \\
= \;& e^{-rT_1} E\left[S_{T_1}\left[e^{-q(T_2-T_1)} N\left(\frac{(r-q+\frac{1}{2}\sigma^2)(T_2-T_1)}{\sigma\sqrt{T_2-T_1}}\right) - e^{-r(T_2-T_1)} N\left(\frac{(r-q-\frac{1}{2}\sigma^2)(T_2-T_1)}{\sigma\sqrt{T_2-T_1}}\right)\right]\right] \\
= \;& e^{-rT_1}\left[e^{-q(T_2-T_1)} N\left(\frac{(r-q+\frac{1}{2}\sigma^2)(T_2-T_1)}{\sigma\sqrt{T_2-T_1}}\right) - e^{-r(T_2-T_1)} N\left(\frac{(r-q-\frac{1}{2}\sigma^2)(T_2-T_1)}{\sigma\sqrt{T_2-T_1}}\right)\right] E\Big[S_{T_1}\Big] \\
= \;& S_0 e^{-qT_1}\left[e^{-q(T_2-T_1)} N\left(\frac{(r-q+\frac{1}{2}\sigma^2)(T_2-T_1)}{\sigma\sqrt{T_2-T_1}}\right) - e^{-r(T_2-T_1)} N\left(\frac{(r-q-\frac{1}{2}\sigma^2)(T_2-T_1)}{\sigma\sqrt{T_2-T_1}}\right)\right] \\
= \;& e^{-qT_1} c(S_0, S_0, T_2 - T_1).
\end{aligned}$$

For a non-dividend-paying stock, $q = 0$ and the value of the forward start option is exactly the same as the value of a regular at-the-money option with the same life as the forward start option.

Alternatively to value the option, we note that the value of an at-the-money call option is proportional to the stock price. The value of forward start option at time T_1 is therefore,

$$c(S_0, S_0, T_2 - T_1)\frac{S_{T_1}}{S_0}.$$

Hence we have

$$\begin{aligned}
& e^{-rT_1} E\Big[c(S_0, S_0, T_2 - T_1)\frac{S_{T_1}}{S_0}\Big] \\
= \;& e^{-rT_1}\frac{c(S_0, S_0, T_2 - T_1)}{S_0} E\Big[S_{T_1}\Big] \\
= \;& e^{-rT_1}\frac{c(S_0, S_0, T_2 - T_1)}{S_0} S_0 e^{(r-q)T_1} \\
= \;& e^{-qT_1} c(S_0, S_0, T_2 - T_1).
\end{aligned}$$

9.6 Shout Option

The shout option is American style option. If the option is shouted at a time τ, when the asset price is S_τ the payoff from the option is

$$\max(0, S_T - S_\tau) + (S_\tau - X)$$

where X is the strike price and S_T is the asset price at time T.

Since there is clearly an element of optimization in the matter of shouting, one would expect to see a free boundary problem occur quite naturally as with American option.

9.7 One-Touch Option

✎ **Up & In One-Touch Option**

The one-touch option is an American version of the binary option. There is no benefit in holding the option once the level has been reached therefore it should be exercised immediately the level is reached for the first time.

As usual, we assume that the underlying process for S_t satisfies the geometric Brownian motion, i.e.

$$dS_t = S_t\Big((r-q)\,dt + \sigma\,dW_t\Big).$$

We define, for $H > S_0$,

$$T_H = \inf\{t : S_t = H\}.$$

i.e. T_H is the first time that the geometric Brownian motion S reaches the level H. Also, define $X_t = \log S_t$,

$$M_X(t) = \sup_{0\le s\le t} X_t, \quad \text{and} \quad m_X(t) = \inf_{0\le s\le t} X_t.$$

By Corollary(6.1.3), we have

$$\begin{aligned}
& P\Big(T_H \le t\Big) \\
= \; & P\Big(M_X(t) \ge \log H\Big) \\
= \; & 1 - P\Big(M_X(t) \le \log H\Big) \\
= \; & 1 - N\left(\frac{\log H - \log S_0 - \mu t}{\sigma\sqrt{t}}\right) + \exp\left(\frac{2\mu(\log H - \log S_0)}{\sigma^2}\right) N\left(-\frac{\log H - \log S_0 + \mu t}{\sigma\sqrt{t}}\right), \qquad (9.7.1)
\end{aligned}$$

where $\mu = r - q - \frac{1}{2}\sigma^2$. By differentiating this with respect to t, we can write the probability density function of T_H, denoted by $f(t)$, as [2]

$$f(t) = \frac{\log\frac{H}{S_0}}{\sqrt{2\pi t}\,t\sigma}\exp\left(-\frac{\Big(\mu t + \log\frac{S_0}{H}\Big)^2}{2t\sigma^2}\right).$$

When $H > S_0$, the one-touch option price, $V(S_0, H, T)$, is given by

$$\begin{aligned}
V(S_0,H,T) &= E[\mathbf{1}_{\{T_H\le T\}} e^{-rT_H}] \\
&= \int_0^T e^{-rt} f(t)dt \\
&= \frac{\log\frac{H}{S_0}}{\sqrt{2\pi}\,\sigma}\int_0^T \frac{1}{\sqrt{t}\,t}\,e^{-rt}\exp\left(-\frac{\Big(\mu t + \log\frac{S_0}{H}\Big)^2}{2t\sigma^2}\right)dt.
\end{aligned}$$

1. First, let us compute the integrand:

$$e^{-rt}\exp\left(-\frac{\Big(\mu t + \log\frac{S_0}{H}\Big)^2}{2t\sigma^2}\right)$$

[2] See Example(6.2.2). $\mu = r - q - \frac{1}{2}\sigma^2, b = \log\frac{H}{S_0}$.

$$
\begin{aligned}
&= \exp\left(-\frac{(\mu^2+2r\sigma^2)t^2+2\mu t\log\frac{S_0}{H}+\left(\log\frac{S_0}{H}\right)^2}{2t\sigma^2}\right) \\
&= \left(\frac{S_0}{H}\right)^{-\frac{\mu}{\sigma^2}}\exp\left(-\frac{(\mu^2+2r\sigma^2)t^2+\left(\log\frac{S_0}{H}\right)^2}{2t\sigma^2}\right) \\
&= \left(\frac{S_0}{H}\right)^{-\frac{\mu}{\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}\left\{(\mu^2+2r\sigma^2)t+\left(\log\frac{S_0}{H}\right)^2\frac{1}{t}\right\}\right).
\end{aligned}
$$

2. Put $y=\frac{1}{\sqrt{t}}$. Then we have

$$
V(S_0,H,T) = \frac{2\log\frac{H}{S_0}}{\sqrt{2\pi}\,\sigma}\left(\frac{S_0}{H}\right)^{-\frac{\mu}{\sigma^2}}\int_{\frac{1}{\sqrt{T}}}^{\infty}\exp\left(-\frac{1}{2\sigma^2}\left\{(\mu^2+2r\sigma^2)\frac{1}{y^2}+\left(\log\frac{S_0}{H}\right)^2y^2\right\}\right)dy.
$$

3. Note that when $a>0, b>0$, we have an indefinite integral

$$
\begin{aligned}
&\int\frac{1}{\sqrt{\pi}}\,e^{-a^2y^2-b^2\frac{1}{y^2}}\;dy \\
&= \frac{1}{2a}\left(e^{-2ab}N\left(\sqrt{2}ay-\frac{\sqrt{2}b}{y}\right)+e^{2ab}N\left(\sqrt{2}ay+\frac{\sqrt{2}b}{y}\right)\right). \qquad (9.7.2)
\end{aligned}
$$

This equality can be easily verified by differentiating (9.7.2).

4. Put $a=\frac{1}{\sqrt{2}\sigma}\log\frac{H}{S_0}$ and $b=\frac{1}{\sqrt{2}\sigma}\sqrt{\mu^2+2r\sigma^2}$. Then we have

$$
\begin{aligned}
V(S_0,H,T) &= \frac{2\log\frac{H}{S_0}}{\sqrt{2}\,\sigma}\left(\frac{S_0}{H}\right)^{-\frac{\mu}{\sigma^2}}\frac{\sqrt{2}\sigma}{2}\frac{1}{\log\frac{H}{S_0}}\,\Xi \\
&= \left(\frac{S_0}{H}\right)^{-\frac{\mu}{\sigma^2}}\Xi.
\end{aligned}
$$

where

$$
\Xi = e^{-2ab}N\left(\sqrt{2}ay-\frac{\sqrt{2}b}{y}\right)+e^{2ab}N\left(\sqrt{2}ay+\frac{\sqrt{2}b}{y}\right)\Bigg|_{y=\frac{1}{\sqrt{T}}}^{y=\infty}.
$$

5. We have

$$
e^{2ab} = \exp\left(\frac{\sqrt{\mu^2+2r\sigma^2}}{\sigma^2}\log\frac{H}{S_0}\right) = \left(\frac{S_0}{H}\right)^{-\frac{\sqrt{\mu^2+2r\sigma^2}}{\sigma^2}}.
$$

Also we have

$$
\begin{aligned}
N\left(\sqrt{2}ay-\frac{\sqrt{2}b}{y}\right)\Bigg|_{y=\frac{1}{\sqrt{T}}}^{y=\infty} &= 1-N\left(\frac{1}{\sigma}\log\frac{H}{S_0}\frac{1}{\sqrt{T}}-\frac{1}{\sigma}\sqrt{\mu^2+2r\sigma^2}\sqrt{T}\right) \\
&= 1-N\left(\frac{\log\frac{H}{S_0}-\sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right)
\end{aligned}
$$

$$= N\left(\frac{\log\frac{S_0}{H}+\sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right),$$

and

$$\begin{aligned}
N\left(\sqrt{2}ay+\frac{\sqrt{2}b}{y}\right)\Bigg|_{y=\frac{1}{\sqrt{T}}}^{y=\infty} &= 1-N\left(\frac{1}{\sigma}\log\frac{H}{S_0}\frac{1}{\sqrt{T}}+\frac{1}{\sigma}\sqrt{\mu^2+2r\sigma^2}\sqrt{T}\right)\\
&= 1-N\left(\frac{\log\frac{H}{S_0}+\sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right)\\
&= N\left(\frac{\log\frac{S_0}{H}-\sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right).
\end{aligned}$$

6. Finally, we have

$$\begin{aligned}
&V(S_0,H,T)\\
=\;& \left(\frac{S_0}{H}\right)^{-\frac{\mu}{\sigma^2}}\left(\frac{S_0}{H}\right)^{\frac{\sqrt{\mu^2+2r\sigma^2}}{\sigma^2}}N\left(\frac{\log\frac{S_0}{H}+\sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right)+\left(\frac{S_0}{H}\right)^{-\frac{\mu}{\sigma^2}}\left(\frac{S_0}{H}\right)^{\frac{-\sqrt{\mu^2+2r\sigma^2}}{\sigma^2}}N\left(\frac{\log\frac{S_0}{H}-\sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right)\\
=\;& \left(\frac{S_0}{H}\right)^{\frac{\sqrt{\mu^2+2r\sigma^2}-\mu}{\sigma^2}}N\left(\frac{\log\frac{S_0}{H}+\sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right)+\left(\frac{S_0}{H}\right)^{\frac{-\sqrt{\mu^2+2r\sigma^2}-\mu}{\sigma^2}}N\left(\frac{\log\frac{S_0}{H}-\sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right).
\end{aligned}$$

Note that when dividend rate $q=0$, we have

$$V(S_0,H,T) = \left(\frac{S_0}{H}\right)N\left(\frac{\log\frac{S_0}{H}+\left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)+\left(\frac{S_0}{H}\right)^{-\frac{2r}{\sigma^2}}N\left(\frac{\log\frac{S_0}{H}-\left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right).$$

Example 9.7.1 (Up & In One-Touch Option with deferred payment). If the stock price reaches level $H > S_0$ before maturity T, the payment 1 occurs at T. This option price, V, is given by

$$V = e^{-rT}P(T_H \le T).$$

From Eq(9.7.1), the hitting probability, $P(T_H \le T)$, is given by

$$\begin{aligned}
&P(T_H \le T)\\
=\;& 1-N\left(\frac{\log H-\log S_0-\mu T}{\sigma\sqrt{T}}\right)+\exp\left(\frac{2\mu(\log H-\log S_0)}{\sigma^2}\right)N\left(-\frac{\log H-\log S_0+\mu T}{\sigma\sqrt{T}}\right)\\
=\;& N\left(\frac{\log S_0-\log H+\mu T}{\sigma\sqrt{T}}\right)+\left(\frac{H}{S_0}\right)^{\frac{2\mu}{\sigma^2}}N\left(-\frac{\log H-\log S_0+\mu T}{\sigma\sqrt{T}}\right)\\
=\;& N\left(\frac{\log\frac{S_0}{H}+\mu T}{\sigma\sqrt{T}}\right)+\left(\frac{H}{S_0}\right)^{\frac{2\mu}{\sigma^2}}N\left(\frac{\log\frac{S_0}{H}-\mu T}{\sigma\sqrt{T}}\right).
\end{aligned}$$

Hence we have

$$V = e^{-rT}N\left(\frac{\log\frac{S_0}{H}+\left(r-q-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)+e^{-rT}\left(\frac{H}{S_0}\right)^{\frac{2(r-q)}{\sigma^2}-1}N\left(\frac{\log\frac{S_0}{H}-\left(r-q-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right).$$

Note that this instrument is up-and-in-bond. Thus V is the same to the up-and-in-bond price. $\square$

✎ Down & In One-Touch Option

We define for $L < S_0$,

$$T_L = \inf\{t : S_t = L\}.$$

The probability density function $h(t)$ for T_L is given by

$$h(t) = -\frac{\log \frac{L}{S_0}}{\sqrt{2\pi t}\, t\sigma} \exp\left(-\frac{\left(\mu t + \log \frac{S_0}{L}\right)^2}{2t\sigma^2}\right).$$

The one-touch option price, $V(S_0, L, T)$ is given by

$$V(S_0, L, T) = -\frac{\log \frac{L}{S_0}}{\sqrt{2\pi}\,\sigma} \int_0^T \frac{1}{\sqrt{t}\, t} e^{-rt} \exp\left(-\frac{\left(\mu t + \log \frac{S_0}{L}\right)^2}{2t\sigma^2}\right) dt.$$

1. Put $a = -\frac{1}{\sqrt{2}\sigma} \log \frac{L}{S_0}$ and $b = \frac{1}{\sqrt{2}\sigma}\sqrt{\mu^2 + 2r\sigma^2}$. Then we have

$$\begin{aligned} V(S_0, L, T) &= -\frac{2\log \frac{L}{S_0}}{\sqrt{2}\,\sigma} \left(\frac{S_0}{L}\right)^{-\frac{\mu}{\sigma^2}} \frac{-\sqrt{2}\sigma}{2\log \frac{L}{S_0}}\, \Xi \\ &= \left(\frac{S_0}{L}\right)^{-\frac{\mu}{\sigma^2}} \Xi. \end{aligned}$$

where

$$\Xi = e^{-2ab} N\left(\sqrt{2}ay - \frac{\sqrt{2}b}{y}\right) + e^{2ab} N\left(\sqrt{2}ay + \frac{\sqrt{2}b}{y}\right)\Bigg|_{\frac{1}{\sqrt{T}}}^{\infty}.$$

2. We have

$$e^{2ab} = \exp\left(-\frac{\sqrt{\mu^2 + 2r\sigma^2}}{\sigma^2} \log \frac{L}{S_0}\right) = \left(\frac{S_0}{L}\right)^{\frac{\sqrt{\mu^2+2r\sigma^2}}{\sigma^2}}.$$

Also we have

$$\begin{aligned} N\left(\sqrt{2}ay - \frac{\sqrt{2}b}{y}\right)\Bigg|_{\frac{1}{\sqrt{T}}}^{\infty} &= 1 - N\left(-\frac{1}{\sigma}\log \frac{L}{S_0}\frac{1}{\sqrt{T}} - \frac{1}{\sigma}\sqrt{\mu^2 + 2r\sigma^2}\sqrt{T}\right) \\ &= 1 - N\left(\frac{-\log \frac{L}{S_0} - \sqrt{\mu^2 + 2r\sigma^2}T}{\sigma\sqrt{T}}\right) \\ &= N\left(\frac{-\log \frac{S_0}{L} + \sqrt{\mu^2 + 2r\sigma^2}T}{\sigma\sqrt{T}}\right), \end{aligned}$$

and

$$N\left(\sqrt{2}ay + \frac{\sqrt{2}b}{y}\right)\Bigg|_{\frac{1}{\sqrt{T}}}^{\infty} = 1 - N\left(-\frac{1}{\sigma}\log \frac{L}{S_0}\frac{1}{\sqrt{T}} + \frac{1}{\sigma}\sqrt{\mu^2 + 2r\sigma^2}\sqrt{T}\right)$$

$$
\begin{aligned}
&= 1 - N\left(\frac{-\log\frac{L}{S_0} + \sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right) \\
&= N\left(\frac{-\log\frac{S_0}{L} - \sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right).
\end{aligned}
$$

3. Finally, we have

$$
\begin{aligned}
&V(S_0,H,T) \\
&= \left(\frac{S_0}{L}\right)^{-\frac{\mu}{\sigma^2}}\left(\frac{S_0}{L}\right)^{\frac{-\sqrt{\mu^2+2r\sigma^2}}{\sigma^2}} N\left(\frac{-\log\frac{S_0}{L} + \sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right) + \left(\frac{S_0}{L}\right)^{-\frac{\mu}{\sigma^2}}\left(\frac{S_0}{L}\right)^{\frac{\sqrt{\mu^2+2r\sigma^2}}{\sigma^2}} N\left(\frac{-\log\frac{S_0}{L} - \sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right) \\
&= \left(\frac{S_0}{L}\right)^{\frac{-\sqrt{\mu^2+2r\sigma^2}-\mu}{\sigma^2}} N\left(\frac{-\log\frac{S_0}{L} + \sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right) + \left(\frac{S_0}{L}\right)^{\frac{\sqrt{\mu^2+2r\sigma^2}-\mu}{\sigma^2}} N\left(\frac{-\log\frac{S_0}{L} - \sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right).
\end{aligned}
$$

Note that when dividend rate $q = 0$, we have

$$
V(S_0,H,T) = \left(\frac{S_0}{L}\right)^{-\frac{2r}{\sigma^2}} N\left(\frac{-\log\frac{S_0}{L} + \left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right) + \left(\frac{S_0}{L}\right) N\left(\frac{-\log\frac{S_0}{L} - \left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right).
$$

Example 9.7.2 (Down & In One-Touch Option with deferred payment). If the stock price reaches level $L < S_0$ before maturity T, the payment 1 occurs at T. This option price, V, is given by

$$
V = e^{-rT}P(T_L \le T).
$$

From Corollary(6.1.3), the hitting probability, $P(T_L \le T)$, is given by

$$
\begin{aligned}
P(T_L \le T) &= P(m_X(T) \le \log L) \\
&= N\left(\frac{\log L - \log S_0 - \mu T}{\sigma\sqrt{T}}\right) + \exp\left(\frac{2\mu(\log L - \log S_0)}{\sigma^2}\right) N\left(\frac{\log L - \log S_0 + \mu T}{\sigma\sqrt{T}}\right) \\
&= N\left(\frac{-\log\frac{S_0}{L} - \mu T}{\sigma\sqrt{T}}\right) + \left(\frac{L}{S_0}\right)^{\frac{2\mu}{\sigma^2}} N\left(\frac{-\log\frac{S_0}{L} + \mu T}{\sigma\sqrt{T}}\right).
\end{aligned}
$$

Hence we have

$$
V = e^{-rT}N\left(\frac{-\log\frac{S_0}{L} - \left(r-q-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right) + e^{-rT}\left(\frac{L}{S_0}\right)^{\frac{2(r-q)}{\sigma^2}-1} N\left(\frac{-\log\frac{S_0}{L} + \left(r-q-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right).
$$

Note that this instrument is down-and-in-bond. Thus V is the same to the down-and-in-bond price. □

✎ One-Touch Asset Option

The one-touch asset option holder should exercise immediately the level(H) is reached for the first time if the dividend rate $q > 0$. If $q = 0$, this option is equivalent to up-and-in(or down-and-in) asset.

When the holder exercises the option, he receives one asset. It is equivalent to H. Hence this option's value is H times one-touch option.

1. Up: $S_0 < H$

$$H\left\{\left(\frac{S_0}{H}\right)^{\frac{\sqrt{\mu^2+2r\sigma^2}-\mu}{\sigma^2}} N\left(\frac{\log\frac{S_0}{H}+\sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right)+\left(\frac{S_0}{H}\right)^{\frac{-\sqrt{\mu^2+2r\sigma^2}-\mu}{\sigma^2}} N\left(\frac{\log\frac{S_0}{H}-\sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right)\right\}.$$

2. Down: $S_0 > L$

$$L\left\{\left(\frac{S_0}{L}\right)^{\frac{-\sqrt{\mu^2+2r\sigma^2}-\mu}{\sigma^2}} N\left(\frac{-\log\frac{S_0}{L}+\sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right)+\left(\frac{S_0}{L}\right)^{\frac{\sqrt{\mu^2+2r\sigma^2}-\mu}{\sigma^2}} N\left(\frac{-\log\frac{S_0}{L}-\sqrt{\mu^2+2r\sigma^2}T}{\sigma\sqrt{T}}\right)\right\}.$$

9.8 Relative Digital Option

A relative digital option pays one unit if the price of stock 1(denoted by S_1) is higher than the price of stock 2(denoted by S_2). More specifically, a call has the payoff X with

$$X = \mathbf{1}_{\{S_1(T)>\lambda S_2(T)\}}$$

and the corresponding put pays

$$X = \mathbf{1}_{\{S_1(T)<\lambda S_2(T)\}}$$

at maturity T, where $\lambda > 0$ is a given constant.

Assume that $S_1(t)$ and $S_2(t)$ satisfy the following SDE under risk neutral measure:

$$\begin{aligned} dS_1(t) &= S_1(t)\Big((r-q_1)dt+\sigma_1 dW_1(t)\Big), \\ dS_2(t) &= S_2(t)\Big((r-q_2)dt+\sigma_2 dW_2(t)\Big) \end{aligned}$$

where $W_1(t)$ and $W_2(t)$ are two standard Brownian motion with correlation ρ. Then the ratio process $\frac{S_1(t)}{S_2(t)}$ is given by

$$\frac{S_1(t)}{S_2(t)} = \frac{S_1(0)}{S_2(0)}\exp\left((q_2-q_1)t+\frac{1}{2}(\sigma_2^2-\sigma_1^2)t+\sigma_1 W_1(t)-\sigma_2 W_2(t)\right).$$

Hence we can see $\log\frac{S_1(T)}{S_2(T)}$ is normally distributed with mean

$$E\left[\log\frac{S_1(T)}{S_2(T)}\right] = \log\frac{S_1(0)}{S_2(0)}+(q_2-q_1)T+\frac{1}{2}(\sigma_2^2-\sigma_1^2)T$$

and variance $\widehat{\sigma}^2 T$ where

$$\widehat{\sigma}^2 := \sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2.$$

So we can see that

$$\begin{aligned} E\left[\frac{S_1(T)}{S_2(T)}\right] &= \frac{S_1(0)}{S_2(0)}\exp\left\{(q_2-q_1)T+\frac{1}{2}(\sigma_2^2-\sigma_1^2)T+\frac{1}{2}\widehat{\sigma}^2 T\right\} \\ &= \frac{S_1(0)}{S_2(0)}\exp\Big\{(q_2-q_1)T+\sigma_2^2 T-\rho\sigma_1\sigma_2 T\Big\}. \end{aligned}$$

From Theorem 1.10.4 the value of relative digital call option is given by

$$e^{-rT} P\left(\frac{S_1(T)}{S_2(T)} > \lambda\right) \quad = \quad e^{-rT} N(d_2)$$

where

$$\begin{aligned} d_2 &= \frac{\log \frac{S_1(0)}{\lambda S_2(0)} + (q_2 - q_1)T + \sigma_2^2 T - \rho\sigma_1\sigma_2 T - \frac{1}{2}\widehat{\sigma}^2 T}{\sqrt{\sigma_1^2 T + \sigma_2^2 T - 2\rho\sigma_1\sigma_2 T}} \\ &= \frac{\log \frac{S_1(0)}{\lambda S_2(0)} + (q_2 - q_1)T + \frac{1}{2}(\sigma_2^2 - \sigma_1^2)T}{\sqrt{\sigma_1^2 T + \sigma_2^2 T - 2\rho\sigma_1\sigma_2 T}}. \end{aligned} \tag{9.8.1}$$

9.9 Relative Outperformance Options

A relative outperformance option is defined by the payoff

$$\left(\frac{S_1(T)}{S_2(T)} - K\right)^+$$

for a call and

$$\left(K - \frac{S_1(T)}{S_2(T)}\right)^+$$

for a put, with some strike K.

Theorem 1.10.4 gives us the call option value

$$e^{-rT}\left\{E\left[\frac{S_1(T)}{S_2(T)}\right] N(d_1) - KN(d_2)\right\}$$

where

$$\begin{aligned} d_1 &:= \frac{\log \frac{E\left[\frac{S_1(T)}{S_2(T)}\right]}{K} + \frac{1}{2}\widehat{\sigma}^2 T}{\widehat{\sigma}\sqrt{T}} \\ &= \frac{\log \frac{S_1(0)}{KS_2(0)} + (q_2 - q_1)T + \frac{1}{2}(\sigma_2^2 - \sigma_1^2)T + \widehat{\sigma}^2 T}{\widehat{\sigma}\sqrt{T}}, \\ d_2 &:= d_1 - \widehat{\sigma}\sqrt{T}. \end{aligned}$$

In fact d_2 is given in (9.8.1). For a detail, see Brockhaus et al.(1999).

Also we can see that the put option's value is given by

$$e^{-rT}\left\{KN(-d_2) - E\left[\frac{S_1(T)}{S_2(T)}\right] N(-d_1)\right\}.$$

9.10 Digital Options on Best or Worst of Two Assets

✎ **Digital Call on Best of Two Assets**

The payoff of a digital call on the best of two asset S_1 and S_2 with strike K and maturity T is given by

$$X = \mathbf{1}_{\{\max(S_1(T),S_2(T))\geq K\}}.$$

Hence the value V can be expressed as

$$\begin{aligned} e^{rT}V &= P(S_1(T)>S_2(T),S_1(T)\geq K)+P(S_1(T)\leq S_2(T),S_2(T)\geq K) \\ &= P\left(\log\frac{S_1(T)}{S_2(T)}>0,\log S_1(T)\geq \log K\right)+P\left(\log\frac{S_1(T)}{S_2(T)}\leq 0,\log S_2(T)\geq\log K\right). \end{aligned}$$

However, $\log(S_1(T)/S_2(T))$ and $\log S_1(T)$ are binormally distributed, with means

$$\begin{aligned} \widetilde{M} &:= E\left[\log\frac{S_1(T)}{S_2(T)}\right] &= \log\frac{S_1(0)}{S_2(0)}+\left(r-q_1-\frac{1}{2}\sigma_1^2\right)T-\left(r-q_2-\frac{1}{2}\sigma_2^2\right)T, \\ M_1 &:= E[\log S_1(T)] &= \log S_1(0)+\left(r-q_1-\frac{1}{2}\sigma_1^2\right)T, \end{aligned}$$

variances

$$\begin{aligned} \mathrm{Cov}\left[\log\frac{S_1(T)}{S_2(T)},\log S_1(T)\right]^T &= \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \sigma_1^2T & \rho\sigma_1\sigma_2T \\ \rho\sigma_1\sigma_2T & \sigma_2^2T \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} (\sigma_1^2-2\rho\sigma_1\sigma_2+\sigma_2^2)T & (\sigma_1^2-\rho\sigma_1\sigma_2)T \\ (\sigma_1^2-\rho\sigma_1\sigma_2)T & \sigma_1^2T \end{bmatrix} \\ &=: \begin{bmatrix} \widetilde{\Sigma}^2 & \Psi_1 \\ \Psi_1 & \Sigma_1^2 \end{bmatrix}. \end{aligned}$$

In addition, $\log(S_1(T)/S_2(T))$ and $\log S_2(T)$ are binormally distributed, with means

$$M_2 := E[\log S_2(T)] = \log S_2(0)+\left(r-q_2-\frac{1}{2}\sigma_2^2\right)T,$$

and variances

$$\begin{aligned} \mathrm{Cov}\left[\log\frac{S_1(T)}{S_2(T)},\log S_2(T)\right]^T &= \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \sigma_1^2T & \rho\sigma_1\sigma_2T \\ \rho\sigma_1\sigma_2T & \sigma_2^2T \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} (\sigma_1^2-2\rho\sigma_1\sigma_2+\sigma_2^2)T & -(\sigma_2^2-\rho\sigma_1\sigma_2)T \\ -(\sigma_2^2-\rho\sigma_1\sigma_2)T & \sigma_2^2T \end{bmatrix} \\ &=: \begin{bmatrix} \widetilde{\Sigma}^2 & -\Psi_2 \\ -\Psi_2 & \Sigma_2^2 \end{bmatrix}. \end{aligned}$$

Thus,

$$\begin{aligned} & e^{rT}V \\ = \; & P\left(\log\frac{S_1(T)}{S_2(T)}>0,\log S_1(T)\geq\log K\right)+P\left(\log\frac{S_1(T)}{S_2(T)}\leq 0,\log S_2(T)\geq\log K\right) \end{aligned}$$

$$
\begin{aligned}
&= P\left(\frac{\log\frac{S_1(T)}{S_2(T)}-\widetilde{M}}{\widetilde{\Sigma}} > \frac{-\widetilde{M}}{\widetilde{\Sigma}}, \frac{\log S_1(T)-M_1}{\Sigma_1} \geq \frac{\log K - M_1}{\Sigma_1}\right) \\
&\qquad + P\left(\frac{\log\frac{S_1(T)}{S_2(T)}-\widetilde{M}}{\widetilde{\Sigma}} \leq \frac{-\widetilde{M}}{\widetilde{\Sigma}}, \frac{\log S_2(T)-M_2}{\Sigma_2} \geq \frac{\log K - M_2}{\Sigma_2}\right) \\
&= M\left(\frac{\widetilde{M}}{\widetilde{\Sigma}}, \frac{M_1-\log K}{\Sigma_1}; \frac{\Psi_1}{\widetilde{\Sigma}\Sigma_1}\right) + M\left(-\frac{\widetilde{M}}{\widetilde{\Sigma}}, \frac{M_2-\log K}{\Sigma_2}; \frac{\Psi_2}{\widetilde{\Sigma}\Sigma_2}\right)
\end{aligned}
$$

where $M(a,b;\rho)$ is the cumulative probability in a standardized bivariate normal distribution that the first variable is less than a and the second variable is less than b, when the coefficient of correlation between the variables is ρ. Hence we have

$$
V = e^{-rT}\Big(M(a,b_1;\rho_1) + M(-a,b_2;\rho_2)\Big) \tag{9.10.1}
$$

where

$$
\begin{aligned}
a &:= \frac{\widetilde{M}}{\widetilde{\Sigma}} = \frac{\log\frac{S_1(0)}{S_2(0)} + \left(r-q_1-\frac{1}{2}\sigma_1^2\right)T - \left(r-q_2-\frac{1}{2}\sigma_2^2\right)T}{\sqrt{(\sigma_1^2-2\rho\sigma_1\sigma_2+\sigma_2^2)T}}, \\
b_1 &:= \frac{M_1-\log K}{\Sigma_1} = \frac{\log\frac{S_1(0)}{K} + \left(r-q_1-\frac{1}{2}\sigma_1^2\right)T}{\sigma_1\sqrt{T}}, \\
b_2 &:= \frac{M_2-\log K}{\Sigma_2} = \frac{\log\frac{S_2(0)}{K} + \left(r-q_2-\frac{1}{2}\sigma_2^2\right)T}{\sigma_2\sqrt{T}}, \\
\rho_1 &:= \frac{\Psi_1}{\widetilde{\Sigma}\Sigma_1} = \frac{\sigma_1-\rho\sigma_2}{\sqrt{\sigma_1^2-2\rho\sigma_1\sigma_2+\sigma_2^2}}, \quad \rho_2 := \frac{\Psi_2}{\widetilde{\Sigma}\Sigma_2} = \frac{\sigma_2-\rho\sigma_1}{\sqrt{\sigma_1^2-2\rho\sigma_1\sigma_2+\sigma_2^2}}.
\end{aligned}
$$

✎ **Digital Put on Best of Two Assets**

The payoff of a digital put on the best of two asset S_1 and S_2 with strike K and maturity T is given by

$$
X = \mathbf{1}_{\{\max(S_1(T),S_2(T))\leq K\}}.
$$

Hence the value V can be expressed as

$$
\begin{aligned}
V &= e^{-rT} - e^{-rT}\Big(M(a,b_1;\rho_1) + M(-a,b_2;\rho_2)\Big) \\
&= e^{-rT}\Big(1 - M(a,b_1;\rho_1) - M(-a,b_2;\rho_2)\Big).
\end{aligned}
$$

✎ **Digital Call on Worst of Two Assets**

The price of a digital call on the worst of two, with payoff

$$
X = \mathbf{1}_{\{\min(S_1(T),S_2(T))\geq K\}}
$$

is given by

$$
e^{rT}V = P\Big(S_1(T)\geq K, S_2(T)\geq K\Big)
$$

$$
\begin{aligned}
&= P\Big(\log S_1(T) \geq \log K, \log S_2(T) \geq \log K\Big) \\
&= P\Big(\frac{\log S_1(T) - M_1}{\Sigma_1} \geq \frac{\log K - M_1}{\Sigma_1}, \frac{\log S_2(T) - M_2}{\Sigma_2} \geq \frac{\log K - M_2}{\Sigma_2}\Big) \\
&= M(b_1, b_2; \rho). \qquad (9.10.2)
\end{aligned}
$$

✎ **Digital Put on Worst of Two Assets**

The price of a digital put on the worst of two, with payoff

$$X = \mathbf{1}_{\{\min(S_1(T), S_2(T)) \leq K\}}$$

is given by

$$V = e^{-rT}\Big(1 - M(b_1, b_2; \rho)\Big).$$

9.11 Options on Best or Worst of Two Assets

✎ **Call on Best of Two Assets**

The payoff of call on the maximum of two assets is given by

$$\Big(\max(S_1(T), S_2(T)) - K\Big)^+.$$

First, consider the following probability:

$$
\begin{aligned}
& P(\max(S_1(T), S_2(T)) \geq K) \\
= \;& P(S_1(T) > S_2(T), S_1(T) \geq K) + P(S_1(T) \leq S_2(T), S_2(T) \geq K) \\
= \;& M(a, b_1; \rho_1) + M(-a, b_2; \rho_2) \quad \text{by (9.10.1).} \qquad (9.11.1)
\end{aligned}
$$

Example 9.11.1. Note that this probability can be written as

$$
\begin{aligned}
& P(\max(S_1(T), S_2(T)) \geq K) \\
= \;& 1 - P(S_1(T) < K, S_2(T) < K) \\
= \;& 1 - P\left(\frac{\log S_1(T) - M_1}{\Sigma_1} < \frac{\log K - M_1}{\Sigma_1}, \frac{\log S_2(T) - M_2}{\Sigma_2} < \frac{\log K - M_2}{\Sigma_2}\right) \\
= \;& 1 - M(-b_1, -b_2; \rho).
\end{aligned}
$$

□

The price of this option is given by

$$
\begin{aligned}
& e^{rT} C_{\max}\big(S_1(0), S_2(0), K, T\big) \\
= \;& E\Big[\Big(\max(S_1(T), S_2(T)) - X\Big)^+\Big] \\
= \;& \int_{\{\max(S_1(T), S_2(T)) \geq K\}} \Big(\max(S_1(T), S_2(T)) - X\Big) dP \\
= \;& \int_{\{S_1(T) > S_2(T)\}} S_1(T) dP + \int_{\{S_1(T) \leq S_2(T)\}} S_2(T) dP
\end{aligned}
$$

$$- XP\Big(\max(S_1(T), S_2(T)) \geq K\Big), \tag{9.11.2}$$

where E is the expectation under the standard risk-neutral measure.

The first part of (9.11.2) can be computed as

$$\begin{aligned}
&\int_{\{S_1(T)>S_2(T)\}} S_1(T)dP \\
=\ & \int_{-b_1}^{\infty}\int_{-a}^{\infty} e^{M_1+\Sigma_1 y}\frac{1}{2\pi\sqrt{1-\rho_1^2}}\exp\left(-\frac{1}{2(1-\rho_1^2)}(x^2+y^2-2\rho_1 xy)\right)dxdy \\
=\ & \frac{e^{M_1}}{2\pi\sqrt{1-\rho_1^2}}\int_{-b_1}^{\infty}\int_{-a}^{\infty} e^{\sigma_1\sqrt{T}y}\exp\left(-\frac{1}{2(1-\rho_1^2)}(x^2+y^2-2\rho_1 xy)\right)dxdy \\
=\ & \frac{S_1(0)e^{(r-q_1)T}}{2\pi\sqrt{1-\rho_1^2}}\int_{-b_1}^{\infty}\int_{-a}^{\infty}\exp\left(-\frac{1}{2(1-\rho_1^2)}\big((x-\rho_1\sigma_1\sqrt{T})^2+(y-\sigma_1\sqrt{T})^2-2\rho_1(x-\rho_1\sigma_1\sqrt{T})(y-\sigma_1\sqrt{T})\big)\right)dxdy \\
=\ & \frac{S_1(0)e^{(r-q_1)T}}{2\pi\sqrt{1-\rho_1^2}}\int_{-b_1-\sigma_1\sqrt{T}}^{\infty}\int_{-a-\rho_1\sigma_1\sqrt{T}}^{\infty}\exp\left(-\frac{1}{2(1-\rho_1^2)}(x^2+y^2-2\rho_1 xy)\right)dxdy \\
=\ & S_1(0)e^{(r-q_1)T}M\Big(b_1+\sigma_1\sqrt{T}, a+\rho_1\sigma_1\sqrt{T};\rho_1\Big).
\end{aligned} \tag{9.11.3}$$

Similarly the second part of (9.11.2) can be computed as

$$\begin{aligned}
&\int_{\{S_1(T)\leq S_2(T)\}} S_2(T)dP \\
=\ & \int_{-b_2}^{\infty}\int_{-\infty}^{-a} e^{M_2+\Sigma_2 y}\frac{1}{2\pi\sqrt{1-\rho_2^2}}\exp\left(-\frac{1}{2(1-\rho_2^2)}(x^2+y^2+2\rho_2 xy)\right)dxdy \\
=\ & \frac{e^{M_2}}{2\pi\sqrt{1-\rho_2^2}}\int_{-b_2}^{\infty}\int_{-\infty}^{-a} e^{\sigma_2\sqrt{T}y}\exp\left(-\frac{1}{2(1-\rho_2^2)}(x^2+y^2+2\rho_2 xy)\right)dxdy \\
=\ & \frac{S_2(0)e^{(r-q_2)T}}{2\pi\sqrt{1-\rho_2^2}}\int_{-b_2}^{\infty}\int_{-\infty}^{-a}\exp\left(-\frac{1}{2(1-\rho_2^2)}\big((x+\rho_2\sigma_2\sqrt{T})^2+(y-\sigma_2\sqrt{T})^2+2\rho_2(x-\rho_2\sigma_2\sqrt{T})(y-\sigma_2\sqrt{T})\big)\right)dxdy \\
=\ & \frac{S_2(0)e^{(r-q_2)T}}{2\pi\sqrt{1-\rho_2^2}}\int_{-b_2-\sigma_2\sqrt{T}}^{\infty}\int_{-\infty}^{-a+\rho_2\sigma_2\sqrt{T}}\exp\left(-\frac{1}{2(1-\rho_2^2)}(x^2+y^2+2\rho_2 xy)\right)dxdy \\
=\ & S_2(0)e^{(r-q_2)T}M\Big(b_2+\sigma_2\sqrt{T}, -a+\rho_2\sigma_2\sqrt{T};\rho_2\Big).
\end{aligned} \tag{9.11.4}$$

Note that the correlation of $\log S_1(T)/S_2(T)$ and $\log S_2(T)$ is $-\rho_2$.

From (9.11.1), (9.11.3) and (9.11.4) we have

$$\begin{aligned}
C_{\max}\big(S_1(0), S_2(0), K, T\big) &= S_1(0)e^{-q_1 T}M\Big(b_1+\sigma_1\sqrt{T}, a+\rho_1\sigma_1\sqrt{T};\rho_1\Big) \\
&\quad + S_2(0)e^{-q_2 T}M\Big(b_2+\sigma_2\sqrt{T}, -a+\rho_2\sigma_2\sqrt{T};\rho_2\Big) \\
&\quad - Ke^{-rT}\Big(M(a, b_1;\rho_1)+M(-a, b_2;\rho_2)\Big) \\
&= S_1(0)e^{-q_1 T}M\Big(b_1+\sigma_1\sqrt{T}, a+\rho_1\sigma_1\sqrt{T};\rho_1\Big) \\
&\quad + S_2(0)e^{-q_2 T}M\Big(b_2+\sigma_2\sqrt{T}, -a+\rho_2\sigma_2\sqrt{T};\rho_2\Big) \\
&\quad - Ke^{-rT}\Big(1-M(-b_1, -b_2;\rho)\Big)
\end{aligned}$$

✎ Put on Best of Two Assets

The payoff of put on the maximum of two assets is given by

$$\Big(K - \max(S_1(T), S_2(T))\Big)^+.$$

By the put-call parity we have

$$P_{\max}\big(S_1(0), S_2(0), K, T\big) = Ke^{-rT} + C_{\max}\big(S_1(0), S_2(0), K, T\big) - e^{-rT}E\big[\max(S_1(T), S_2(T))\big].$$

The expectation of maximum of two asset is given by

$$\begin{aligned} e^{-rT}E\big[\max(S_1(T), S_2(T))\big] &= C_{\max}\big(S_1(0), S_2(0), 0, T\big) \\ &= \lim_{K\to 0} C_{\max}\big(S_1(0), S_2(0), K, T\big) \\ &= S_1(0)e^{-q_1T}N(a + \rho_1\sigma_1\sqrt{T}) + S_2(0)e^{-q_2T}N(-a + \rho_2\sigma_2\sqrt{T}). \end{aligned}$$

✎ Call on worst of Two Assets

The payoff of call on the maximum of two assets is given by

$$\Big(\min(S_1(T), S_2(T)) - K\Big)^+.$$

By (9.10.2) we have

$$P\Big(\min(S_1(T), S_2(T)) \geq K\Big) = M(b_1, b_2; \rho). \tag{9.11.5}$$

The price of this option is given by

$$\begin{aligned} & e^{rT}C_{\min}\big(S_1(0), S_2(0), K, T\big) \\ =\ & E\Big[\Big(\min(S_1(T), S_2(T)) - X\Big)^+\Big] \\ =\ & \int_{\{\min(S_1(T),S_2(T))\geq K\}} \Big(\min(S_1(T), S_2(T)) - X\Big)dP \\ =\ & \int_{\{S_1(T)>S_2(T)\}} S_2(T)dP + \int_{\{S_1(T)\leq S_2(T)\}} S_1(T)dP \\ & - XP\Big(\min(S_1(T), S_2(T)) \geq K\Big), \end{aligned} \tag{9.11.6}$$

The first part of (9.11.6) can be computed as

$$\begin{aligned} & \int_{\{S_1(T)>S_2(T)\}} S_2(T)dP \\ =\ & \int_{-b_2}^{\infty}\int_{-a}^{\infty} e^{M_2+\Sigma_2 y}\frac{1}{2\pi\sqrt{1-\rho_2^2}}\exp\left(-\frac{1}{2(1-\rho_2^2)}(x^2+y^2+2\rho_2 xy)\right)dxdy \\ =\ & \frac{e^{M_2}}{2\pi\sqrt{1-\rho_2^2}}\int_{-b_2}^{\infty}\int_{-a}^{\infty} e^{\sigma_2\sqrt{T}y}\exp\left(-\frac{1}{2(1-\rho_2^2)}(x^2+y^2+2\rho_2 xy)\right)dxdy \\ =\ & \frac{S_2(0)e^{(r-q_2)T}}{2\pi\sqrt{1-\rho_2^2}}\int_{-b_2}^{\infty}\int_{-a}^{\infty}\exp\left(-\frac{1}{2(1-\rho_2^2)}\big((x+\rho_2\sigma_2\sqrt{T})^2+(y-\sigma_2\sqrt{T})^2+2\rho_2(x-\rho_2\sigma_2\sqrt{T})(y-\sigma_2\sqrt{T})\big)\right)dxdy \end{aligned}$$

$$
\begin{aligned}
&= \frac{S_2(0)e^{(r-q_2)T}}{2\pi\sqrt{1-\rho_2^2}} \int_{-b_2-\sigma_2\sqrt{T}}^{\infty} \int_{-a+\rho_2\sigma_2\sqrt{T}}^{\infty} \exp\left(-\frac{1}{2(1-\rho_2^2)}(x^2+y^2+2\rho_2 xy)\right) dxdy \\
&= S_2(0)e^{(r-q_2)T} M\Big(b_2+\sigma_2\sqrt{T}, a-\rho_2\sigma_2\sqrt{T}; -\rho_2\Big). \qquad (9.11.7)
\end{aligned}
$$

Similarly we can calculate the second part of (9.11.6) as

$$
\begin{aligned}
&\int_{\{S_1(T)\le S_2(T)\}} S_1(T)dP \\
&= \int_{-b_1}^{\infty}\int_{-\infty}^{-a} e^{M_1+\Sigma_1 y} \frac{1}{2\pi\sqrt{1-\rho_1^2}} \exp\left(-\frac{1}{2(1-\rho_1^2)}(x^2+y^2-2\rho_1 xy)\right) dxdy \\
&= \frac{e^{M_1}}{2\pi\sqrt{1-\rho_1^2}} \int_{-b_1}^{\infty}\int_{-\infty}^{-a} e^{\sigma_1\sqrt{T}y} \exp\left(-\frac{1}{2(1-\rho_1^2)}(x^2+y^2-2\rho_1 xy)\right) dxdy \\
&= \frac{S_1(0)e^{(r-q_1)T}}{2\pi\sqrt{1-\rho_1^2}} \int_{-b_1}^{\infty}\int_{-\infty}^{-a} \exp\left(-\frac{1}{2(1-\rho_1^2)}\big((x-\rho_1\sigma_1\sqrt{T})^2+(y-\sigma_1\sqrt{T})^2-2\rho_1(x-\rho_1\sigma_1\sqrt{T})(y-\sigma_1\sqrt{T})\big)\right) dxdy \\
&= \frac{S_1(0)e^{(r-q_1)T}}{2\pi\sqrt{1-\rho_1^2}} \int_{-b_1-\sigma_1\sqrt{T}}^{\infty}\int_{-\infty}^{-a-\rho_1\sigma_1\sqrt{T}} \exp\left(-\frac{1}{2(1-\rho_1^2)}(x^2+y^2-2\rho_1 xy)\right) dxdy \\
&= S_1(0)e^{(r-q_1)T} M\Big(b_1+\sigma_1\sqrt{T}, -a-\rho_1\sigma_1\sqrt{T}; -\rho_1\Big). \qquad (9.11.8)
\end{aligned}
$$

Hence (9.11.5), (9.11.7) and (9.11.8) gives us

$$
\begin{aligned}
C_{\min}\big(S_1(0),S_2(0),K,T\big) &= S_1(0)e^{-q_1T} M\Big(b_1+\sigma_1\sqrt{T}, -a-\rho_1\sigma_1\sqrt{T}; -\rho_1\Big) \\
&\quad + S_2(0)e^{-q_2T} M\Big(b_2+\sigma_2\sqrt{T}, a-\rho_2\sigma_2\sqrt{T}; -\rho_2\Big) \\
&\quad - Ke^{-rT} M(b_1,b_2;\rho).
\end{aligned}
$$

✎ **Put on Worst of Two Assets**

The payoff of put on the minimum of two assets is given by

$$
\Big(K-\min(S_1(T),S_2(T))\Big)^+.
$$

By the put-call parity we have

$$
P_{\min}\big(S_1(0),S_2(0),K,T\big) = Ke^{-rT} + C_{\min}\big(S_1(0),S_2(0),K,T\big) - e^{-rT}E\big[\min(S_1(T),S_2(T))\big].
$$

The expectation of min of two asset is given by

$$
\begin{aligned}
e^{-rT}E\big[\min(S_1(T),S_2(T))\big] &= C_{\min}\big(S_1(0),S_2(0),0,T\big) \\
&= \lim_{K\to 0} C_{\min}\big(S_1(0),S_2(0),K,T\big) \\
&= S_1(0)e^{-q_1T}N(-a-\rho_1\sigma_1\sqrt{T}) + S_2(0)e^{-q_2T}N(a-\rho_2\sigma_2\sqrt{T}).
\end{aligned}
$$

Part IV

Interest Models

Chapter 10

Interest Rate Models

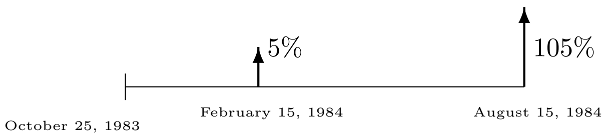
10.1 Basics

We calculate the price we pay for the bond(invoice price), including both the quoted price and accrued interest.

$$\boxed{\text{invoice price} = \text{quoted price} + \text{accrued interest}}$$

A regular bond pays one-half of its coupon rate times its principal value every six months up to and including the maturity date.

Example 10.1.1. A bond with 10% coupon maturing on August 15, 1984, valued on October 25, 1983:



Example 10.1.2 (Accrued Interest). Accrued interest on a Treasury bond is computed using an actual-over-actual day count method. Consider, for example, a 12% bond paying a 6% coupon on May 15 and November 15 each year.
The accrued interest to September 15 is

$$\frac{123}{184} \times 6\% = 4.0109\%.$$

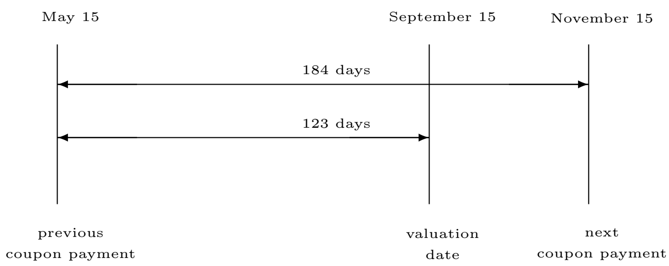
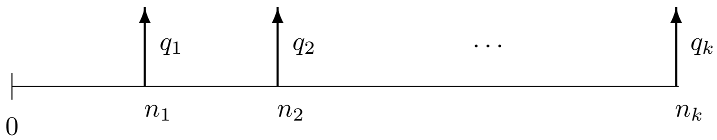


Figure 10.1: Calculating accrued interest.

10.2 Constant Interest Rate



For convenience, we use r to denote the interest per period.

✎ **Present Value:**

$$\begin{aligned}
P &= \frac{q_1}{(1+r)^{n_1}} + \cdots + \frac{q_k}{(1+r)^{n_k}} \\
&= \sum_{i=1}^{k} q_i (1+r)^{-n_i}. \\
\frac{dP}{dr} &= -\sum_{i=1}^{k} n_i q_i (1+r)^{-n_i-1}. \\
\frac{dP}{P} &= -\frac{1}{1+r} \frac{\displaystyle\sum_{i=1}^{k} n_i q_i (1+r)^{-n_i}}{\displaystyle\sum_{i=1}^{k} q_i (1+r)^{-n_i}} dr \\
&= -\frac{D}{1+r} dr,
\end{aligned}$$

where

$$\begin{aligned}
D &:= \frac{\sum_{i=1}^{k} n_i q_i (1+r)^{-n_i}}{\sum_{i=1}^{k} q_i (1+r)^{-n_i}} \\
&= \text{Weighted average of } n_i \text{ with the weighting factor } q_i(1+r)^{-n_i}.
\end{aligned}$$

D is referred to as the duration.

$$\frac{dP}{dr} = -\frac{PD}{1+r},$$

$$
\begin{aligned}
\frac{d^2P}{dr^2} &= \frac{PD}{(1+r)^2} - \frac{1}{1+r}\Big(\frac{dD}{dr}P + \frac{dP}{dr}D\Big) \\
&= \frac{PD}{(1+r)^2} - \frac{1}{1+r}\Big(-\frac{C}{1+r}P - \frac{PD}{1+r}D\Big), \quad \text{by (10.2.1)} \\
&= \frac{C+D+D^2}{(1+r)^2}\,P. \\
\frac{\frac{\partial^2 P}{\partial r^2}}{P} &= \frac{C+D+D^2}{(1+r)^2} \quad =: \quad C_g (\text{the general definition of convexity}).
\end{aligned}
$$

It is approximately true that

$$
\Delta P = -P\frac{D}{1+r}\Delta r = -PD^*\Delta r, \quad \text{where } D^* \text{is the modified duration.}
$$

The preceding analysis is based on the assumption that r is expressed with annual compounding. If r is expressed with continuous compounding, a similar analysis to that given above shows that

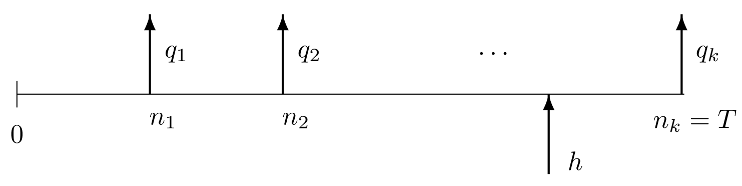
$$
\frac{dP}{dr} = -DP.
$$

More generally, if r is expressed with a compounding frequency of m times per year,

$$
\begin{aligned}
\frac{dP}{dr} &= -\frac{PD}{1+r/m}, \\
D^* &= \frac{D}{1+r/m}.
\end{aligned}
$$

The duration of a bond portfolio can be defined as a weighted average of the durations of the individual bonds in the portfolio with the weights begin proportional to the bond prices.

✎ **Time horizon(used for ALM):**



Two kind of risks for bond investors

- Reinvestment Risk.
- Market Risk.

The value of W of the bond at the time horizon h is

$$
W = P(1+r)^h
$$

$$= \sum_i q_i(1+r)^{-n_i+h}.$$

$$\begin{aligned}\frac{dW}{dr} &= \frac{dP}{dr}(1+r)^h + Ph(1+r)^{h-1}\\ &= (1+r)^h\Big[\frac{dP}{dr} + \frac{h}{1+r}P\Big]\\ &= (1+r)^h\ \frac{h-D}{1+r}\ P.\end{aligned}$$

1. If $h < D$(long maturity bond), $\frac{dW}{dr} < 0$.
2. If $h > D$(short maturity bond), $\frac{dW}{dr} > 0$.
3. If $h = D$, $\frac{dW}{dr} = 0$(immunization).

The duration D is given by

$$D = \frac{\sum_{i=1}^k n_i q_i(1+r)^{-n_i}}{P}.$$

Now, let us compute $\dfrac{dD}{dr}$:

$$\begin{aligned}\frac{dD}{dr} &= -\frac{1}{P^2}\frac{dP}{dr}DP - \frac{\sum_{i=1}^k n_i^2 q_i(1+r)^{-n_i-1}}{P}\\ &= \frac{1}{P^2}\frac{DP}{1+r}DP - \frac{\sum_{i=1}^k n_i^2 q_i(1+r)^{-n_i-1}}{P}\\ &= \frac{D^2}{1+r} - \frac{\sum_{i=1}^k n_i^2 q_i(1+r)^{-n_i-1}}{P}\\ &= -\frac{\sum_{i=1}^k (n_i^2 - D^2) q_i(1+r)^{-n_i}}{P(1+r)}\\ &= -\frac{\sum_{i=1}^k (n_i - D)^2 q_i(1+r)^{-n_i}}{P(1+r)}\\ &= -\frac{C}{1+r}. \qquad (10.2.1)\end{aligned}$$

where

$$\begin{aligned}C &= \frac{\sum_{i=1}^k (n_i - D)^2 q_i(1+r)^{-n_i}}{P}\\ &= \text{Convexity},\\ &= \text{Weighted variance of } n_i \text{ with weighting factor } q_i(1+r)^{-n_i}.\end{aligned}$$

Remark 10.2.1. Note that this definition of convexity is different from the general definition of convexity by

$$C_g := \frac{\frac{\partial^2 P}{\partial r^2}}{P}.$$

Let us calculate $\frac{d^2W}{dr^2}$:

$$\begin{aligned}
\frac{d^2W}{dr^2} &= (h-1)(1+r)^{h-2}(h-D)P + (1+r)^{h-1}\Big(-\frac{dD}{dr}P + (h-D)\frac{dP}{dr}\Big) \\
&= (h-1)(1+r)^{h-2}(h-D)P + (1+r)^{h-1}\Big(\frac{C}{1+r}P - (h-D)\frac{D}{1+r}P\Big) \\
&= (1+r)^{h-2}\Big((h-1)(h-D) + CP - (h-D)D\Big)P \\
&= (1+r)^{h-2}\Big((h-D)^2 - (h-D) + C\Big)P
\end{aligned}$$

✎ **Rate of return:**

1. (theoretic) bond price P with respect to r.

$$P(r) = \sum_{i=1}^{k} q_i(1+r)^{-n_i}.$$

2. Yield to maturity(YTM, $\bar{r}$): Let $\bar{P}$ be the market price of the bond. The yield to maturity of the bond is define implicitly in equation

$$\begin{aligned}
\bar{P} &= P(\bar{r}), \quad \text{i.e.} \\
\bar{P} &= \sum_{i=1}^{k} q_i(1+\bar{r})^{-n_i}.
\end{aligned}$$

3. Define **future value** FV by

$$\mathrm{FV}(h,r) = (1+r)^h \sum_{i=1}^{k} q_i(1+r)^{-n_i} =: W(r).$$

Note that

$$W'(r) = (1+r)^h \, \frac{h-D}{1+r} \, P.$$

4. (average) rate of return: If the bond is purchased at a price of $\bar{P}$, the average **rate of return**(R) on the bond over the time interval h is the value of R that satisfies the equation

$$\bar{P} = \frac{\mathrm{FV}}{(1+R)^h}. \tag{10.2.2}$$

Since FV depends on r, the value of R defined implicitly in equation (10.2.2) will also depend on r, i.e. $R(r)$. Note that $R(\bar{r}) = \bar{r}$.

To get a second order approximation to the rate of return about $\bar{r}$, we need $R'(\bar{r})$ and $R''(\bar{r})$.

$$W(r) = \bar{P}\,\Big(1+R(r)\Big)^h,$$

$$
\begin{aligned}
W'(r) &= \bar{P}\ h \left(1 + R(r)\right)^{h-1} R'(r), \\
R'(r) &= \frac{1}{\bar{P}\ h \left(1 + R(r)\right)^{h-1}} \times (1+r)^h\ \frac{h-D}{1+r}\ P, \\
R'(\bar{r}) &= 1 - \frac{D(\bar{r})}{h}.
\end{aligned}
$$

After a tedious calculation, we get

$$
R''(\bar{r}) = \frac{C - D + \frac{D^2}{h}}{h(1+\bar{r})} = \frac{C - \frac{1}{h}(h-D)D}{h(1+\bar{r})}
$$

$$
\begin{aligned}
R(r) &= R(\bar{r}) + R'(\bar{r})(r-\bar{r}) + \frac{1}{2}R''(\bar{r})(r-\bar{r})^2 + \cdots \\
&= \bar{r} + \left(1 - \frac{D}{h}\right)(r-\bar{r}) + \frac{1}{2}\left(\frac{C - \frac{1}{h}(h-D)D}{h(1+\bar{r})}\right)(r-\bar{r})^2 + \cdots
\end{aligned}
$$

If $D(\bar{r}) = h$,

$$
R(r) = \bar{r} + \frac{1}{2}\ \frac{C}{h(1+\bar{r})}\ (r-\bar{r})^2.
$$

Remark 10.2.2 (Continuous Compounding). Let m be a compounding frequency per year.

Discrete compounding	Continuous compounding
$P = \sum_{i=1}^{k} q_i (1+r/m)^{-mn_i}$	$P = \sum_{i=1}^{k} q_i e^{-rn_i}$
$D = \sum_{i=1}^{k} n_i \frac{q_i(1+r/m)^{-mn_i}}{P}$	$D = \sum_{i=1}^{k} n_i \frac{q_i e^{-rn_i}}{P}$
$\frac{dP}{dr} = -\sum_{i=1}^{k} n_i q_i (1+r/m)^{-mn_i-1} = -\frac{PD}{1+r/m}$	$\frac{dP}{dr} = -\sum_{i=1}^{k} n_i q_i e^{-rn_i} = -PD$

Table 10.1: Comparision between Compoundings

Remark 10.2.3. When convexity is taken into account, it can be shown that

$$
P(r+\Delta r) \approx P(r) + \frac{dP}{dr}\Delta r + \frac{1}{2}\frac{d^2P}{dr^2}\Delta r^2
$$

$$
\begin{aligned}
&= P(r) - \frac{PD}{1+r/m}\Delta r + \frac{1}{2}PC_g(\Delta r)^2, \\
\frac{\Delta P}{P} &\approx -\frac{D}{1+r/m}\Delta r + \frac{1}{2}C_g \Delta r^2.
\end{aligned}
$$

10.3 Principal Components of Interest Rates

In this section, we give the results of a principal component analysis about interest rate.

date	tcm1m	tcm3m	tcm6m	tcm1y	tcm2y	tcm3y	tcm5y	tcm7y	tcm10y	tcm20y
2001-07-31	3.67	3.54	3.47	3.53	3.79	4.06	4.57	4.86	5.07	5.61
2001-08-01	3.65	3.53	3.47	3.56	3.83	4.09	4.62	4.9	5.11	5.63
2001-08-02	3.65	3.53	3.46	3.57	3.89	4.17	4.69	4.97	5.17	5.68
2001-08-03	3.63	3.52	3.47	3.57	3.91	4.22	4.72	4.99	5.2	5.7
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 10.2: Interest data

date	tcm1m	tcm3m	tcm6m	tcm1y	tcm2y	tcm3y	tcm5y	tcm7y	tcm10y	tcm20y
2001-08-01	-0.02	-0.01	0	0.03	0.04	0.03	0.05	0.04	0.04	0.02
2001-08-02	0	0	-0.01	0.01	0.06	0.08	0.07	0.07	0.06	0.05
2001-08-03	-0.02	-0.01	0.01	0	0.02	0.05	0.03	0.02	0.03	0.02
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 10.3: Daily differences of interest rates

eigenvector	Prin 1	Prin 2	Prin 3	Prin 4	Prin 5
tcm1m	0.165	0.554	0.644	-0.473	0.108
tcm3m	0.229	0.532	0.03	0.531	-0.383
tcm6m	0.294	0.381	-0.345	0.319	0.25
tcm1y	0.337	0.137	-0.405	-0.27	0.588
tcm2y	0.355	-0.048	-0.261	-0.366	-0.314
tcm3y	0.357	-0.098	-0.159	-0.266	-0.421
tcm5y	0.357	-0.166	0.07	0.003	-0.204
tcm7y	0.349	-0.223	0.173	0.081	-0.019
tcm10y	0.341	-0.251	0.241	0.179	0.108
tcm20y	0.317	-0.306	0.338	0.277	0.323

Table 10.4: Principal Components of difference of interest rate

1. The first principal component is made up by approximately equal weights of the original variables, and can therefore be intuitively interpreted as the **average level** of the yield curve.
2. The second is made up by weight of similar magnitude and opposite signs at the opposite end of the maturity spectrum, and therefore lends itself to the interpretation of being the **slope** of the yield curve.
3. The third is made up by weight of similar magnitude and identical signs at the both end of the maturity spectrum, and approximately twice as large and of opposite sign in the middle.

	Eigenvalue	Proportion	Cumulative
Prin 1	7.20597463	0.7206	0.7206
Prin 2	1.86924183	0.1869	0.9075
Prin 3	0.43191883	0.0432	0.9507
Prin 4	0.21611144	0.0216	0.9723
Prin 5	0.13927787	0.0139	0.9863
Prin 6	0.04821816	0.0048	0.9911
Prin 7	0.04035971	0.004	0.9951
Prin 8	0.01922379	0.0019	0.997
Prin 9	0.01677408	0.0017	0.9987
Prin 10	0.01289966	0.0013	1

Table 10.5: Eigenvalues of the Correlation Matrix

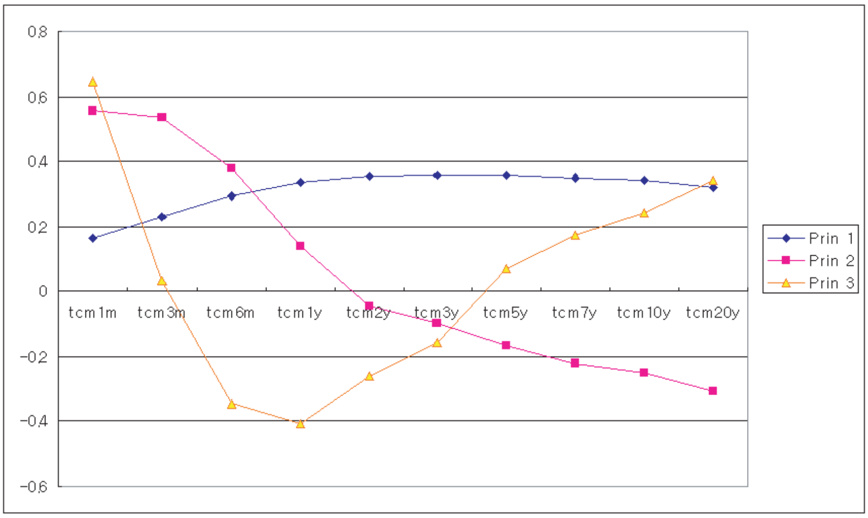


Figure 10.2: Principal component of interest(spot rate)

This feature warrants the interpretation of the third component as the **curvature** of the yield curve.

As for the explanatory power of these new variables, in most currencies one then finds(see Table 10.5) that the level often accounts for up to $80-90\%$ of the total variance, and that the first three principal components taken together often describe up to $90-99\%$ of the inter-maturity variability.

A general principal components analysis is explained in Appendix H.

10.4 Interest Rate Derivatives

Interest rate derivatives are instruments whose payoff are dependent in some way on the level of interest rates. A key challenge for derivatives traders is to find good, robust procedures for *pricing* and *hedging* these products.

Interest rate derivatives are more difficult to value than equity and foreign exchange derivatives. There are a number of reasons for this:

- The behavior of an individual interest rate is more complicated than that of a stock price or exchange rate.
- For the valuation of many products, it is necessary to develop a model describing the behavior of the entire yield curve.
- The volatilities of different points on the yield curve are different.
- Interest rates are used for discounting as well as for defining the payoff from the derivatives.

10.4.1 Black's Model

Consider a European call option on a variable whose value is V. Define

1. T: Maturity date of the option.
2. F: Forward price of V for a contract with maturity T.
3. F_0: Value of F at time zero.
4. X: Strike price of the option.
5. $P(t,T)$: Price at time t of a zero-coupon bond paying 1 at time T.
6. σ: Volatility of F

Black's model calculates the expected payoff from the option assuming

1. V_T has a lognormal distribution with the standard deviation of $\log V_T$ equal to $\sigma\sqrt{T}$.
2. the expected value of V_T is F_0.

An important feature of Black's model is that we do not have to assume geometric Brownian motion for the evolution of either V or F. All that we require is that V_T be lognormal at time T The parameter, σ is to define the standard deviation of $\log V_T$ by means of the relationship

$$\text{Standard Deviation of } \log V_T = \sigma\sqrt{T}.$$

The payoff from the option is

$$\max(V_T - X, 0)$$

at time T. By Theorem(1.10.4), the lognormal assumption implies that the expected payoff is

$$E[V_T]N(d_1) - XN(d_2),$$

where

$$\begin{aligned} d_1 &= \frac{\log \frac{E[V_T]}{X} + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}, \\ d_2 &= d_1 - \sigma\sqrt{T}. \end{aligned}$$

Because we are assuming that $E[V_T] = F_0$, the value of the option is

$$P(0,T)\Big[F_0 N(d_1) - XN(d_2)\Big], \qquad (10.4.1)$$

where

$$\begin{aligned} d_1 &= \frac{\log \frac{F_0}{X} + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}, \\ d_2 &= d_1 - \sigma\sqrt{T}. \end{aligned}$$

(10.4.1) is called the Black formula.

Example 10.4.1. The variable V can be as the followings:

1. Bond price.
2. LIBOR rate.
3. Swap rate.

□

10.4.2 Bond Options

The assumption usually made is that the bond price at the maturity of the option is lognormal. (10.4.1) can be used to price the option with F_0 equal to the forward bond price. The variable σ is defined so that $\sigma\sqrt{T}$ is the standard deviation of the logarithm of the bond price at the maturity of the option.

Since a bond is a security that provides a known cash income, F_0 can be calculated using the formula

$$F_0 = \frac{B_0 - I}{P(0,T)}$$

where B_0 is the bond price at time zero and I is the present value of the coupons will paid during the life of the option. In this formula, both the spot bond price and forward bond price are cash prices(dirty prices) rather than quoted prices(clean prices). The strike price, X, in (10.4.1) should be the cash strike price. If, as is more common, the strike price is the quoted price applicable when the option is exercised, X should be set equal to the strike price plus accrued interest at the expiration date of the option or

$$X = \text{strike price} + \text{accrued interest}.$$

✎ **Yield Volatility v.s. Forward Price Volatility**
The volatilities that are quoted for bond options are often forward yield volatilities rather than forward price volatilities. The relationship between the change in the bond's price, B, and its yield, y, at the maturity of the option is

$$\frac{\Delta B}{B} \approx -D\Delta y$$

or

$$\frac{\Delta B}{B} \approx -Dy\frac{\Delta y}{y}$$

where D is the modified duration of the forward bond underlying the option.

This suggests that the forward price volatility, σ, used in Black's model can be approximately related to the corresponding forward yield volatility, σ_y, using

$$\sigma = D \times y_0 \times \sigma_y$$

where y_0 is the initial forward yield.

10.4.3 Caps And Floors

Consider a cap with a total life T, a principal of L, and a cap rate of R_X. Suppose that the reset dates are $t_1, t_2, \cdots, t_n$ and define $t_{n+1} = T$. Define R_k as the LIBOR rate for the period between t_k and t_{k+1} observed at time $t_k (1 \le k \le n)$. The cap leads to a payoff at time $t_{k+1} (k = 1, 2, \cdots, n)$ of

$$L\delta_k \max(R_k - R_X, 0)$$

where $\delta_k = t_{k+1} - t_k$.

If the LIBOR R_k is assumed to be lognormal with volatility σ_k, (10.4.1) gives the value of caplet as

$$L\delta_k P(0, t_{k+1})\Big[F_k N(d_1) - R_X N(d_2)\Big] \tag{10.4.2}$$

where

$$\begin{aligned} d_1 &= \frac{\log\frac{F_k}{R_X} + \frac{1}{2}\sigma_k^2 t_k}{\sigma_k\sqrt{t_k}}, \\ d_2 &= \frac{\log\frac{F_k}{R_X} - \frac{1}{2}\sigma_k^2 t_k}{\sigma_k\sqrt{t_k}} = d_1 - \sigma_k\sqrt{t_k} \end{aligned}$$

and F_k is the forward rate for the period between time t_k and t_{k+1}.

✎ **Spot Volatility v.s. Flat Volatility**
Each caplet of a cap must be valued separately using (10.4.2).

1. One approach is to use a different volatility for each caplet. The volatilities are then referred to as spot volatilities.
2. An alternative approach is to use the same volatility for all the caplets comprising any particular cap but to vary this volatility according to the maturity of the cap. The volatilities used are then referred to as flat volatilities.

10.4.4 Swap Options

Swap options or swaption are options on interest rate swap(IRS). An IRS can be regarded as an agreement to exchange a fixed-rate bond for a floating-rate bond. At the start of a swap, the value of the floating-rate bond always equals the principal amount of the swap. Hence the value of the fixed-rate bond equals the principal amount of the swap. A swaption can, therefore, be regarded as an option to exchange a fixed-rate bond for the principal amount of the swap.

1. If a swaption gives the holder the right to pay fixed and receive floating, it is a put option on the fixed-rate bond with strike price equal to the principal.

2. If a swaption gives the holder the right to pay floating and receive fixed, it is a call option on the fixed-rate bond with strike price equal to the principal.

We assume that the swap rate at the maturity of the option is lognormal. Consider a swaption where we have the right to pay a rate R_X and receive LIBOR on a swap that will last n years starting in T. We suppose that there are m payments per year under the swap and that the principal L. Suppose that the payment dates are $t_1, t_2, \cdots, t_{mn}$. Using (10.4.1), the value of the cash flow received at time t_i is

$$\frac{L}{m}P(0,t_i)\Big[F_0 N(d_1) - R_X N(d_2)\Big]$$

where

$$\begin{aligned} d_1 &= \frac{\log\frac{F_0}{R_X} + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}, \\ d_2 &= \frac{\log\frac{F_0}{R_X} - \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T} \end{aligned}$$

and F_0 is the forward swap rate.

The total value of the swaption is

$$\frac{L}{m}\Big[F_0 N(d_1) - R_X N(d_2)\Big]\sum_{i=1}^{mn} P(0,t_i).$$

10.5 Convexity Adjustments

The forward yield, $f(t,T,T')$, on a bond is defined as the yield calculated from the forward bond prices. i.e.

$$e^{-f(t,T,T')(T'-T)} = \frac{P(t,T')}{P(t,T)}.$$

Suppose that $B_T\big(=P(T,T')\big)$ is the price of a bond at time T, y_T is its yield, and the relationship between B_T and y_T is

$$B_T = G(y_T),$$

where $G(x) = e^{-x(T'-T)}$. This means that $y_T = f(T, T, T')$. Define F_0 as the forward bond price at time zero for a contract maturing at time T and y_0 as the forward bond yield at time zero. This definition of a forward bond yield means that

$$F_0 = G(y_0).$$

This means that

$$F_0 = \frac{P(0,T')}{P(0,T)}, \quad y_0 = f(0,T,T').$$

The function G is non-linear. This implies that, when the expected future bond price equals the forward bond price (so that we are a world that is T-forward measure), the expected future bond yield does not equal the forward bond yield. i.e. even if

$$F_0 = E_T[B_T], \qquad \left(\text{i.e. } \frac{P(0,T')}{P(0,T)} = E_T[P(T,T')]\right),$$

$$y_0 \neq E_T[y_T], \qquad \left(\text{i.e. } f(0,T,T') \neq E_T[f(T,T,T')]\right).$$

Expanding $G(y_T)$ in a Taylor series about $y_T = y_0$ yields the following approximation

$$B_T = G(y_0) + (y_T - y_0)G'(y_0) + \frac{1}{2}(y_T - y_0)^2 G''(y_0).$$

Taking expectations in a T-forward measure

$$E_T[B_T] = G(y_0) + E_T[y_T - y_0]G'(y_0) + \frac{1}{2}E_T[(y_T - y_0)^2]G''(y_0).$$

Since $E_T[B_T] = F_0 = G(y_0)$ we have

$$E_T[y_T - y_0]G'(y_0) + \frac{1}{2}E_T[(y_T - y_0)^2]G''(y_0) = 0.$$

The expression $E_T[(y_T - y_0)^2]$ is approximately $\sigma_y^2 y_0^2 T$, where σ_y is the volatility of the forward yield y_0. Hence it is approximately true that

$$E_T[y_T] = y_0 - \frac{1}{2}y_0^2\sigma_y^2 T\frac{G''(y_0)}{G'(y_0)}.$$

10.6 Non-Constant Interest

10.6.1 Generalities

1. $P(t,T)$: the price at time t of the zero-coupon bond paying 1 at time T. i.e. $P(T,T) = P(t,t) = 1$.

2. The $(T-t)$ year zero rate is the rate of interest earned on an investment that start t and last for $(T-t)$. All the interest and principal is realized at the time T. There is no intermediate payment. The **spot rate(zero rate)** for the period $[t,T]$ is given by

$$R(t,T) \quad = \quad -\frac{\log P(t,T)}{T-t} \tag{10.6.1}$$

i.e.

$$P(t,T) = e^{-R(t,T)(T-t)}, \quad P(t,T)e^{R(t,T)(T-t)} = 1.$$

$R(t,T)$ is the average yield of the bond over its remaining lifetime. A chart showing the zero rate as a function of maturity is known as zero curve.

3. At time t, consider a portfolio: 1 long position in $P(t,T_2)$ and k short position in $P(t,T_1)$ $(t < T_1 < T_2)$. The value at time t of this position is

$$V_t = P(t,T_2) - kP(t,T_1).$$

To make the value of this contract 0, k must be

$$\frac{P(t,T_2)}{P(t,T_1)}.$$

4. Alternative approach: forward contract

	Stock	Bond
underlying	S_t	$P(t,T_2)$
maturity	T	T_1
forward price at t	$S_t e^{r(T-t)}$	$E^{T_1}\left[P(T_1,T_2)\vert\mathcal{F}_t\right]$

Table 10.6: Stock forward vs bond forward

We can compute $E\left[P(T_1,T_2)\vert\mathcal{F}_t\right]$ as follows:

$$\begin{aligned} 0 &= E_Q\left[\Big(P(T_1,T_2)-k\Big)\exp\left(-\int_t^{T_1} r(s)ds\right)\Big|\mathscr{F}_t\right] \\ &= B_t E_Q\left[\frac{P(T_1,T_2)}{B_{T_1}}\Big|\mathscr{F}_t\right] - kP(t,T_1) \\ &= P(t,T_2) - kP(t,T_1), \end{aligned}$$

where $B_t = \exp\left(\int_0^t r(s)ds\right)$. Furthermore, we can get the following additional results:

$$\begin{aligned}
k &= \frac{1}{P(t,T_1)}\; B_t E_Q\left[\frac{P(T_1,T_2)}{B_{T_1}}\Big|\mathscr{F}_t\right] \\
&= \frac{1}{P(t,T_1)}\; P(t,T_1)E^{T_1}\left[P(T_1,T_2)|\mathcal{F}_t\right] \quad = \quad E^{T_1}\left[P(T_1,T_2)|\mathcal{F}_t\right] \\
&= \frac{1}{P(t,T_1)}\; P(t,T_2)\; E^{T_2}\left[\, 1\; |\mathcal{F}_t\right] \quad = \quad \frac{P(t,T_2)}{P(t,T_1)}.
\end{aligned}$$

If we denote forward price[1] of T_2-bond with maturity T_1 at time t by $F_t(T_1,T_2)$, the above k is $F_t(T_1,T_2)$ and

$$\begin{aligned}
F_t(T_1,T_2) &= E^{T_1}[P(T_1,T_2)|\mathscr{F}_t] \\
&= E^{T_1}[F_{T_1}(T_1,T_2)|\mathscr{F}_t]. \qquad (10.6.2)
\end{aligned}$$

This means that $\{F_t(T_1,T_2)\}$ is a martingale under the forward measure of T_1-bond.

In fact these results come from the martingale property [2].

Now, define the **forward rate** $f(t,T_1,T_2)$ to satisfy the following equation:

$$ke^{f(t,T_1,T_2)(T_2-T_1)} \quad = \quad 1.$$

Thus

$$\begin{aligned}
e^{f(t,T_1,T_2)(T_2-T_1)} &= \frac{1}{k} = \frac{P(t,T_1)}{P(t,T_2)}. \\
f(t,T_1,T_2) &= -\frac{\log P(t,T_2) - \log P(t,T_1)}{T_2-T_1} \qquad (10.6.3) \\
&= \frac{R(t,T_2)(T_2-t) - R(t,T_1)(T_1-t)}{T_2-T_1} \\
&= R(t,T_2) + (T_1-t)\;\frac{R(t,T_2)-R(t,T_1)}{T_2-T_1}, \\
f(t,T) &:= \lim_{T_1\to T} f(t,T_1,T) \\
&= R(t,T) + (T-t)\;\frac{\partial R}{\partial T} \\
&= \frac{\partial}{\partial T}\Big((T-t)R(t,T)\Big),
\end{aligned}$$

[1] *The forward price* is the delivery price in a forward contract that causes the contract to be worth zero. (10.6.2) shows that the forward price is its expected future spot price under T_1-forward measure. *The futures price* is the expected future spot price under the traditional risk-neutral measure.

[2] Note that

$$\begin{aligned}
\frac{P(t,T_2)}{B_t} &= E\left[\frac{P(T_1,T_2)}{B_{T_1}}\Big|\mathscr{F}_t\right], \\
\frac{P(t,T_2)}{P(t,T_1)} &= E^{T_1}\left[P(T_1,T_2)\Big|\mathscr{F}_t\right].
\end{aligned}$$

$$\begin{aligned}\int_t^T f(t,s)ds &= (T-t)R(t,T),\\ R(t,T) &= \frac{1}{T-t}\int_t^T f(t,s)ds.\end{aligned}$$

These equations means that the forward rate is defined as the yield calculated from the forward bond price

$$\frac{P(t,T_2)}{P(t,T_1)}.$$

i.e.

$$\frac{P(t,T_2)}{P(t,T_1)} = e^{-f(t,T_1,T_2)(T_2-T_1)}.$$

From Equation (10.6.3),

$$f(t,T) = -\frac{\partial}{\partial T}\log P(t,T). \tag{10.6.4}$$

Since $P(t,t) = 1$,

$$\begin{aligned}P(t,T) &= \exp\Big(-\int_t^T f(t,s)ds\Big), \quad \text{also}, \qquad (10.6.5)\\ &= \exp\Big(-(T-t)R(t,T)\Big).\end{aligned}$$

Example 10.6.1 (Bootstrap Method). We now discuss how zero-coupon interest rates can be calculated from the prices of coupon-bearing instruments. One approach is known as the bootstrap method.

내용 채워야 할 곳...

□

Example 10.6.2 (Yield To Maturity). The *yield to maturity* or *internal rate of return* on a coupon-bearing bond is the discount rate that equates the cash flows on the bond to its market value.

Example 10.6.3 (Par Yield). The *par yield* for a certain maturity is the coupon rate that causes the bond price to equal its face value. Usually the bond is assumed to provide semiannual coupons.

10.6.2 Swaps

For given $T_0, T_1, \cdots, T_n$ with constant $\delta = T_{i+1} - T_i$, the contract is made at $t(\leq T_0)$.

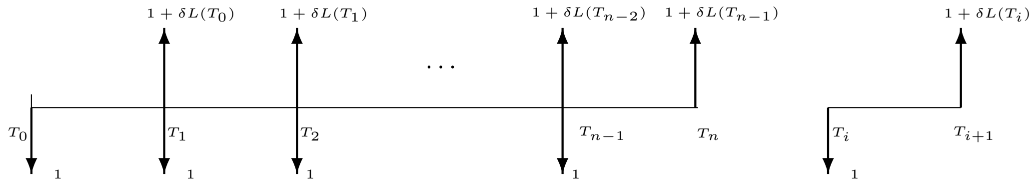


Figure 10.3: Floating Rate 1

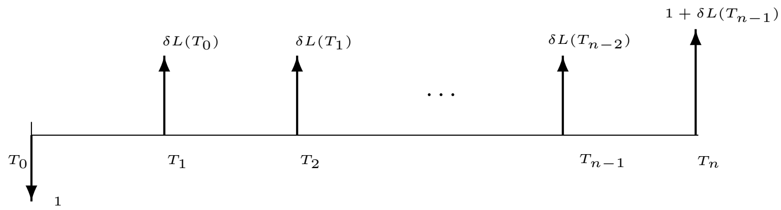


Figure 10.4: Floating Rate 2

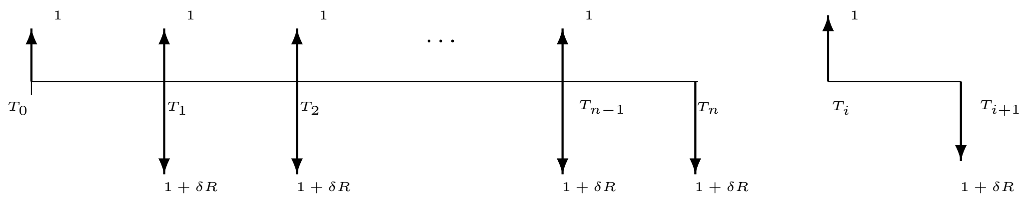


Figure 10.5: Fixed Rate 1

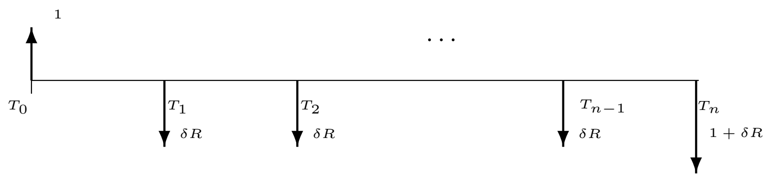


Figure 10.6: Fixed Rate 2

Swap rate $L(T_i)$ is determined at T_i and it will be working over the period $[T_i, T_{i+1}]$. At T_i, the value of cash flow $1 + \delta L(T_i)$ of T_{i+1} should be 1.

$$
\begin{aligned}
P(T_i, T_{i+1})\big(1 + \delta L(T_i)\big) &= 1, \qquad (10.6.6)\\
P(T_i, T_{i+1}) &= \frac{1}{1 + \delta L(T_i)}.
\end{aligned}
$$

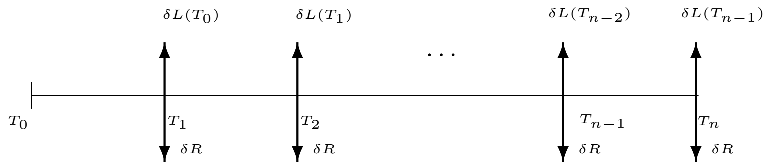


Figure 10.7: Swap

From Figure 10.6, we can calculate the forward swap rate [3] R:

$$P(t,T_0) - \delta R \sum_{i=1}^{n} P(t,T_i) - P(t,T_n) = 0, \tag{10.6.7}$$

$$R(t) := \frac{1}{\delta}\frac{P(t,T_0) - P(t,T_n)}{\sum_{i=1}^{n} P(t,T_i)}, \quad t \le T_0 \tag{10.6.8}$$

In particular, if $t = T_0$ and R is given by (10.6.8), we have

$$\text{the value of the floating-rate bond} = 1 = \text{the value of the fixed-rate bond.}$$

Example 10.6.4 (Discounted Expectation of Floating Rate). Let us calculate the discounted expectation of floating rate $\delta L(T_i)$ payed at time T_{i+1}. From (10.6.6) we have

$$\delta L(T_i) = \frac{1}{P(T_i,T_{i+1})} - 1.$$

First, note that

$$P(T_i,T_{i+1}) = B_{T_i} E\left[\frac{1}{B_{T_{i+1}}}\Big|\mathscr{F}_{T_i}\right]. \tag{10.6.9}$$

The expectation can be calculated as follows:

$$\begin{aligned}
& B_{T_0} E\left[\frac{\delta L(T_i)}{B_{T_{i+1}}}\Big|\mathscr{F}_{T_0}\right] \\
= \; & B_{T_0} E\left[\frac{1}{B_{T_{i+1}}}\left(\frac{1}{P(T_i,T_{i+1})} - 1\right)\Big|\mathscr{F}_{T_0}\right] \\
= \; & B_{T_0} E\left[\frac{1}{B_{T_{i+1}}}\frac{1}{P(T_i,T_{i+1})}\Big|\mathscr{F}_{T_0}\right] - P(T_0,T_{i+1}) \\
= \; & B_{T_0} E\left[E\left[\frac{1}{B_{T_{i+1}}}\Big|\mathscr{F}_{T_i}\right]\frac{1}{P(T_i,T_{i+1})}\Big|\mathscr{F}_{T_0}\right] - P(T_0,T_{i+1}) \\
= \; & B_{T_0} E\left[B_{T_i}|\mathscr{F}_{T_0}\right] - P(T_0,T_{i+1}) \quad \text{by (10.6.9)} \\
= \; & P(T_0,T_i) - P(T_0,T_{i+1}).
\end{aligned} \tag{10.6.10}$$

This price also suggest the hedge of selling a T_{i+1}-bond and buying a T_i-bond. When the T_i-bond matures, we buy $\frac{1}{P(T_i,T_{i+1})}$ units of the T_{i+1}-bond, and we left with exactly the right payoff at time T_{i+1}. Also we obtain that the present value of the cash flows given in Figure 10.8 is 0 since

[3]Swap rate is the fixed rate in an interest rate swap that causes the swap to have a value of zero.

Figure 10.8: $L(T_i)$

$$\begin{aligned}
& B_{T_0} E\left[\frac{1+\delta L(T_i)}{B_{T_{i+1}}} - \frac{1}{B_{T_i}}\Big|\mathscr{F}_{T_0}\right] && (10.6.11)\\
= \; & B_{T_0} E\left[\frac{1}{B_{T_{i+1}}} + \frac{\delta L(T_i)}{B_{T_{i+1}}} - \frac{1}{B_{T_i}}\Big|\mathscr{F}_{T_0}\right] \\
= \; & P(T_0, T_{i+1}) + P(T_0, T_i) - P(T_0, T_{i+1}) - P(T_0, T_i) \\
= \; & 0.
\end{aligned}$$

These results means that portfolio $\delta L(T_{i+1})$ is equivalent to a long position of T_i bond and a short of T_{i+1} bond. $\square$

Now let us calculate the swap contract which is already made before t. In contrast to the cap formula which depends on the term structure model, the swap formula is generic.

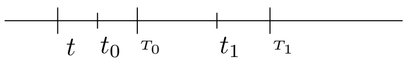


Figure 10.9: Time horizon

1. At time $t_0 (t \le t_0 \le T_0)$: The first swap rate $L(T_0)$ is not yet determined at time t_0. The value of floating rate contract is zero. See Figure 10.6. The entire swap contract value $\Pi(t_0)$ is

$$\Pi(t_0) = P(t_0, T_0) - \delta R \sum_{i=1}^{n} P(t_0, T_i) - P(t_0, T_n).$$

2. At time $t_1 (T_0 \le t_1 < T_1)$: Only the first swap rate $L(T_0)$ has already been determined

$$\begin{aligned}
\Pi(t_1) &= \sum_{i=1}^{n} \text{value}(\delta L(T_{i-1})) - \sum_{i=1}^{n} \text{value}(\delta R(T_i)) \\
&= \Big(1 + \delta L(T_0)\Big) P(t_1, T_1) - P(t_1, T_n) - \delta R \sum_{i=1}^{n} P(t_1, T_i). && (10.6.12)
\end{aligned}$$

Note that the first part of (10.6.12) can be decomposed as follows:

$$\begin{aligned}
& \Big(1 + \delta L(T_0)\Big) P(t_1, T_1) - P(t_1, T_n) \\
= \; & \delta L(T_0) P(t_1, T_1) + P(t_1, T_1) - P(t_1, T_n)
\end{aligned}$$

$$
\begin{aligned}
&= \delta L(T_0)P(t_1,T_1) + \sum_{i=1}^{n-1}\Big(P(t_1,T_i) - P(t_1,T_{i+1})\Big) \\
&= \delta L(T_0)P(t_1,T_1) + \sum_{i=1}^{n-1}\left(\frac{P(t_1,T_i)}{P(t_1,T_{i+1})} - 1\right)P(t_1,T_{i+1}).
\end{aligned}
$$

Hence $\left(\frac{P(t_1,T_i)}{P(t_1,T_{i+1})} - 1\right)$ can be consider as a deterministic cash flow at time T_{i+1}. Also we note that

$$
\begin{aligned}
\delta L(T_i) &\neq \frac{P(t_1,T_i)}{P(t_1,T_{i+1})} - 1, \\
E\big[\delta L(T_i)|\mathscr{F}_{t_1}\big] &\neq \frac{P(t_1,T_i)}{P(t_1,T_{i+1})} - 1.
\end{aligned}
$$

3. Alternative approach: From Figure 10.7 we have

$$
\begin{aligned}
\Pi(t_0) &= \sum_{i=1}^{n} E\left[\frac{B_{t_0}}{B_{T_i}}\Big(\delta L(T_{i-1}) - \delta R\Big)\Big|\mathscr{F}_{t_0}\right] \\
&= \sum_{i=1}^{n} E\left[\frac{B_{t_0}}{B_{T_i}}\Big(\delta L(T_{i-1}) + 1\Big)\Big|\mathscr{F}_{t_0}\right] - \sum_{i=1}^{n} E\left[\frac{B_{t_0}}{B_{T_i}}\Big(1 + \delta R\Big)\Big|\mathscr{F}_{t_0}\right] \\
&= \sum_{i=1}^{n} E\left[\frac{B_{t_0}}{B_{T_i}}\frac{1}{P(T_{i-1},T_i)}\Big|\mathscr{F}_{t_0}\right] - \sum_{i=1}^{n} P(t_0,T_i)(1+\delta R) \\
&= \sum_{i=1}^{n} E\left[\frac{B_{t_0}}{B_{T_{i-1}}}E\left[\frac{B_{T_{i-1}}}{B_{T_i}}\Big|\mathscr{F}_{T_{i-1}}\right]\frac{1}{P(T_{i-1},T_i)}\Big|\mathscr{F}_{t_0}\right] - \sum_{i=1}^{n} P(t_0,T_i)(1+\delta R) \\
&= \sum_{i=1}^{n} E\left[\frac{B_{t_0}}{B_{T_{i-1}}}\Big|\mathscr{F}_{t_0}\right] - \sum_{i=1}^{n} P(t_0,T_i)(1+\delta R) \\
&= \sum_{i=1}^{n} P(t_0,T_{i-1}) - \sum_{i=1}^{n} P(t_0,T_i)(1+\delta R) \\
&= P(t_0,T_0) - \delta R\sum_{i=1}^{n} P(t_0,T_i) - P(t_0,T_n).
\end{aligned}
$$

Remark 10.6.5. Note that $L(T_i)$ is $\mathscr{F}_{T_i}$-measurable. At time t, the cash flow 1 of T_i is equivalent to the cash flow $1+\delta L(T_i)$ of T_{i+1}, see Figure 10.8. Consider the following equations:

$$
\begin{aligned}
P(t,T_i) &= \big(1+\delta L(T_i)\big)P(t,T_{i+1}), && (10.6.13) \\
L(T_i) &= \frac{1}{\delta}\left(\frac{P(t,T_i)}{P(t,T_{i+1})} - 1\right) && (10.6.14) \\
&= \frac{1}{\delta}\left(e^{\delta f(t,T_i,T_{i+1})} - 1\right) \\
&\approx f(t,T_i,T_{i+1}), \quad (\delta \ll 1).
\end{aligned}
$$

The equation (10.6.13) and (10.6.14) are wrong. If (10.6.13) is true, $L(T_i)$ is $\mathscr{F}_t$-measurable. It is false. The correct equations are

$$
E\left[\exp\left(-\int_t^{T_i} r(s)ds\right)\Big|\mathscr{F}_t\right] = E\left[\Big(1+\delta L(T_i)\Big)\exp\left(-\int_t^{T_{i+1}} r(s)ds\right)\Big|\mathscr{F}_t\right],
$$

$$\begin{aligned}
P(t,T_i) &= E\left[\Big(1+\delta L(T_i)\Big)\exp\left(-\int_t^{T_{i+1}} r(s)ds\right)\Big|\mathscr{F}_t\right] \qquad (10.6.15)\\
&= E\left[\exp\left(-\int_t^{T_{i+1}} r(s)ds\right)\Big|\mathscr{F}_t\right] + E\left[\delta L(T_i)\exp\left(-\int_t^{T_{i+1}} r(s)ds\right)\Big|\mathscr{F}_t\right]\\
&= P(t,T_{i+1}) + E\left[\delta L(T_i)\exp\left(-\int_t^{T_{i+1}} r(s)ds\right)\Big|\mathscr{F}_t\right]\\
&= P(t,T_{i+1}) + E\left[\delta L(T_i)\frac{B_t}{B_{T_{i+1}}}\Big|\mathscr{F}_t\right],\\
P(t,T_i) - P(t,T_{i+1}) &= E\left[\delta L(T_i)\frac{B_t}{B_{T_{i+1}}}\Big|\mathscr{F}_t\right].
\end{aligned}$$

This equation has already been proved by (10.6.11) in Example 10.6.4. Also note that we can change the risk-neutral measure to T_{i+1}-forward measure. i.e.

$$\begin{aligned}
E\left[\delta L(T_i)\frac{B_t}{B_{T_{i+1}}}\Big|\mathscr{F}_t\right] &= E_{T_{i+1}}\left[\delta L(T_i)\frac{B_t}{B_{T_{i+1}}}\times\frac{B_{T_{i+1}}}{B_t}P(t,T_{i+1})\Big|\mathscr{F}_t\right]\\
&= E_{T_{i+1}}\left[\delta L(T_i)P(t,T_{i+1})\Big|\mathscr{F}_t\right]\\
&= P(t,T_{i+1})\ E_{T_{i+1}}\left[\delta L(T_i)\Big|\mathscr{F}_t\right],\\
E_{T_{i+1}}\left[L(T_i)\Big|\mathscr{F}_t\right] &= \frac{1}{\delta}\left(\frac{P(t,T_i)}{P(t,T_{i+1})}-1\right).
\end{aligned}$$

Also note that the (10.6.15) can be written as

$$\begin{aligned}
P(t,T_i) &= E\left[\Big(1+\delta L(T_i)\Big)\exp\left(-\int_t^{T_{i+1}} r(s)ds\right)\Big|\mathscr{F}_t\right]\\
&= E\left[\Big(1+\delta L(T_i)\Big)\exp\left(\int_0^{t} r(s)ds\right)\exp\left(-\int_0^{T_{i+1}} r(s)ds\right)\Big|\mathscr{F}_t\right]\\
&= E\left[\Big(1+\delta L(T_i)\Big)\ \frac{B_t}{B_{T_{i+1}}}\Big|\mathscr{F}_t\right]\\
&= B_t\ E\left[\frac{1+\delta L(T_i)}{B_{T_{i+1}}}\Big|\mathscr{F}_t\right].
\end{aligned}$$

Since $P(t,T_i) = B_t E\left[\frac{1}{B_{T_i}}|\mathscr{F}_t\right]$, we have

$$B_t E\left[\frac{1}{B_{T_i}}\ |\ \mathscr{F}_t\right] = B_t\ E\left[\frac{1+\delta L(T_i)}{B_{T_{i+1}}}\ \Big|\ \mathscr{F}_t\right].$$

□

Example 10.6.6. Equation(10.6.6) can be considered as definition of $L(T_i)$. From this we can know that the cash flow -1 at T_i and $1+\delta L(T_i)$ at T_{i+1} is valuated zero at the present(t):

$$\Big(1+\delta L(T_i)\Big)P(t,T_{i+1}) - P(t,T_i)$$

$$
\begin{aligned}
&= \Big(1+\delta L(T_i)\Big)\exp\Big(-\int_t^{T_{i+1}} f(t,s)ds\Big) - P(t,T_i)\\
&= \Big(1+\delta L(T_i)\Big)\exp\Big(-\int_t^{T_i} f(t,s)ds\Big)\ \exp\Big(-\int_{T_i}^{T_{i+1}} f(t,s)ds\Big) - P(t,T_i)\\
&= \Big(1+\delta L(T_i)\Big)P(t,T_i)P(T_i,T_{i+1}) - P(t,T_i)\\
&= P(t,T_i) - P(t,T_i)\\
&= 0.
\end{aligned}
$$

This equations are false because

$$
\exp\Big(-\int_{T_i}^{T_{i+1}} f(t,s)ds\Big) \quad \neq \quad \exp\Big(-\int_{T_i}^{T_{i+1}} f(T_i,s)ds\Big) \quad = \quad P(T_i,T_{i+1}).
$$

□

Summing Up

1. For $t \leq T_i$, $E_{T_{i+1}}\Big[L(T_i)\Big|\mathscr{F}_t\Big] = \frac{1}{T_{i+1}-T_i}\Big(\frac{P(t,T_i)}{P(t,T_{i+1})}-1\Big)$.
2. $P(t,T_{i+1})\ E_{T_{i+1}}\Big[L(T_i)\Big|\mathscr{F}_t\Big] = \frac{1}{T_{i+1}-T_i}\Big(P(t,T_i)-P(t,T_{i+1})\Big)$.

✎ **Convexity Adjustments**

The equation

$$
E_{T_1}[P(T_1,T_2) \mid \mathscr{F}_t] \quad = \quad \frac{P(t,T_2)}{P(t,T_1)}
$$

can be rewritten as

$$
E_{T_1}[e^{-f(T_1,T_1,T_2)(T_2-T_1)} \mid \mathscr{F}_t] \quad = \quad e^{-f(t,T_1,T_2)(T_2-T_1)}.
$$

This means that the expected bond price eqauls the forward price under the forward measure. Consider now a derivatives that provides a payoff dependent on the forward rate $f(T_1,T_1,T_2)$ at time T_1. We, therefore, need to know the value of the expected forward rate $f(T_1,T_1,T_2)$ when the expected bond price equals the forward bond price.

Let $G(x) := e^{-x(T_2-T_1)}$. Then we have

$$
E_{T_1}\left[G\big(f(T_1,T_1,T_2)\big)\right] \quad = \quad G\big(f(0,T_1,T_2)\big).
$$

We need to get

$$
E_{T_1}[f(T_1,T_1,T_2) \mid \mathscr{F}_t].
$$

10.7 Instantaneous Short Rate

The forward rate is the forward price of instantaneous borrowing at time T. The forward rate for borrowing now, at time $T = t$, is exactly the **instantaneous short rate**, that is

$$f(t,t) = r_t$$

1. Short rate dynamics:

$$dr_t = a(t)dt + b(t)dW_t. \tag{10.7.1}$$

2. Bond price dynamics: For all $T \geq t$,

$$dP(t,T) = P(t,T)m(t,T)dt + P(t,T)v(t,T)dW_t. \tag{10.7.2}$$

3. Forward rate dynamics: For all $T \geq t$,

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW_t. \tag{10.7.3}$$

Theorem 10.7.1. The relations between short rate, bond price and forward rate dynamics are given as follow.

1. $\Big[(10.7.2) \Rightarrow (10.7.3)\Big]$ If $P(t,T)$ satisfies (10.7.2), then for the forward rate dynamics we have

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW_t.$$

where α and σ are given by

$$\begin{cases} \alpha(t,T) = v_T(t,T) \cdot v(t,T) - m_T(t,T), \\ \sigma(t,T) = -v_T(t,T). \end{cases}$$

2. $\Big[(10.7.3) \Rightarrow (10.7.1)\Big]$ If $f(t,T)$ satisfies (10.7.3) then the short rate satisfies

$$dr_t = a(t)dt + b(t)dW_t,$$

where

$$\begin{cases} a(t) = f_T(t,t) + \alpha(t,t), \\ b(t) = \sigma(t,t). \end{cases}$$

3. $\Big[(10.7.3) \Rightarrow (10.7.2)\Big]$ If $f(t,T)$ satisfies (10.7.3) then $P(t,T)$ satisfies

$$dP(t,T) = P(t,T)\left(r_t + A(t,T) + \frac{1}{2}\|\Sigma(t,T)\|^2\right)dt + P(t,T)\Sigma(t,T)dW_t,$$

where

$$\begin{cases} A(t,T) = -\int_t^T \alpha(t,s)ds, \\ \Sigma(t,T) = -\int_t^T \sigma(t,s)ds. \end{cases}$$

Proof.

1.

$$
\begin{aligned}
f(t,T) &= -\frac{\partial}{\partial T}\log P(t,T),\\
df(t,T) &= -\frac{\partial}{\partial T}\left(\frac{1}{P(t,T)}dP(t,T) - \frac{1}{2}\frac{1}{P(t,T)^2}\big(dP(t,T)\big)^2\right)\\
&= -\frac{\partial}{\partial T}\left(m(t,T)dt + v(t,T)dW_t - \frac{1}{2}v(t,T)^2dt\right)\\
&= \Big(v_T(t,T)\cdot v(t,T) - m_T(t,T)\Big)dt - v_T(t,T)dW_t.
\end{aligned}
$$

2. Since $r_t = f(t,t)$,

$$
\begin{aligned}
dr_t &= \frac{\partial}{\partial T}f(t,T)\Big|_{T=t}dt + df(t,T)\Big|_{T=t}\\
&= f_T(t,t)dt + \alpha(t,t)dt + \sigma(t,t)dW_t.
\end{aligned}
$$

3. Let $Y(t,T) = -\int_t^T f(t,s)ds$.

$$
\begin{aligned}
P(t,T) &= \exp\Big(Y(t,T)\Big),\\
dP(t,T) &= P(t,T)dY + \frac{1}{2}P(t,T)(dY)^2,\\
dY(t,T) &= f(t,t)dt - \int_t^T df(t,s)ds\\
&= f(t,t)dt - \int_t^T \alpha(t,s)dtds - \int_t^T \sigma(t,s)dW_tds\\
&= f(t,t)dt + A(t,T)dt + \Sigma(t,T)dW_t
\end{aligned}
$$

Thus

$$
\begin{aligned}
dP(t,T) &= P(t,T)\Big(f(t,t) + A(t,T)\Big)dt + P(t,T)\Sigma(t,T)dW_t + \frac{1}{2}P(t,T)\|\Sigma(t,T)\|^2dt\\
&= P(t,T)\left(r_t + A(t,T) + \frac{1}{2}\|\Sigma(t,T)\|^2\right)dt + P(t,T)\Sigma(t,T)dW_t.
\end{aligned}
$$

■

Theorem 10.7.2 (Market Price of Risk). If

$$
dr_t = a\big(t,r(t)\big)dt + b\big(t,r(t)\big)dW_t, \tag{10.7.4}
$$

the bond price of maturity T, $P(t,r(t),T) = P^T(t,r)$ satisfies

$$
dP^T = P^Tm^Tdt + P^Tv^TdW_t.
$$

where

$$
\begin{cases}
m^T = \frac{P_t^T + aP_r^T + \frac{1}{2}b^2P_{rr}^T}{P^T},\\
v^T = \frac{bP_r^T}{P^T}.
\end{cases}
$$

Also, there exists a process γ such that the relation

$$\frac{m^T(t)-r_t}{v^T(t)} = \gamma(t)$$

holds for all t and for every choice of maturity time T.

Proof.

$$\begin{aligned} dP^T &= P_t dt + P_r^T dr_t + \frac{1}{2}P_{rr}^T (dr_t)^2 \\ &= P_t^T dt + P_r^T (a\,dt + b\,dW_t) + \frac{1}{2}P_{rr}^T b^2 dt \\ &= P^T\left(\frac{P_t^T + aP_r^T + \frac{1}{2}P_{rr}^T b^2}{P^T}\right)dt + P^T\left(\frac{bP_r^T}{P^T}\right)dW_t. \end{aligned}$$

Consider a portfolio $V(t)$ which consists of T-bond and S-bond with the relative portfolio (u^T, u^S). i.e.

$$\begin{aligned} V(t) &= \frac{u^T V}{P^T}P^T + \frac{u^T V}{P^S}P^S, \quad u^T + u^S = 1, \\ dV &= \frac{u^T V}{P^T}dP^T + \frac{u^T V}{P^S}dP^S \\ &= \frac{u^T V}{P^T}\left(P^T m^T dt + P^T v^T dW_t\right) + \frac{u^S V}{P^S}\left(P^S m^S dt + P^S v^S dW_t\right) \\ &= V\{u^T m^T + u^S m^S\}dt + V\{u^T v^T + u^S v^S\}dW_t. \end{aligned}$$

Now we choose the portfolio weights as the solution(for each t) to the system

$$\begin{cases} u^T + u^S = 1, \\ v^T u^T + v^S u^S = 0. \end{cases} \tag{10.7.5}$$

The solution of (10.7.5) is given by

$$\begin{cases} u^T = -\frac{v^S}{v^T - v^S}, \\ u^S = \frac{v^T}{v^T - v^S}. \end{cases}$$

and we end up with the following V-dynamics

$$dV = V\left\{\frac{m^S v^T - m^T v^S}{v^T - v^S}\right\}dt$$

Since $dV = r_t V dt$, we have

$$r_t = \frac{m^S v^T - m^T v^S}{v^T - v^S}.$$

Thus we have, after some manipulations,

$$\frac{m^T(t)-r_t}{v^T(t)} = \frac{m^S(t)-r_t}{v^S(t)}.$$

The left hand side of this equation does not depend upon the choice of T, whereas the right hand side does not depend on S. This quotient is often termed **the market price of risk**.

$$\gamma_t = \frac{m^T(t)-r_t}{v^T(t)}. \tag{10.7.6}$$

On an arbitrage free market all bonds will have the same market price of risk. ■

Note that although $a(t)$ and $b(t)$ are given in Equation (10.7.4), we cannot know m^T and v^T. Hence the market price of risk γ cannot be determined, because m^T and v^T is not determined from the dynamics of the short rate r. Therefore we cannot determine the market price of risk. Hence the followings are equivalent.

1. m^T and v^T are determined.
2. The martingale measure is determined.
3. The market price of risk is determined.

physical measure	risk neutral measure: γ is given
$$\begin{aligned} dr_t &= a(t,r_t)dt + b(t,r_t)dW_t, \\ dP &= P\Big(m(t,T)dt + v(t,T)dW_t\Big), \\ 0 &= P_t^T + (a-\lambda b)P_r^T + \frac{1}{2}b^2 P_{rr}^T - rP^T, \\ 1 &= P^T(T,r). \end{aligned}$$	$$\begin{aligned} dr_t &= \big(a(t,r_t) - \gamma_t b(t,r_t)\big)dt + b(t,r_t)d\widetilde{W}_t, \\ &= \widetilde{a}(t,r_t)dt + \widetilde{b}(t,r_t)d\widetilde{W}_t, \\ dP &= P\Big(r_t dt + v(t,T)d\widetilde{W}_t\Big), \\ 0 &= P_t^T + \widetilde{a}P_r^T + \frac{1}{2}b^2 P_{rr}^T - rP^T, \\ 1 &= P^T(T,r). \end{aligned}$$

Theorem 10.7.3 (Term Structure Equation). If the bond market is free of arbitrage then P^T will satisfy the equation

$$\begin{cases} P_t^T + (a-b\gamma)P_r^T + \frac{1}{2}b^2 P_{rr}^T - rP^T &= 0, \\ P^T(T,r) &= 1, \end{cases}$$

if

$$dr_t = a(t)dt + b(t)dW_t,$$

where γ_t is the market price of risk.

Proof. From Equation (10.7.6),

$$\begin{aligned} m^T &= r_t + v^T\gamma_t, \\ \frac{P_t^T + aP_r^T + \frac{1}{2}b^2 P_{rr}^T}{P^T} &= r_t + \frac{bP_r^T}{P^T}\gamma_t, \quad \text{by Theorem (10.7.2).} \end{aligned}$$

■

It is natural to view bonds as derivatives, with the short rate of interest as the underlying object, in the same way that options are derivatives of the stock price process in the Black-Scholes model. The first natural question to ask is then whether bond prices are uniquely determined by a specification of the P-dynamics or r as in Eq. (10.7.4), plus a requirement that the bond market should be free of arbitrage.

The answer to this question is emphatically NO, and the reason is that our a priori specified market is not complete. The only exogenously given asset is the risk free one. In particular there is no possibility of replicating any interesting contingent claim, not even the simple claim associated with a zero coupon bond.

Summing up

1. The price of bond is not uniquely determined by a specification of the P-dynamics of the short rate of interest, plus the requirement of an arbitrage free bond market.
2. This is due to the fact that arbitrage pricing is a matter of pricing a derivative in terms of a given underlying price process. In the a priori given market, we do not have a sufficiently rich family of underlying assets.
3. Bonds of different maturity must satisfy certain internal consistency relations in order to avoid arbitrage possibilities.
4. If we include one single bond in the exogenously given market then we ought to be able to price all other bonds in terms of this benchmark bond.

10.7.1 Measure Change

The simplest cash product is the account, or bond formed by starting with one dollar at time zero and reinvesting continually at instantaneous short rate. Then the B_t is a stochastic process satisfying the SDE

$$dB_t = r_t B_t dt, \quad B_0 = 1, \quad \text{or } B_t = \exp\Big(\int_0^t r_s ds\Big).$$

Let

$$Z(t,T) = \frac{P(t,T)}{B_t}, \quad 0 \le t \le T,$$

where

$$dP(t,T) = P(t,T)m(t,T)dt + P(t,T)v(t,T)dW_t. \tag{10.7.7}$$

Note that $m(t,T)$ and $v(t,T)$ cannot be obtained from the short rate dynamics. If we are already given $m(t,T)$ and $v(t,T)$ or the market price of risk, the risk neutral measure can be uniquely determined.

Now, let us find a risk neutral measure, $\widetilde{Q}$, under which $Z(t,T)$ is a martingale.

$$\begin{aligned} dZ(t,T) &= -\frac{1}{B_t^2}P(t)dB_t + \frac{1}{B_t}dP(t,T) \\ &= -Z_t r_t dt + Z_t\Big(m(t,T)dt + v(t,T)dW_t\Big) \\ &= Z_t\Big(m(t,T) - r_t\Big)dt + Z_t v(t,T)dW_t. \end{aligned}$$

Let

$$\gamma_t = \frac{m(t,T) - r_t}{v(t,T)}. \tag{10.7.8}$$

Then by the Girsanov theorem,

$$d\widetilde{W}_t = dW_t + \gamma_t dt$$

is a new Brownian motion with respect to martingale measure $\widetilde{Q}$ under which Z_t is a martingale. Hence equation (10.7.7) is converted

$$dP(t,T) = P(t,T)\Big(r_t dt + v(t,T)d\widetilde{W}_t\Big). \tag{10.7.9}$$

The price of T-bond $P(t,T)$:

$$\begin{aligned} P(t,T) &= B_t E_{\widetilde{Q}}\left[\frac{1}{B_T}|\mathscr{F}_t\right] \\ &= E_{\widetilde{Q}}\left[\exp\Big(-\int_t^T r_s ds\Big)\Big|\mathscr{F}_t\right], \\ \text{where} \qquad dr_t &= (a-\gamma b)dt + b\, d\widetilde{W}_t. \end{aligned}$$

10.7.2 Replicating Strategies

We have found the martingale measure and the process for $P(t,T)$ under it. But we ought to check that we can produce replicating strategies for claim. Suppose we have a claim X which pays off at time S. If we are going to hedge this with a discount bond maturing at time T, our only restriction is that S should come before T - we cannot hedge a long-term product with a shorter-term instrument [4]. Suppose, for simplicity, we choose to use a bond with maturity T larger than S.

As before, our second step to replication is to form the conditional $\widetilde{Q}$-expectation of the discounted claim $B_S^{-1}X$, rather than the raw claim X. That is, we define E_t to be the $\widetilde{Q}$-martingale

$$E_t := E_{\widetilde{Q}}\left[\frac{X}{B_S}\;\Big|\;\mathscr{F}_t\right].$$

For the martingale representation theorem to be used, we also need that the bond volatility $v(t,T)$ is never zero before T, in which case, we apply the representation theorem to the martingale $Z(t,T)$ and the discounted claim process E_t. This gives us that

$$E_t = E_0 + \int_0^t \phi_s\, dZ(s,T),$$

for some $\mathscr{F}_t$-previsible process ϕ.

Our trading strategy will be a combination of both a holding in the T-bond and a holding in the cash bond B_t. Specifically, we

- hold ϕ_t units of the T-bond a time t,
- hold $\psi_t := E_t - \phi_t Z(t,T)$ units of the cash bond(B_t) at time t.

[4]Unless we split the time-period up into shorter subsections, and roll over short-term bonds from section to section

The value of this portfolio at time t is

$$V_t = B_t E_t = B_t E_{\widetilde{Q}}\left[\frac{X}{B_S} \,\Big|\, \mathscr{F}_t\right].$$

This strategy (ϕ_t, ψ_t) will be self-financing if $dV_t = \phi_t dP(t,T) + \psi_t dB_t$, or equivalently if

$$dE_t = \phi_t dZ(t,T).$$

Which is ensured by the representation of E_t in terms of ϕ_t. The portfolio (ϕ_t, ψ_t) is self-financing. Thus if X is the payoff of a derivative maturing at time T, then its value at time t is

$$V_t = B_t E_{\widetilde{Q}}\left[\frac{X}{B_T} \,\Big|\, \mathscr{F}_t\right] = B_t E_{\widetilde{Q}}\left[\exp\left(-\int_t^T r_s ds\right) X \,\Big|\, \mathscr{F}_t\right].$$

10.8 Interest Rate Products

This section follows Baxter and Rennie(1996).

10.8.1 Forward Condtract

We agree, at the current time t, to make a payment of an amount k at a future time T_1, and in return to receive a dollar at the later time T_2. What should the amount k be?

According to the pricing formula (under whatever model we are in), the value of the claim now is

$$V_t = B_t E_{\widetilde{Q}}\left[\frac{1}{B_{T_2}} \mid \mathscr{F}_t\right] - B_t E_{\widetilde{Q}}\left[\frac{k}{B_{T_1}} \mid \mathscr{F}_t\right],$$

under the martingale measure $\widetilde{Q}$.

Recalling that $B_t E_{\widetilde{Q}}\left[\frac{1}{B_T} \mid \mathscr{F}_t\right]$ is just $P(t,T)$, we see that

$$V_t = P(t,T_2) - kP(t,T_1).$$

For this contract to have null value current net worth, we merely choose k at time t to be

$$k = \frac{P(t,T_2)}{P(t,T_1)}.$$

This price makes sense, as saying that the forward yield from T_1 to T_2 is

$$-\frac{\log P(t,T_2) - \log P(t,T_1)}{T_2 - T_1}.$$

For T_1 and T_2 very close together, this approximations to be instantaneous forward rate of borrowing

$$-\frac{\partial}{\partial T}\log P(t,T) = f(t,T).$$

The price also gives us a clue to the hedging strategy. Suppose we were, at time t, to buy k units of the T_1-bond ans sell one unit of the T_2-bond. The initial cost of that deal is zero, and the portfolio pays us k at time T_1 and exactly absorbs the dollar we receive at time T_2.

In this particular example, the answer is independent of our particular term structure model, as the hedging strategy is static. There are other important cases where this also happens.

10.8.2 Multiple Payment Contracts

Most interest rate products don't just make a single payment X at time T. Instead the contract specifies a sequence of payments X_i made at a sequence of times $T_i (i = 1, \cdots, n)$. Each payment X_i may depend on price movements up to its payment time T_i, and even on any previous payment. As long as we bear that in mind, this causes no serious problem, and indeed there are two different ways to keep things clear.

Divide and rule We can treat each payment X_i separately. On tis own, it is just a claim at time T_i, so its worth at time t is

$$V_i(t) = B_t E_{\widetilde{Q}}\left[\frac{X_i}{B_{T_i}} \mid \mathscr{F}_t\right] = P(t,T_i)E_{T_i}\left[X_i \mid \mathscr{F}_t\right],$$

where E_{T_i} is the expection under the T_i-forward measure. This approach will always work, but the forward measure, if used, will have to be changed for each i.

Savings account We could instead roll up the payments into savings as we get them, and keep them till the last payment date $T = T_n$. That is, as each payment is made, we use it to buy a T-bond(or invest it in the bank account process B_t till time T). Then the payoff is a single payment at time T of

$$X = \sum_{i=1}^{n} \frac{X_i}{P(T_i,T)}$$

with worth at time t

$$V_t = B_t E_{\widetilde{Q}}\left[\frac{X}{B_T} \mid \mathscr{F}_t\right] = P(t,T)E_T\left[X \mid \mathscr{F}_t\right].$$

10.8.3 Floating Rate Notes

A bond might also pay off a coupon which was not fixed, but depend on current interest rates. One interesting case is where the interest paid over on inverval from time S to time T is the same as the yield of the T-bond bought as time S.

Suppose a bond pays its dollar principal at time T_n, and also payments at time $T_i = T_0 + i\delta(i = 1,\cdots,n)$ of varying amounts. The amount of payment made at time T_i is determined by the LIBOR rate set at time T_{i-1}

$$L(T_{i-1}) := \frac{1}{\delta}\left(\frac{1}{P(T_{i-1},T_i)} - 1\right).$$

The actual payment made at time T_i is $\delta L(T_{i-1}) = P(T_{i-1},T_i)^{-1} - 1$, which is the amount of interest we would receive by buying a dollar's worth of the T_i bond at time T_{i-1}.

The value to us now, at time T_0, of the T_i payment is

$$B_{T_0} E_{\widetilde{Q}}\left[\frac{1}{B_{T_i}}\left(\frac{1}{P(T_{i-1},T_i)} - 1\right) \mid \mathscr{F}_{T_0}\right].$$

Using (10.6.10) we can rewrithe the value of the T_i payment as

$$P(T_0,T_{i-1}) - P(T_0,T_i).$$

This price also suggest the hedge of selling a T_i-bond and buying a T_{i-1}-bond. When the T_{i-1}-bond matures, we buy $P(T_i,T_{i-1})^{-1}$ units of the T_i-bond, and we are left with exactly the right payoff at time T_i.

The total value of the variable coupon bond is the sum of its components. That is,

$$V_0 = P(T_0,T_n) + \sum_{i=1}^{n}\Big(P(T_0,T_{i-1}) - P(T_0,T_i)\Big)$$

$$= \quad 1.$$

Surprisingly, the bond has a fixed price equal to the face value of its principal. Why this is so, is because the bond is equivalent to this simple sequence of trades:

1. Take a dollar and buy $\frac{1}{P(T_0,T_1)}$ units T_1-bonds.
2. Take the interest from the bonds at T_1 as a coupon, and buy $\frac{1}{P(T_1,T_2)}$ units T_2-bonds with the dollar principal.
3. Repeat until we are left with the dollar at time T_n.

This has exactly the same cash flows as the floating rate note, so the initial prices must match.

10.8.4 Swaps

This very popular contract simply exchanges a stream of varying payments for a stream of fixed amount payments (or vice versa). That is, we swap a floting interest rate for a fixed one.

Typically, we might offer a contract where we receive a regular sequence of fixed amounts and at each payments date we pay an amount depending on prevailing interest rates. In practice, only the net difference is exchanged.

A standard definition of the variable payment is that of the interest paid by a bond over the previous time period. If the payment dates are $T_i = T_0 + i\delta(i = 0, \cdots, n)$, then the ith payment will be determined by the δ-period LIBOR rate set at time T_{i-1}. The payment made is

$$\delta L(T_{i-1}) \quad = \quad \frac{1}{P(T_{i-1}, T_i)} - 1.$$

Suppose the swap pays at a fixed rate k at each time period. Then the swap looks like a portfolio which is long a fixed coupon bond and short a floating coupon bond. We know that the former is worth

$$P(T_0, T_n) + k\delta \sum_{i=1}^{n} P(T_0, T_i),$$

and the latter costs a dollar. The fixed rate needed to give the swap initial null value is

$$k \quad = \quad \frac{1 - P(T_0, T_n)}{\delta \sum_{i=1}^{n} P(T_0, T_i)}.$$

10.8.5 Forward Swaps

In a forward swap agreement, we have chosen to receive fixed payments at rate k, starting at time T_0 wiht payments at times $T_i = T_0 + i\delta(i = 0, \cdots, n)$. The value of this swap at time T_0 will be

$$X \quad = \quad P(T_0, T_n) + k\delta \sum_{i=1}^{n} P(T_0, T_i) - 1.$$

The present value of X at time t before T_0 is given by the formula:

$$\begin{aligned} V_t &= B_t E_{\tilde{Q}}\left[\frac{X}{B_{T_0}} \mid \mathscr{F}_t\right] \\ &= P(t,T_n) + k\delta \sum_{i=1}^{n} P(t,T_i) - P(t,T_0). \end{aligned}$$

The fixed rate needed to give the forward swap initial null value at time t is

$$k = \frac{P(t,T_0) - P(t,T_n)}{\delta \sum_{i=1}^{n} P(t,T_i)}.$$

This rate k is the forward swap rate. An alternative formulation of this expression is

$$k = \frac{1 - F_t(T_0,T_n)}{\delta \sum_{i=1}^{n} F_t(T_0,T_i)},$$

where $F_t(T_0,T_i)$ is the forward price at time t for purchasing a T_i-bond at time T_0. That is

$$F_t(T_0,T_i) = \frac{P(t,T_i)}{P(t,T_0)}.$$

In this form the expression resembles the instantaneous swap rate.

10.9 Short Rate Models

The difference between an equilibrium and a no arbitrage model is as follows. In an equilibrium model, today's term structure of interest rates is an output. In a no arbitrage model, today's term structure of interest rates is an input.

1. **(One Factor) Equilibrium Models**:
   The process r involves only one source of uncertainty. Usually the short rate can be described in a risk-neutral world by an Ito process of the form
   $$dr = m(r)dt + s(r)d\widetilde{W}.$$
   The instantaneous drift m, and instantaneous volatility, s, are assumed to be function of r, but are **independent of time**. i.e. the drift of the short rate is not usually a function of time. The disadvantage of the equilibrium models is that they do not automatically fit today's term structure.

2. **No-Arbitrage Models**:
   A no arbitrage model is a model designed to be exactly consistent with today's term structure of interest rates.

✎ **Several Short Rate Models**
If a parameter is time dependent, this is written explicitly. Otherwise all parameters are constant.

1. Vasicek(1977): (Ornstein Uhlenbeck process)
   $$dr = (b - ar)dt + \sigma d\widetilde{W}.$$
2. Cox-Ingersoll-Ross(1985): (Bessel Process)
   $$dr = (b - ar)dt + \sigma\sqrt{r}d\widetilde{W}.$$
3. Dothan:
   $$dr = ardt + \sigma rd\widetilde{W}.$$
4. Black-Derman-Toy:
   $$dr = a(t)rdt + \sigma(t)rd\widetilde{W}.$$
5. Ho-Lee(1896):
   $$dr = \theta_t dt + \sigma d\widetilde{W}.$$
6. Hull-White(extended Vasicek, 1990):
   $$dr = (\theta_t - a(t)r)dt + \sigma(t)d\widetilde{W}.$$
7. Hull-White(extended CIR):
   $$dr = (\theta_t - a(t)r)dt + \sigma(t)\sqrt{r}\ d\widetilde{W}.$$

8. Black-Karasinski(1991):

$$dX_t = \big(\theta_t - a(t)X_t\big)dt + \sigma(t)d\widetilde{W}, \quad X_t = \log r_t.$$

Example 10.9.1. Consider a SDE:

$$dr_t \;=\; \Big(g(t) + h(t)r_t\Big)dt + \sigma(t)dW_t,$$

with g, h, and σ all deterministic functions of time. To solve the SED, first, let

$$H(t) \;=\; \int_0^t h(s)ds.$$

Applying Ito's formula, we get

$$\begin{aligned} d\Big(e^{-H(t)}r_t\Big) &= -e^{-H(t)}h(t)r_t dt + e^{-H(t)}\Big[\Big(g(t) + h(t)r_t\Big)dt + \sigma(t)dW_t\Big] \\ &= e^{-H(t)}\Big(g(t)dt + \sigma(t)dW_t\Big). \end{aligned}$$

Hence we obtain that for $u < t$

$$\begin{aligned} e^{-H(t)}r_t &= e^{-H(u)}r_u + \int_u^t e^{-H(s)}g(s)ds + \int_u^t e^{-H(s)}\sigma(s)dW_s, \\ r_t &= e^{H(t)-H(u)}r_u + \int_u^t e^{H(t)-H(s)}g(s)ds + \int_u^t e^{H(t)-H(s)}\sigma(s)dW_s. \end{aligned}$$

In particular, we have

$$r_t \;=\; e^{H(t)}r_0 + \int_0^t e^{H(t)-H(s)}g(s)ds + \int_0^t e^{H(t)-H(s)}\sigma(s)dW_s.$$

□

10.10 Parameter Estimations(= Inverting The Yield Curve = Yield Curve Fitting)

We have choose to model our r-process by giving the $\widetilde{Q}$-dynamics, which means that the parameter in short dynamics are the one which hold under the martingale measure $\widetilde{Q}$. When we make observation in the real world we are not observing r under the martingale measure $\widetilde{Q}$, but under the physical measure P. This means that if we apply standard statistical procedures to our observed date we will not get our $\widetilde{Q}$-parameters. However, the diffusion term is the same under P and under $\widetilde{Q}$. It is because a Girsanov transformation will only affect the drift terms of a diffusion but not the diffusion term.

When it comes to the estimation of parameters affecting the drift term of r we have to use completely different methods. Thus, in order to obtain information about the $\widetilde{Q}$-drift parameters we have to collect price information from the market, and the typical approach is that of **fitting the yield curve**. For a detail, see Björk(1998) pp. 254.

10.11 Affine Term Structures

Definition 10.11.1 (Affine Term Structure(ATS)). If bond prices are given by $P(t,T) = P^T(t,r(t))$ where P^T has the form

$$P^T(t,r) = e^{A(t,T)-B(t,T)r_t}, \tag{10.11.1}$$

and where $A(t,T)$ and $B(t,T)$ are deterministic functions, the model is said to possess an **Affine Term Structure**. □

We assume that $\widetilde{Q}$ is the risk neutral measure, and that the short rate have the following $\widetilde{Q}$-dynamics

$$dr_t = \mu(t,r_t)dt + \sigma(t,r_t)d\widetilde{W}_t,$$

where $\widetilde{W}$ is a $\widetilde{Q}$-Brownian motion. From theorem (10.7.3) or the fact that the drift of dP equals r_t we have

$$A_t(t,T) - \Big(1 + B_t(t,T)\Big)r - \Big(\mu(t,r) - \lambda\sigma(t,r)\Big)B(t,T) + \frac{1}{2}\sigma^2(t,r)B^2(t,T) = 0.$$

Since $\lambda = 0$ under risk neutral measure we have

$$A_t(t,T) - \Big(1 + B_t(t,T)\Big)r - \mu(t,r)B(t,T) + \frac{1}{2}\sigma^2(t,r)B^2(t,T) = 0. \tag{10.11.2}$$

The boundary condition $P^T(T,r) \equiv 1$ implies

$$\begin{aligned} A(T,T) &= 0, \\ B(T,T) &= 0. \end{aligned}$$

We see immediately that if μ and σ^2 both are affine in r, the equation (10.11.2) becomes separable. We thus make the additional assumption that μ and σ have the form

$$\begin{aligned} \mu(t,r) &= \alpha(t)r + \beta(t), \\ \sigma(t,r) &= \sqrt{\gamma(t)r + \delta(t)}. \end{aligned} \tag{10.11.3}$$

Then we have

$$A_t(t,T) - \beta(t)B(t,T) + \frac{1}{2}\delta(t)B^2(t,T) - \Big\{1 + B_t(t,T) + \alpha(t)B(t,T) - \frac{1}{2}\gamma(t)B^2(t,T)\Big\}r = 0.$$

Theorem 10.11.2. Assume that μ and σ are of the form equation (10.11.3). Then the model has an affine term structure of the form (10.11.1), where A and B satisfy the system

$$1 + B_t(t,T) + \alpha(t)B(t,T) - \frac{1}{2}\gamma(t)B^2(t,T) = 0, \quad B(T,T) = 0, \tag{10.11.4}$$

$$A_t(t,T) = \beta(t)B(t,T) - \frac{1}{2}\delta(t)B^2(t,T), \quad A(T,T) = 0. \tag{10.11.5}$$

We note equation (10.11.4) is a Riccati equation [5] for B(for each fixed T). Having solved (10.11.4) we can then easily obtain A by simply integration equation (10.11.5).

[5]Riccati Equation:

$$\frac{dy}{dx} = q_1(x) + q_2(x)y + q_3(x)y^2.$$

If y_1 is a particular solution, $y = y_1 + \frac{1}{v(x)}$ is the general solution if $v(x)$ satisfies the first order equation

$$v' = -(q_2 + 2q_3y_1)v - q_3.$$

Remark 10.11.3. If we want to fit the theoretical prices to the observed prices it is convenient to do this using forward rates rather than bond prices. From Equation (10.6.4) and (10.11.1), we have

$$f(0,T) = B_T(0,T)r_0 - A_T(0,T). \tag{10.11.6}$$

More generally we have

$$f(t,T) = B_T(t,T)r_t - A_T(t,T).$$

Example 10.11.4 (The Bond Price Dynamics). From (10.11.1), by Ito' Lemma we have

$$\begin{aligned}
dP^T &= \left(\frac{\partial A}{\partial t}dt - \frac{\partial B}{\partial t}r_t dt - B dr_t\right) + \frac{1}{2}P^T B^2 (dr_t)^2 \\
&= P^T\left\{\left(\frac{\partial A}{\partial t} - \frac{\partial B}{\partial t}r_t + \frac{1}{2}B^2\sigma^2 - B\mu\right)dt - \sigma B d\widetilde{W}_t\right\} \\
&= P^T\Big(r_t dt - \sigma B d\widetilde{W}_t\Big), \quad \text{by (10.11.2)} \\
&= P^T\Big(r_t dt - \sigma(t,r_t)B(t,T)d\widetilde{W}_t\Big).
\end{aligned}$$

□

Example 10.11.5 (Change of Numeraire). The forward measure for any date T_F is the measure associated with taking the T_F-maturity bond $P(t,T_F)$ as numeraire asset. From Girsanov's Theorem, it follows that the process $W_t^{T_F}$ defined by

$$dW_t^{T_F} := d\widetilde{W}_t + \sigma(t,r_t)B(t,T_F)dt$$

is a standard Brownian motion under the forward measure Q^{T_F}. Accordingly, the dynamics of the affine term structure model become

$$dr_t = \Big(\mu(t,r_t) - \sigma^2(t,r_t)B(t,T_F)\Big)dt + \sigma(t,r_t)dW_t^{T_F}.$$

Suppose we want to price a derivative security making a payoff of $g(r_{T_F})$ at time T_F. Under the risk-neutral measure, we would price the security by computing

$$E_{\widetilde{Q}}\left[\exp\left(-\int_0^{T_F} r_u du\right)g(r_{T_F})\right].$$

In fact, g could be a function of the path of r_t rather than just its terminal value. Switching to the forward measure, this become

$$P(0,T_F)E_{Q^{T_F}}\Big[g(r_{T_F})\Big],$$

where $E_{Q^{T_F}}$ denotes expectation under the forward measure. Thus, we may price the derivative security by simulating r_t under the forward measure Q^{T_F}, estimating the expectation $g(r_{T_F})$ and multiplying by $P(0,T_F)$. Notice that discounting in this case is deterministic - we do not need to simulate a discount factor. This apparent simplification results form inclusion of the additional term $-\sigma^2(t,r_t)B(t,T_F)$ in the drift of r_t.

A consequence of working under the forward measure is that the simulation prices the bond maturing at T_F exactly: pricing this bond corresponds to taking $g(r_{T_F}) \equiv 1$. Again, this apparent simplification is really a consequence of the form of the drift of r_t under the forward measure. □

Example 10.11.6. For $T \leq T_F$, we have

$$\begin{aligned} X_t &= B_t E_{\tilde{Q}} \left[\frac{X_T}{B_T} \mid \mathscr{F}_t \right] \\ &= P(t, T_F) E_{Q^{T_F}} \left[\frac{X_T}{P(T, T_F)} \mid \mathscr{F}_t \right] \end{aligned}$$

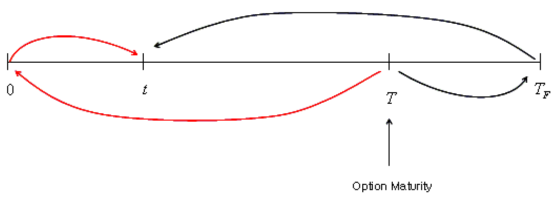


Figure 10.10: Discounting Factor

10.11.1 Vasicek Models(Affine Term Structure)

The Vasicek model takes the following form:

$$dr = (b - ar)dt + \sigma dW.$$

The model is mean-reverting to a constant level, which is a good property, but interest rates can easily go negative, which is a very bad property.

✎ **Method 1: Affine Term Structure Formula**

Then we have $\alpha(t) \equiv -a, \beta(t) \equiv b, \delta(t) \equiv \sigma^2, \gamma(t) \equiv 0$ in (10.11.3). Hence

$$\begin{aligned} 1 + B_t(t,T) - aB(t,T) = 0, &\quad B(T,T) = 0, \\ A_t(t,T) = bB(t,T) - \frac{1}{2}\sigma^2 B^2(t,T), &\quad A(T,T) = 0. \end{aligned}$$

The solution is

$$\begin{aligned} B(t,T) &= \frac{1 - e^{-a(T-t)}}{a}, \\ A(t,T) &= -b\int_t^T B(s,T)ds + \frac{1}{2}\sigma^2 \int_t^T B^2(s,T)ds, \quad \text{put } y := B(s,T), \\ &= -b\int_{B(t,T)}^0 \frac{y}{ay-1}dy + \frac{1}{2}\sigma^2 \int_{B(t,T)}^0 \frac{y^2}{ay-1}dy \\ &= -b\int_{B(t,T)}^0 \Big\{\frac{1}{a} + \frac{1}{a(ay-1)}\Big\}dy + \frac{1}{2}\sigma^2 \int_{B(t,T)}^0 \Big\{\frac{1}{a}y + \frac{1}{a^2} + \frac{1}{a^2(ay-1)}\Big\}dy \\ &= -b\Big(-\frac{B(t,T)}{a} - \frac{1}{a^2}\log|aB(t,T)-1|\Big) + \frac{1}{2}\sigma^2\Big(-\frac{B^2(t,T)}{2a} - \frac{B(t,T)}{a^2} - \frac{1}{a^3}\log|aB(t,T)-1|\Big) \\ &= \frac{b}{a}\Big(B(t,T) - (T-t)\Big) + \frac{1}{2}\sigma^2\Big(\frac{1}{a^2}(T-t) - \frac{B^2(t,T)}{2a} - \frac{B(t,T)}{a^2}\Big) \\ &= \frac{b}{a}\Big(B(t,T) - (T-t)\Big) + \frac{\sigma^2}{2a^2}\Big((T-t) - B(t,T)\Big) - \frac{\sigma^2 B^2(t,T)}{4a} \qquad (10.11.7) \\ &= \frac{\big(B(t,T) - (T-t)\big)\big(ab - \frac{\sigma^2}{2}\big)}{a^2} - \frac{\sigma^2 B^2(t,T)}{4a}. \end{aligned}$$

The bond price is

$$P(t,T) = e^{A(t,T) - B(t,T)r(t)}.$$

Using (10.6.1) yields

$$\begin{aligned} R(t,T) &= -\frac{1}{T-t}A(t,T) + \frac{1}{T-t}B(t,T)r(t), \\ dR(t,T) &= \Big(\cdots\Big)dt + \frac{\sigma}{T-t}B(t,T)dW_t \end{aligned}$$

This means that the volatility of zero rate $R(t,T)$ is given by

$$\frac{\sigma}{T-t}B(t,T).$$

✎ **Method 2: Direct Calculation(exponential martingale)**
The bond price is given by

$$P(t,T) = E_{\widetilde{Q}}\Big[\exp\Big(-\int_t^T r_s ds\Big)\Big|r_t\Big].$$

Note that $d\Big(e^{as}(r_s - \frac{b}{a})\Big) = \sigma e^{as} dW_s$. For $s \geq t$,

$$\begin{aligned} e^{as}(r_s - \frac{b}{a}) &= e^{at}(r_t - \frac{b}{a}) + \sigma \int_t^s e^{au} dW_u, \\ r_s - \frac{b}{a} &= e^{-a(s-t)}(r_t - \frac{b}{a}) + \sigma e^{-as} \int_t^s e^{au} dW_u \\ r_s &= \frac{b}{a}(1 - e^{-a(s-t)}) + e^{-a(s-t)} r_t + \sigma e^{-as} \int_t^s e^{au} dW_u. \end{aligned}$$

This gives us

$$\begin{aligned} -\int_t^T r_s ds &= -\frac{b}{a}\int_t^T (1 - e^{-a(s-t)})ds - r_t \int_t^T e^{-a(s-t)} ds - \sigma \int_t^T e^{-as} \int_t^s e^{au} dW_u ds \\ &= -\frac{b}{a}\int_t^T (1 - e^{-a(T-s)})ds - r_t \int_t^T e^{-a(T-s)} ds - \sigma \int_t^T e^{-as} \int_t^s e^{au} dW_u ds \\ &= -b\int_t^T B(s,T)ds - r_t \int_t^T e^{-a(T-s)} ds - \sigma \int_t^T e^{-as} \int_t^s e^{au} dW_u ds. \end{aligned}$$

Put $B(t,T) = \frac{1-e^{-a(T-t)}}{a}$.

$$\begin{aligned} -\frac{b}{a}\int_t^T (1 - e^{-a(s-t)})ds &= -\frac{b\Big(a(T-t) + e^{-a(T-t)} - 1\Big)}{a^2} \\ &= -\frac{b}{a}(T-t) + \frac{b}{a}B(t,T) \\ &= \frac{b}{a}\big(B(t,T) - (T-t)\big). \end{aligned}$$

$$\begin{aligned} -r_t \int_t^T e^{-a(s-t)} ds &= \frac{r_t}{a}\Big(e^{-a(T-t)} - 1\Big) \\ &= -B(t,T) r_t. \end{aligned}$$

$$\begin{aligned} -\sigma \int_t^T e^{-as} \int_t^s e^{au} dW_u ds &= -\sigma \int_t^T \int_t^s e^{-a(s-u)} dW_u ds \\ &= -\sigma \int_t^T \int_u^T e^{-a(s-u)} ds dW_u \\ &= -\sigma \int_t^T -\frac{(e^{-a(T-u)} - 1)}{a} dW_u \\ &= \frac{\sigma}{a}\int_t^T (e^{-a(T-u)} - 1) dW_u. \end{aligned}$$

Let $\gamma_u = \frac{\sigma}{a}(e^{-a(T-u)} - 1) = -\sigma B(u,T)$.

$$\exp\Big(\int_t^T \gamma_s dW_s\Big) = \exp\Big(\int_t^T \gamma_s dW_s - \frac{1}{2}\int_t^T \gamma^2 ds + \frac{1}{2}\int_t^T \gamma^2 ds\Big).$$

Note that

$$Y_t := \exp\Big(\int_t^T \gamma_s dW_s - \frac{1}{2}\int_t^T \gamma^2 ds\Big)$$

is a martingale.

$$\begin{aligned} dY_t &= Y_t(-\gamma_t dW_t + \frac{1}{2}\gamma_t^2 dt) + \frac{1}{2}Y_t\ d\langle Y\rangle_t \\ &= Y_t(-\gamma_t dW_t + \frac{1}{2}\gamma_t^2 dt) + \frac{1}{2}Y_t\ \big(-\gamma_t^2 dt\big) \quad \gamma_t^2 \text{ is decreasing,} \\ &= -Y_t\gamma_t dW_t. \end{aligned}$$

$$\begin{aligned} E_{\widetilde{Q}}\Big[\exp\Big(\int_t^T \gamma dW_s\Big)\Big] &= E_{\widetilde{Q}}\Big[\exp\Big(\frac{1}{2}\int_t^T \gamma^2 ds\Big)\Big] \\ &= \exp\Big(\frac{1}{2}\int_t^T \gamma^2 ds\Big), \quad \Big[= \exp\Big(\frac{1}{2}\sigma^2\int_t^T B^2(s,T)ds\Big)\Big], \\ \frac{1}{2}\int_t^T \gamma^2 ds &= \frac{\sigma^2}{4a^2}\,\frac{-3 - e^{-2a(T-t)} + 4e^{-a(T-t)} + 2a(T-t)}{a} \\ &= \frac{\sigma^2}{4a^2}\,\frac{-1 - e^{-2a(T-t)} + 2e^{-a(T-t)} + 2e^{-a(T-t)} - 2 + 2a(T-t)}{a} \\ &= -\frac{\sigma^2}{4a}B(t,T)^2 - \frac{\sigma^2}{2a^2}B(t,T) + \frac{\sigma^2}{2a^2}(T-t) \\ &= -\frac{\sigma^2}{4a}B(t,T)^2 - \frac{\sigma^2}{2a^2}\big(B(t,T) - (T-t)\big). \end{aligned}$$

Put

$$A(t,T) = \frac{\big(B(t,T) - (T-t)\big)\big(ab - \frac{\sigma^2}{2}\big)}{a^2} - \frac{\sigma^2 B(t,T)^2}{4a}.$$

Then

$$P(t,T) = \exp\Big(A(t,T) - B(t,T)r_t\Big)$$

✎ **Direct Calculation(Ornstein Uhlenbeck process):**

$$\begin{aligned} -\sigma\int_t^T e^{-as}\int_t^s e^{au}dW_u ds &= \frac{\sigma}{a}\int_t^T (e^{-a(T-u)} - 1)dW_u \\ &= \int_t^T \gamma_s dW_s \\ &= \sigma\int_t^T B(s,T)ds. \end{aligned}$$

To calculate this, let us change the time. Let $X_t = \int_t^T \gamma_s dW_s$.

$$
\begin{aligned}
dX_t &= -\gamma_t dW_t, \\
d\langle X\rangle_t &= \gamma_t^2 dt, \\
\langle X\rangle_t &= -\int_t^T \gamma_s^2 ds.
\end{aligned}
$$

Then, by Lemma (2.2.9),

$$
\begin{aligned}
X_t &= W_{\langle X\rangle_t} = W\Big(\langle X\rangle_t\Big), \\
E[\exp(X_t)] &= \exp\Big(\frac{1}{2}\langle X\rangle_t\Big) = \exp\Big(\frac{1}{2}\int_t^T \gamma_s^2 ds\Big).
\end{aligned}
$$

Example 10.11.7. Under the Vasicek model

$$
dr = (b - ar)dt + \sigma dW,
$$

from Example (10.9.1) we have that

$$
\begin{aligned}
H(t) &= -\int_0^t a ds = -at, \\
r_t &= e^{-at} r_0 + \int_0^t e^{-a(t-s)} b ds + \int_0^t e^{-a(t-s)} \sigma dW_s \\
&= e^{-at} r_0 + \frac{b}{a}(1 - e^{-at}) + \sigma \int_0^t e^{-a(t-s)} dW_s.
\end{aligned}
$$

Futhermore we can see that

$$
r_t \sim N\Big(e^{-at} r_0 + \frac{b}{a}(1 - e^{-at}), \frac{\sigma^2}{2a}(1 - e^{-2at})\Big).
$$

□

10.11.2 Cox-Ingersoll-Ross Model(Affine Term Structure)

Bond price with CIR model

The Cox-Ingersoll-Ross model is the simplest one which avoids negative interest rate.

$$dr = (b - ar)dt + \sigma\sqrt{r}d\widetilde{W}.$$

In Equation (10.11.3), set $\alpha = -a, \beta = b, \gamma = \sigma^2, \delta = 0$. By Theorem (10.11.2), we see that

$$1 + B_t(t,T) - a(t)B(t,T) - \frac{1}{2}\sigma^2 B^2(t,T) = 0, \quad B(T,T) = 0, \tag{10.11.8}$$

$$A_t(t,T) = bB(t,T), \quad A(T,T) = 0. \tag{10.11.9}$$

✎ **Solving Equation (10.11.8)**

Let $\gamma = \sqrt{a^2 + 2\sigma^2}$.

1. $B_1(t) = \frac{-a+\gamma}{\sigma^2}$ is one particular solution of Equation 10.11.8.
2. $B(t) = B_1(t) + \frac{1}{v(t)}$ with

$$v'(t) = -\gamma v(t) - \frac{\sigma^2}{2}, \quad v(T) = \frac{\sigma^2}{a-\gamma}.$$

3. we can easily see that

$$v(t) = -\frac{\sigma^2}{2\gamma} + \Big(\frac{\sigma^2}{2\gamma} + \frac{\sigma^2}{a-\gamma}\Big)e^{\gamma(T-t)}.$$

4. The solution is given by

$$\begin{aligned}
B(t,T) &= \frac{-a+\gamma}{\sigma^2} + \frac{1}{-\frac{\sigma^2}{2\gamma} + \Big(\frac{\sigma^2}{2\gamma} + \frac{\sigma^2}{a-\gamma}\Big)e^{\gamma(T-t)}} \\
&= \frac{-a+\gamma}{\sigma^2} + \frac{2\gamma}{\sigma^2}\cdot\frac{1}{-1 + \Big(1 + \frac{2\gamma}{a-\gamma}\Big)e^{\gamma(T-t)}} \\
&= \frac{a - \gamma + (-a + \gamma - 2\gamma)e^{\gamma(T-t)} + 2\gamma}{\sigma^2\Big(-1 + \Big(1 + \frac{2\gamma}{a-\gamma}\Big)e^{\gamma(T-t)}\Big)} \\
&= \frac{a + \gamma + (-a-\gamma)e^{\gamma(T-t)}}{-\sigma^2 + \sigma^2\frac{a+\gamma}{a-\gamma}e^{\gamma(T-t)}} \\
&= \frac{2\big(e^{\gamma(T-t)} - 1\big)}{\frac{2\sigma^2}{a+\gamma} - \frac{2\sigma^2}{a-\gamma}e^{\gamma(T-t)}} \\
&= \frac{2\big(e^{\gamma(T-t)} - 1\big)}{-(a-\gamma) + (a+\gamma)e^{\gamma(T-t)}}
\end{aligned}$$

$$= \frac{2\big(e^{\gamma(T-t)}-1\big)}{2\gamma+(a+\gamma)(e^{\gamma(T-t)}-1)}.$$

✎ **Solving Equation (10.11.9)**

1.

$$\begin{aligned} A(t,T) &= -b\int_t^T B(s,T)ds. \\ &= -b\int \frac{2\big(e^{\gamma(T-s)}-1\big)}{2\gamma+(a+\gamma)(e^{\gamma(T-s)}-1)}ds \end{aligned}$$

2. Let $x = e^{\gamma(T-s)}-1$. Then $ds = -\frac{1}{\gamma(1+x)}dx$

$$\begin{aligned} A(t,T) &= -b\int_0^{e^{\gamma(T-t)}-1} \frac{2x}{(\gamma+a)x+2\gamma}\cdot\frac{1}{\gamma(1+x)}dx \\ &= -\frac{2b}{\gamma}\int_0^{e^{\gamma(T-t)}-1}\left\{\frac{1}{(a-\gamma)(x+1)}+\frac{\gamma}{\sigma^2 x+\gamma(a-\gamma)}\right\}dx \\ &= -\frac{2b}{\gamma}\left[\frac{1}{a-\gamma}(T-t)\gamma+\frac{\gamma}{\sigma^2}\log\frac{e^{\gamma(T-t)}-1+\frac{\gamma(\gamma-a)}{\sigma^2}}{\frac{\gamma(\gamma-a)}{\sigma^2}}\right] \\ &= -\frac{2b}{\sigma^2}\left[\frac{\sigma^2}{a-\gamma}(T-t)+\log\left\{\frac{\sigma^2}{\gamma(\gamma-a)}\big(e^{\gamma(T-t)}-1\big)+1\right\}\right] \\ &= -\frac{2b}{\sigma^2}\left[\frac{a+\gamma}{2}(T-t)+\log\left\{\frac{a+\gamma}{2\gamma}\big(e^{\gamma(T-t)}-1\big)+1\right\}\right] \\ &= -\frac{2b}{\sigma^2}\log\left[\frac{\frac{a+\gamma}{2\gamma}\big(e^{\gamma(T-t)}-1\big)+1}{e^{\frac{(a+\gamma)(T-t)}{2}}}\right]. \end{aligned}$$

3.

$$\exp A(t,T) = \left[\frac{2\gamma e^{\frac{(a+\gamma)(T-t)}{2}}}{(a+\gamma)\big(e^{\gamma(T-t)}-1\big)+2\gamma}\right]^{\frac{2b}{\sigma^2}}.$$

Now we can see that

$$P(t,T) = \left[\frac{2\gamma e^{\frac{(a+\gamma)(T-t)}{2}}}{(a+\gamma)\big(e^{\gamma(T-t)}-1\big)+2\gamma}\right]^{\frac{2b}{\sigma^2}} \exp\left[-\frac{2\big(e^{\gamma(T-t)}-1\big)}{2\gamma+(a+\gamma)(e^{\gamma(T-t)}-1)}r_t\right].$$

Recall that

$$e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!} \cong 1+x, \quad \text{if } x \ll 1.$$

One can write

$$\begin{aligned}
\exp A(t,T) &\cong \left[\frac{2\gamma(1+\frac{(a+\gamma)(T-t)}{2})}{(a+\gamma)\big(1+\gamma(T-t)-1\big)+2\gamma}\right]^{\frac{2b}{\sigma^2}} = 1 \quad (T-t\ll 1). \\
B(t,T) &\cong \frac{2\gamma(T-t)}{2\gamma+(a+\gamma)(\gamma(T-t))} = \frac{T-t}{1+\frac{a+\gamma}{2}(T-t)} \cong (T-t) \quad (T-t\ll 1).
\end{aligned}$$

This shows, as one should intuitively expect, that for the short times the discounted bond price approximately is

$$P(t,T) \cong e^{r_t(T-t)}.$$

Using Equation (10.6.1) yields

$$R(t,T) = -\frac{1}{T-t}A(t,T) + \frac{1}{T-t}B(t,T)r(t).$$

CIR Short Rate Process

Let X_j be the solution to the stochastic differential equation

$$dX_j(t) = -\frac{1}{2}\beta X_j(t)dt + \frac{1}{2}\sigma dW_j(t).$$

X_j is called the Ornstein-Uhlenbeck process. The solution to this stochastic differential equation is, by Example (2.2.6)

$$X_j(t) = e^{-\frac{1}{2}\beta t}\Big[X_j(0) + \frac{1}{2}\sigma\int_0^t e^{\frac{1}{2}\beta s}dW_j(s)\Big].$$

Define

$$r(t) := X_1^2(t) + X_2^2(t) + \cdots + X_n^2(t).$$

Ito's formula implies

$$\begin{aligned}
dr(t) &= \sum_{i=1}^n 2X_i\,dX_i + \frac{1}{2}\sum_{i=1}^n 2dX_i\cdot dX_i \\
&= \sum_{i=1}^n 2X_i\Big(-\frac{1}{2}\beta X_j(t)dt + \frac{1}{2}\sigma dW_j(t)\Big) + \sum_{i=1}^n \frac{1}{4}\sigma^2 dt \\
&= -\beta r(t)dt + \sigma\sum_{i=1}^n X_i dW_i + \frac{n\sigma^2}{4}dt \\
&= \Big(\frac{n\sigma^2}{4} - \beta r(t)\Big)dt + \sigma\sqrt{r(t)}\sum_{i=1}^n \frac{X_i}{\sqrt{r(t)}}dW_i.
\end{aligned}$$

Define

$$W(t) = \sum_{i=1}^n \int_0^t \frac{X_i(s)}{\sqrt{r(s)}}dW_i(s).$$

Then W is a martingale,

$$dW = \sum_{i=1}^{n} \frac{X_i(t)}{\sqrt{r(t)}} dW_i(t), \quad dW \cdot dW = \sum_{i=1}^{n} \frac{X_i^2}{r} dt = dt,$$

so W is a Brownian motion. We have

$$dr(t) = \Big(\frac{n\sigma^2}{4} - \beta r(t)\Big) dt + \sigma \sqrt{r(t)} dW(t).$$

10.11.3 Ho and Lee Model(Affine Term Structure)

The short rate is given by the SDE:

$$dr_t = \theta_t dt + \sigma d\widetilde{W}_t,$$

for some θ_t deterministic and bounded, and σ constant.
By Theorem (10.11.2),

$$1 + B_t(t,T) = 0, \quad B(T,T) = 0.$$

Hence $B(t,T) = T - t$. From this, we have

$$\begin{aligned} A_t(t,T) &= \theta_t(T-t) - \frac{1}{2}\sigma^2(T-t)^2, \quad A(T,T) = 0, \\ A(t,T) &= -\int_t^T \theta_s(T-s)ds + \frac{1}{2}\sigma^2 \int_t^T (T-s)^2 ds \\ &= -\int_t^T \theta_s(T-s)ds + \frac{1}{6}\sigma^2(T-t)^3. \end{aligned}$$

Therefore

$$\begin{aligned} P(t,T) &= \exp(A(t,T) - B(t,T)r_t) \\ &= \exp\Big(-\int_t^T \theta_s(T-s)ds + \frac{1}{6}\sigma^2(T-t)^3 - (T-t)r_t \Big). \end{aligned}$$

Note that from (10.11.6) we have

$$\begin{aligned} f(0,T) &= B_T(0,T)r_0 - A_T(0,T) = r_0 + \int_0^T \theta_s ds - \frac{1}{2}\sigma^2 T^2, \\ \frac{\partial}{\partial T} f(0,T) &= \theta_T - \sigma^2 T, \\ \theta_t &= \frac{\partial}{\partial T} f(0,T) + \sigma^2 t. \end{aligned}$$

and

$$\begin{aligned} & \log\Big\{ \frac{P(0,t)}{P(0,T)} P(t,T) \Big\} \\ = \; & \log\Big\{ \frac{P(0,t)}{P(0,T)} \exp(A(t,T) - B(t,T)r_t) \Big\} \\ = \; & -\int_0^t \theta_s(t-s)ds + \frac{1}{6}\sigma^2 t^3 - tr_0 + \int_0^T \theta_s(T-s)ds - \frac{1}{6}\sigma^2 T^3 + Tr_0 \\ & -\int_t^T \theta_s(T-s)ds + \frac{1}{6}\sigma^2(T-t)^3 - (T-t)r_t \\ = \; & (T-t)\int_0^t \theta_s ds + \frac{1}{6}\Big(\sigma^2(T-t)^3 + t^3 - T^3\Big) - (T-t)(r_t - r_0) \\ = \; & (T-t)\Big(f(0,t) - f(0,0) + \frac{\sigma^2}{2}t^2\Big) + \frac{\sigma^2}{6}\Big((T-t)^3 + t^3 - T^3\Big) - (T-t)(r_t - r_0) \\ = \; & (T-t)f(0,t) - \frac{\sigma^2}{2} t(T-t)^2 - (T-t)r_t \end{aligned}$$

which give us

$$P(t,T) = \frac{P(0,T)}{P(0,t)} \exp\Big((T-t)f(0,t) - \frac{\sigma^2}{2}\, t(T-t)^2 - (T-t)r_t\Big).$$

Since $P(0,T), P(0,T)$ and $f(0,t)$ are observed from the market. If we write the observed data $P^*(0,T), P^*(0,T)$ and $f^*(0,t)$,

$$P(t,T) = \frac{P^*(0,T)}{P^*(0,t)} \exp\Big((T-t)f^*(0,t) - \frac{\sigma^2}{2} t(T-t)^2 - (T-t)r_t\Big).$$

Theorem 10.11.8. If we denote the $\Delta t-$period interest rate at time t by $R(t)$, i.e $R(t) = R(t, t{+}\Delta t)$, then

$$P(t,T) = e^{\widehat{A}(t,T) - R(t)(T-t)}$$

where,

$$\widehat{A}(t,T) = \log \frac{P(0,T)}{P(0,t)} - \frac{T-t}{\Delta t} \log \frac{P(0,t+\Delta t)}{P(0,t)} - \frac{1}{2}\sigma^2 t(T-t)\Big((T-t) - \Delta t\Big).$$

Proof. By the definition of $R(t)$,

$$\begin{aligned} P(t,t+\Delta t) &= e^{A(t,t+\Delta t) - r_t \Delta t} = e^{-R(t)\Delta t}, \\ R(t) &= -\frac{1}{\Delta t} A(t, t+\Delta t) + r_t. \\ &= -\frac{1}{\Delta t} \log \frac{P(0,t+\Delta t)}{P(0,t)} - f(0,t) + \frac{\sigma^2}{2}\, t(\Delta t) + r_t. \end{aligned}$$

Since

$$P(t,T) = e^{A(t,T) - r_t(T-t)} = e^{A(t,T) - r_t(T-t) + R(t)(T-t)} e^{-R(t)(T-t)}$$

we have that

$$\begin{aligned} \widehat{A}(t,T) &= A(t,T) - r_t(T-t) + R(t)(T-t) \\ &= \log \frac{P(0,T)}{P(0,t)} + \cancel{(T-t)f(0,t)} - \frac{\sigma^2}{2}\, t(T-t)^2 - \cancel{r_t(T-t)} \\ &\quad - \frac{T-t}{\Delta t} \log \frac{P(0,t+\Delta t)}{P(0,t)} - \cancel{(T-t)f(0,t)} + \frac{\sigma^2}{2}\, t(\Delta t)(T-t) + \cancel{r_t(T-t)} \\ &= \log \frac{P(0,T)}{P(0,t)} - \frac{T-t}{\Delta t} \log \frac{P(0,t+\Delta t)}{P(0,t)} - \frac{1}{2}\sigma^2 t(T-t)\Big((T-t) - \Delta t\Big). \end{aligned}$$

■

Theorem(10.11.8) demonstrates that we do not require the initial zero curve to be differentiable.

Example 10.11.9 (Forward Rate v.s. Futures Rate). We define the forward rate and future rate as follows. For a fixed $\Delta t = T_2 - T_1$,

$$\text{forward rate} = f(0, T_1, T_2)$$

$$\text{futures rate} \quad = \quad E[R(T_1, T_2)].$$

From

$$P(T_1, T_2) \quad = \quad e^{-R(T_1,T_2)(T_2-T_1)}$$

taking log gives us

$$\begin{aligned}
R(T_2, T_1) &= -\frac{1}{T_2 - T_1} \log P(T_1, T_2) \\
&= -\frac{1}{T_2 - T_1} \log \frac{P(0, T_2)}{P(0, T_1)} - f(0, T_1) + \frac{\sigma^2}{2} T_1(T_2 - T_1) + r_{T_1} \\
&= f(0, T_1, T_2) - f(0, T_1) + \frac{\sigma^2}{2} T_1(T_2 - T_1) + f(0, T_1) + \sigma W_{T_1} + \frac{1}{2}\sigma^2 T_1^2 \\
&= f(0, T_1, T_2) + \frac{\sigma^2}{2} T_1 T_2 + \sigma W_{T_1}.
\end{aligned}$$

After taking the expectation, we have

$$\text{futures rate} \quad = \quad \text{forward rate} + \frac{\sigma^2}{2} T_1 T_2$$

□

10.11.4 (Simplified)Hull-White Model(Affine Term Structure, Extended Vasicek)

The short rate dynamics are given by

$$dr_t = (\Phi(t) - ar_t)dt + \sigma d\widetilde{W}_t.$$

By Theorem (10.11.2),

$$\begin{aligned} B_t(t,T) &= aB(t,T) - 1, \quad B(T,T) = 0, && (10.11.10) \\ A_t(t,T) &= \Phi(t)B(t,T) - \frac{1}{2}\sigma^2 B^2(t,T), \quad A(T,T) = 0. && (10.11.11) \end{aligned}$$

The solution to this system is given by

$$\begin{aligned} B(t,T) &= \frac{1 - e^{-a(T-t)}}{a}, \\ A(t,T) &= \int_t^T \Big(\frac{1}{2}\sigma^2 B^2(s,T) - \Phi(s)B(s,T)\Big)ds. \end{aligned}$$

Now let us find Φ. From Equation (10.11.6),

$$\begin{aligned} f(0,T) &= B_T(0,T)r_0 - A_T(0,T) \\ &= e^{-aT}r_0 - \int_0^T \Big(\sigma^2 aB(s,T)e^{-a(T-s)} - \Phi(s)e^{-a(T-s)}\Big)ds \\ &= e^{-aT}r_0 + \frac{\sigma^2}{2}B^2(s,T)\Big|_0^T + \int_0^T \Phi(s)e^{-a(T-s)}ds \\ &= e^{-aT}r_0 - \frac{\sigma^2}{2}B^2(0,T) + \int_0^T \Phi(s)e^{-a(T-s)}ds \\ &= e^{-aT}r_0 - \frac{\sigma^2}{2a^2}(1 - e^{-aT})^2 + \int_0^T \Phi(s)e^{-a(T-s)}ds. && (10.11.12) \end{aligned}$$

Given our observed forward rate curve $f^*(0,T)$, defined by

$$f^*(0,T) = -\frac{\partial}{\partial T}\log P^*(0,T),$$

we now look for a function Φ.

Lemma 10.11.10.

$$\Phi(T) = f_T^*(0,T) + af^*(0,T) + \frac{\sigma^2}{2a}(1 - e^{-2aT}).$$

Proof. Let

$$\begin{aligned} x(T) &= e^{-aT}r_0 + \int_0^T \Phi(s)e^{-a(T-s)}ds, \\ g(T) &= \frac{\sigma^2}{2a^2}(1 - e^{-aT})^2. \end{aligned}$$

Then $f^*(0,T) = x(T) - g(T)$ and $x'(T) = -ax(T) + \Phi(T), x(0) = r(0)$. Hence

$$\begin{aligned}
\Phi(T) &= x'(T) + ax(T) \\
&= f_T^*(0,T) + g'(T) + ax(T) \\
&= f_T^*(0,T) + g'(T) + a\Big(f^*(0,T) + g(T)\Big) \\
&= f_T^*(0,T) + af^*(0,T) + g'(T) + ag(T) \\
&= f_T^*(0,T) + af^*(0,T) + \frac{\sigma^2}{2a}(1 - e^{-2aT}).
\end{aligned}$$

■

Theorem 10.11.11 (Bond Price under Hull-White Model).

$$P(t,T) = \frac{P(0,T)}{P(0,t)} \exp\Big\{B(t,T)f(0,t) - \frac{\sigma^2}{4a}B^2(t,T)(1 - e^{-2at}) - B(t,T)r_t\Big\}.$$

Proof. From

$$\begin{aligned}
P(t,T) &= \exp\Big(A(t,T) - B(t,T)r_t\Big) \\
&= \exp\bigg(\int_t^T \Big(\frac{1}{2}\sigma^2 B^2(s,T) - \Phi(s)B(s,T)\Big)ds - B(t,T)r_t\bigg),
\end{aligned}$$

we have that

$$\begin{aligned}
& \log\Big\{\frac{P(0,t)}{P(0,T)}P(t,T)\Big\} \\
= & -\int_0^t \Phi(s)B(s,t)ds + \int_0^T \Phi(s)B(s,T)ds - \int_t^T \Phi(s)B(s,T)ds \\
& + \int_0^t \frac{\sigma^2}{2}B^2(s,t)ds - \int_0^T \frac{\sigma^2}{2}B^2(s,T)ds + \int_t^T \frac{\sigma^2}{2}B^2(s,T)ds \\
& -B(0,t)r_0 + B(0,T)r_0 - B(t,T)r_t \\
= & -\int_0^t \Phi(s)B(s,t)ds + \int_0^t \Phi(s)B(s,T)ds + \int_0^t \frac{\sigma^2}{2}B^2(s,t)ds - \int_0^t \frac{\sigma^2}{2}B^2(s,T)ds \\
& -B(0,t)r_0 + B(0,T)r_0 - B(t,T)r_t \\
= & -\int_0^t \Phi(s)\big(B(s,t) - B(s,T)\big)ds + \frac{\sigma^2}{2}\int_0^t \big(B^2(s,t) - B^2(s,T)\big)ds \\
& -\big(B(0,t) - B(0,T)\big)r_0 - B(t,T)r_t.
\end{aligned}$$

1. To get the first part, consider that

$$\begin{aligned}
e^{as}\Phi(s) &= e^{as}f_T^*(0,s) + ae^{as}f^*(0,s) + e^{as}\frac{\sigma^2}{2a}(1 - e^{-2as}) \\
&= \frac{\partial}{\partial T}\{e^{as}f^*(0,s)\} + \frac{\sigma^2}{2a}(e^{as} - e^{-as}).
\end{aligned}$$

From this we get the followings:

$$-\int_0^t \Phi(s)\big(B(s,t) - B(s,T)\big)ds$$

$$
\begin{aligned}
&= e^{-at}B(t,T)\int_0^t \Phi(s)e^{as}ds \\
&= e^{-at}B(t,T)\big(e^{at}f(0,t)-f(0,0)+\frac{\sigma^2}{2a}\int_0^t(e^{as}-e^{-as})ds\big) \\
&= B(t,T)f(0,t)-r_0e^{-at}B(t,T)+\frac{\sigma^2}{2a^2}\big(e^{at}+e^{-at}-2\big)e^{-at}B(t,T).
\end{aligned}
$$

2. Part two:

$$
\begin{aligned}
&\frac{\sigma^2}{2}\int_0^t\big(B^2(s,t)-B^2(s,T)\big)ds \\
=\;& \frac{\sigma^2}{2}\int_{B(0,t)}^0\frac{y^2}{ay-1}dy-\frac{\sigma^2}{2}\int_{B(0,T)}^{B(t,T)}\frac{y^2}{ay-1}dy \\
=\;& \frac{\sigma^2}{2}\Big[\frac{y^2}{2a}+\frac{y}{a^2}+\frac{\log|1-ay|}{a^3}\Big]_{B(0,t)}^0-\frac{\sigma^2}{2}\Big[\frac{y^2}{2a}+\frac{y}{a^2}+\frac{\log|1-ay|}{a^3}\Big]_{B(0,T)}^{B(t,T)} \\
=\;& \frac{\sigma^2}{2}\Big(-\frac{1}{2a}B^2(0,t)-\frac{1}{a^2}B(0,t)-\frac{t}{a^2}\Big) \\
&-\frac{\sigma^2}{2}\Big(\frac{1}{2a}B^2(t,T)-\frac{1}{2a}B^2(0,T)+\frac{1}{a^2}B(t,T)-\frac{1}{a^2}B(0,T)+\frac{1}{a^2}(T-t-T)\Big) \\
=\;& \frac{\sigma^2}{2}\Big(-\frac{1}{2a}B^2(0,t)-\frac{1}{a^2}B(0,t)-\frac{1}{2a}B^2(t,T)+\frac{1}{2a}B^2(0,T)-\frac{1}{a^2}B(t,T)+\frac{1}{a^2}B(0,T)\Big).
\end{aligned}
$$

3. Part Three:

$$
-\big(B(0,t)-B(0,T)\big)r_0-B(t,T)r_t \;=\; e^{-at}B(t,T)r_0-B(t,T)r_t.
$$

Hence we have

$$
\begin{aligned}
&\log\Big\{\frac{P(0,t)}{P(0,T)}P(t,T)\Big\} \\
=\;& B(t,T)f(0,t)-B(t,T)r_t+\frac{\sigma^2}{2a^2}\big(e^{at}+e^{-at}-2\big)e^{-at}B(t,T) \\
&+\frac{\sigma^2}{2}\Big(-\frac{1}{2a}B^2(0,t)-\frac{1}{a^2}B(0,t)-\frac{1}{2a}B^2(t,T)+\frac{1}{2a}B^2(0,T)-\frac{1}{a^2}B(t,T)+\frac{1}{a^2}B(0,T)\Big) \\
=\;& B(t,T)f(0,t)-\frac{\sigma^2}{4a}B^2(t,T)(1-e^{-2at})-B(t,T)r_t.
\end{aligned}
$$

This completes the proof. ■

Theorem 10.11.12. If we denote the $\Delta t-$period interest rate at time t by $R(t)$, i.e $R(t)=R(t,t+\Delta t)$, then

$$
P(t,T) \;=\; e^{\widehat{A}(t,T)-\widehat{B}(t,T)R(t)}
$$

where,

$$
\widehat{A}(t,T) \;=\; \log\frac{P(0,T)}{P(0,t)}-\frac{B(t,T)}{B(t,t+\Delta t)}\log\frac{P(0,t+\Delta t)}{P(0,t)}
$$

$$
\begin{aligned}
&-\frac{\sigma^2}{4a}(1-e^{-2at})B(t,T)\Big[B(t,T)-B(t,t+\Delta t)\Big],\\
\widehat{B}(t,T) \quad &= \quad \frac{B(t,T)}{B(t,t+\Delta t)}\Delta t.
\end{aligned}
$$

Proof. By the definition of $R(t)$,

$$
\begin{aligned}
P(t,t+\Delta t) \quad &= \quad e^{A(t,t+\Delta t)-B(t,t+\Delta t)r_t} \quad = \quad e^{-R(t)\Delta t},\\
R(t) \quad &= \quad -\frac{1}{\Delta t}A(t,t+\Delta t)+\frac{1}{\Delta t}B(t,t+\Delta t)r_t.\\
&= \quad -\frac{1}{\Delta t}\log\frac{P(0,t+\Delta t)}{P(0,t)}-\frac{1}{\Delta t}B(t,t+\Delta t)f(0,t)\\
&\quad +\frac{\sigma^2}{4a\,\Delta t}B(t,t+\Delta t)^2(1-e^{-2at})+\frac{1}{\Delta t}B(t,t+\Delta t)r_t.
\end{aligned}
$$

Put $\widehat{B}(t,T):=\frac{B(t,T)}{B(t,t+\Delta t)}\Delta t$. Then we get

$$
\begin{aligned}
R(t) \quad &= \quad -\frac{1}{\Delta t}\log\frac{P(0,t+\Delta t)}{P(0,t)}-\frac{1}{\Delta t}B(t,t+\Delta t)f(0,t)\\
&\quad +\frac{\sigma^2}{4a\,\Delta t}B(t,t+\Delta t)^2(1-e^{-2at})+\frac{B(t,T)}{\widehat{B}(t,T)}r_t,\\
\widehat{B}(t,T)R(t) \quad &= \quad -\frac{B(t,T)}{B(t,t+\Delta t)}\log\frac{P(0,t+\Delta t)}{P(0,t)}-B(t,T)f(0,t)\\
&\quad +\frac{\sigma^2}{4a}B(t,t+\Delta t)B(t,T)(1-e^{-2at})+B(t,T)r_t.
\end{aligned}
$$

Since

$$
P(t,T) \quad = \quad e^{A(t,T)-B(t,T)r_t} \quad = \quad e^{A(t,T)-B(t,T)r_t+\widehat{B}(t,T)R(t)}e^{-\widehat{B}(t,T)R(t)}
$$

we have that

$$
\begin{aligned}
\widehat{A}(t,T) \quad &= \quad A(t,T)-B(t,T)r_t+\widehat{B}(t,T)R(t)\\
&= \quad A(t,T)-\frac{B(t,T)}{B(t,t+\Delta t)}\log\frac{P(0,t+\Delta t)}{P(0,t)}-B(t,T)f(0,t)+\frac{\sigma^2}{4a}B(t,t+\Delta t)B(t,T)(1-e^{-2at})\\
&= \quad \log\frac{P(0,T)}{P(0,t)}+\cancel{B(t,T)f(0,t)}-\frac{\sigma^2}{4a}B^2(t,T)(1-e^{-2at})\\
&\quad -\frac{B(t,T)}{B(t,t+\Delta t)}\log\frac{P(0,t+\Delta t)}{P(0,t)}-\cancel{B(t,T)f(0,t)}+\frac{\sigma^2}{4a}B(t,t+\Delta t)B(t,T)(1-e^{-2at})\\
&= \quad \log\frac{P(0,T)}{P(0,t)}-\frac{B(t,T)}{B(t,t+\Delta t)}\log\frac{P(0,t+\Delta t)}{P(0,t)}-\frac{\sigma^2}{4a}B(t,T)(1-e^{-2at})\Big[B(t,T)-B(t,t+\Delta t)\Big].
\end{aligned}
$$

■

Example 10.11.13. From Example (10.9.1)

$$
H(t) \quad = \quad -\int_0^t a\,ds \quad = \quad -at,
$$

and

$$\begin{aligned} r_t &= e^{-at}r_0 + \int_0^t e^{-a(t-s)}\Phi(s)ds + \sigma\int_0^t e^{-a(t-s)}dW_s \\ &= f(0,t) + \frac{\sigma^2}{2a^2}(1-e^{-at})^2 + \sigma\int_0^t e^{-a(t-s)}dW_s \quad \text{by (10.11.12).} \end{aligned}$$

By Lemma (2.2.9),

$$\sigma\int_0^t e^{-a(t-s)}dW_s \sim N\left(0, \frac{\sigma^2}{2a}(1-e^{-2at})\right).$$

Thus we can see that under risk neutral measure($\widetilde{Q}$) r_t is normally distributed with mean $f(0,t) + \frac{\sigma^2}{2a^2}(1-e^{-at})^2$ and variance $\frac{\sigma^2}{2a}(1-e^{-2at})$, i.e.

$$r_t \sim N\left(f(0,t) + \frac{\sigma^2}{2a^2}(1-e^{-at})^2, \frac{\sigma^2}{2a}(1-e^{-2at})\right).$$

Similarly for $u < t$ and given r_u we have

$$\begin{aligned} r_t &= e^{-a(t-u)}r_u + \int_u^t e^{-a(t-s)}\Phi(s)ds + \sigma\int_u^t e^{-a(t-s)}dW_s \\ &= f(u,t) + \frac{\sigma^2}{2a^2}(1-e^{-a(t-u)})^2 + \sigma\int_u^t e^{-a(t-s)}dW_s \end{aligned}$$

and

$$r_t \sim N\left(f(u,t) + \frac{\sigma^2}{2a^2}(1-e^{-a(t-u)})^2, \frac{\sigma^2}{2a}(1-e^{-2a(t-u)})\right).$$

□

Example 10.11.14 (Under Forward Measure). Under the forward measure Q^{T_F}, from Example (10.11.5) the short rate is given by

$$dr_t = \Big(\Phi(t) - \sigma^2 B(t,T_F) - ar_t\Big)dt + \sigma dW_t^{T_F}.$$

Since

$$\sigma^2\int_0^t e^{-a(t-s)}B(s,T_F)\,ds = \frac{\sigma^2}{a}\left(\frac{1-e^{-at}}{a} - \frac{e^{-a(T_F-t)} - e^{-a(T_F+t)}}{2a}\right)$$

and

$$\frac{\sigma^2}{2a^2}(1-e^{-at})^2 - \sigma^2\int_0^t e^{-a(t-s)}B(s,T_F)\,ds = -\frac{\sigma^2(1-e^{-2at})(1-e^{-a(T_F-t)})}{2a^2},$$

we have that under T_F-forward measure(Q^{T_F}) r_t is normally distributed, i.e.

$$r_t \sim N\left(f(0,t) - \frac{\sigma^2(1-e^{-2at})(1-e^{-a(T_F-t)})}{2a^2}, \frac{\sigma^2}{2a}(1-e^{-2at})\right).$$

Similarly for $u < t$ we obtain that

$$r_t \quad \sim \quad N\left(f(u,t) - \frac{\sigma^2(1-e^{-2a(t-u)})(1-e^{-a(T_F-t)})}{2a^2}, \frac{\sigma^2}{2a}(1-e^{-2a(t-u)})\right)$$

for given r_u. Also we can see that

$$\begin{aligned} E_{Q^{T_F}}[r_{T_F}] &= f(0,T_F), \\ E_{Q^{T_F}}[r_{T_F} \mid \mathscr{F}_u] &= f(u,T_F). \end{aligned}$$

□

10.12 Some Comments about One Factor Short Rate Models

Two limitations of the one factor short rate models are

1. They involve only one factor. This is, one source of uncertainty.
2. They do not give the user complete freedom in choosing the volatility structure. The models can be made to provide a perfect fit to volatilities observed in the market at time zero, but the user then has no control over the volatilities at subsequent times. Future volatility pattern are liable to be quite different from those observed in the market today.

Chapter 11

Heath-Jarrow-Morton Model

Up to this point we have studied interest models where the short rate r is the only explanatory variable. The main advantage with such models are:

1. Specifying r as the solution of an SDE allows us to use Markov process theory, so we may work within a PDE framework.
2. In particular it is often possible to obtain analytic formulas for bond prices and derivatives.

The main drawbacks of short rate models are as follows.

1. From an economic point of view it seems unreasonable to assume that the entire money market is governed by only one explanatory variable.
2. It is hard to obtain a realistic volatility structure for the forward rates without introduction a very complicating short rate models.
3. As the short rate model becomes more realistic, the inversion of the yield curve becomes increasingly more difficult.

11.1 Single-factor HJM

The Heath-Jarrow-Morton model is given by

$$df(t,T) \quad = \quad \alpha(t,T)dt + \sigma(t,T)dW_t. \tag{11.1.1}$$

In this equation and throughout, the differential df is with respect to t and not maturity T. The volatility $\sigma(t,T)$ and the drift $\alpha(t,T)$ can depend on the history of the Brownian motion W_t and on the rates themselves up to time t. We see that (11.1.1) is an infinite dimensional stochastic system(one equation for each fixed T).

By theorem (10.7.1)

$$dP(t,T) \quad = \quad P(t,T)\left(r_t + A(t,T) + \frac{1}{2}\Sigma(t,T)^2\right)dt + P(t,T)\Sigma(t,T)dW_t,$$

where

$$\begin{cases} A(t,T) &= -\int_t^T \alpha(t,s)ds, \\ \Sigma(t,T) &= -\int_t^T \sigma(t,s)ds. \end{cases} \tag{11.1.2}$$

By (10.7.9) and (11.1.1), we get

$$dP(t,T) = P(t,T)\Big(r_t dt + \Sigma(t,T)d\widetilde{W}_t\Big). \tag{11.1.3}$$

The drifts under physical measure are restricted by the need to be a simple change of measure away from a martingale. The expected return is r_t because a zero-coupon bond is a tradable security providing no income.

$$\begin{aligned} \gamma_t &= \frac{\left(r_t + A(t,T) + \frac{1}{2}\Sigma(t,T)^2\right) - r_t}{\Sigma(t,T)} = \frac{A(t,T) + \frac{1}{2}\Sigma(t,T)^2}{\Sigma(t,T)}, \\ A(t,T) &= -\frac{1}{2}\Sigma(t,T)^2 + \Sigma(t,T)\gamma_t. \end{aligned} \tag{11.1.4}$$

Note that γ_t is independent of maturity T. Differentiating (11.1.4) with respect to T, we see that $-\alpha(t,T) = \sigma(t,T)\Sigma(t,T) - \sigma(t,T)\gamma_t$, that is

$$\alpha(t,T) = \sigma(t,T)\Big(\gamma_t - \Sigma(t,T)\Big).$$

Then we have that

$$\begin{aligned} df(t,T) &= \alpha(t,T)dt + \sigma(t,T)(d\widetilde{W}_t - \gamma_t dt) \\ &= -\sigma(t,T)\Sigma(t,T)dt + \sigma(t,T)d\widetilde{W}_t, && (11.1.5) \\ f(t,T) &= f(0,T) - \int_0^t \sigma(s,T)\Sigma(s,T)ds + \int_0^t \sigma(s,T)d\widetilde{W}_s, && (11.1.6) \\ r_t &= f(0,t) - \int_0^t \sigma(s,t)\Sigma(s,t)ds + \int_0^t \sigma(s,t)d\widetilde{W}_s. \end{aligned}$$

Remark 11.1.1 (Multi-Factor Case). Note that

$$df(t,T) = \left(\sigma(t,T)^T \int_t^T \sigma(t,u)du\right) dt + \sigma(t,T)^\top dW_t. \tag{11.1.7}$$

Using a subscript $j = 1, \cdots, d$ to indicate vector components, we can write as

$$\begin{aligned} df(t,T) &= \sum_{j=1}^d \left(\sigma_j(t,T)\int_t^T \sigma_j(t,u)du\right) dt + \sum_{j=1}^d \sigma_j(t,T)dW_j(t), \\ dP(t,T) &= P(t,T)\left(r_t dt + \sum_{j=1}^d - \int_t^T \sigma_j(t,u)du\; dW_j(t)\right). \end{aligned}$$

□

It follows from Eq. (11.1.5) and Theorem (10.7.1) that

$$dr_t = f_T(t,t)dt + \sigma(t,t)d\widetilde{W}_t$$

$$= \left[f_T(0,t) - \int_0^t \sigma_T(s,t)\Sigma(s,t)ds - \int_0^t \sigma(s,t)\Sigma_T(s,t)ds + \int_0^t \sigma_T(s,t)d\widetilde{W}_s\right]dt + \sigma(t,t)d\widetilde{W}_t$$
$$= \left[f_T(0,t) - \int_0^t \sigma_T(s,t)\Sigma(s,t)ds + \int_0^t \sigma^2(s,t)ds + \int_0^t \sigma_T(s,t)d\widetilde{W}_s\right]dt + \sigma(t,t)d\widetilde{W}_t. \quad (11.1.8)$$

It is necessary to examine Eq. (11.1.8). Equation

$$-\int_0^t \sigma_T(s,t)\Sigma(s,t)ds + \int_0^t \sigma^2(s,t)ds$$

depends on the history of Σ and equation

$$\int_0^t \sigma_T(s,t)d\widetilde{W}_s$$

depends on the history of both Σ and $\widetilde{W}$. These terms are liable to cause the process r_t to be non-Markov.

This highlights the key problem in implementing a general HJM model. We have to use Monte Carlo simulation. Trees create difficulties. This means that a tree for r is non-recombining. An up movement followed by a down movement does not lead to the same node as a down movement followed by an up movement.

Remark 11.1.2. In (11.1.5), the drift is determined once σ is specified. This contrasts with the dynamics of the short rate models where parameters of the drift could be specified independent of the diffusion coefficient introducing arbitrage.

1. Indeed, choosing parameters of the drift is essential in calibrating short rate models to an observed set of bond prices.
2. In contrast, an HJM model is automatically calibrated to an initial set of bond prices $P(0,T)$ if the initial forward curve $f(0,T)$ is simply chosen consistent with these bond prices through
$$f(t,T) = -\frac{\partial}{\partial T}\log P(t,T).$$
3. Put slightly differently, calibrating an HJM model to an observed set of bond prices is a matter of choosing an appropriate initial condition rather than choosing a parameter of the model dynamics.
4. The effort in calibrating an HJM model lies in choosing σ to match market prices of interest rate derivatives in addition to matching bond prices.

□

Example 11.1.3. Consider an HJM model with $\sigma(t,T) = \sigma\exp(-\alpha(T-t))$ for some constant $\sigma, \alpha > 0$. In this case, the diffusion term $\sigma(t,T)dW_t$ moves forward rates for short maturities more than forward rates for long maturities. The drift is given by

$$\sigma^2 e^{-\alpha(T-t)}\int_t^T e^{-\alpha(T-u)}du = \frac{\sigma^2}{\alpha}\left(e^{-2\alpha(T-t)} - e^{-\alpha(T-t)}\right)$$

□

Example 11.1.4. It is tempting to consider a specification of the form

$$\sigma(t,T) = \widetilde{\sigma}(t,T)f(t,T)$$

for some deterministic $\widetilde{\sigma}$ depending only on t and T. This would make $\widetilde{\sigma}(t,T)$ the volatility of the forward rate $f(t,T)$ and would suggest that the distribution of $f(t,T)$ is approximately lognormal. However this choice of σ is inadmissible. It produces forward rates that grow to infinity in finite time with positive probability. The difficulty, speaking loosely, is that if σ is proportional to the level of rates, then the drift is proportional to the rates squared. This violates the linear growth condition ordinarily required for the existence and uniqueness of solutions to SDEs. Market conventions often presuppose the existence of a proportional volatility for forward rates, so the failure of this example could be viewed as a shortcoming of the HJM framework. We will see that the difficulty can be avoided by working with simple rather than continuously compounding forward rate(BGM model). □

11.1.1 HJM in terms of the short rate

Suppose that r_t is a Markov diffusion(though not necessarily time homogeneous) with volatility $\rho(r_t,t)$ and drift $\nu(r_t,t)$. That is

$$dr_t = \nu(r_t,t)dt + \rho(r_t,t)d\widetilde{W}_t \tag{11.1.9}$$

where $\rho(x,t)$ and $\nu(x,t)$ are deterministic functions of space and time. Then Let

$$g(x,t,T) = -\log E_{\widetilde{Q}}\Big[\exp\Big(-\int_t^T r_s ds\Big)\Big|r_t = x\Big].$$

Then

$$\int_t^T f(t,u)du = -\log P(t,T) = g(r_t,t,T).$$

Theorem 11.1.5 (Short rate model in HJM terms). If the short rate, r_t, satisfies Equation (11.1.9) under martingale measure, we have

$$df(t,T) = -\sigma(t,T)\Sigma(t,T)dt + \sigma(t,T)d\widetilde{W}_t,$$

where

$$\sigma(t,T) = \rho(r_t,t)\frac{\partial^2 g}{\partial x\partial T}(r_t,t,T),$$
$$\Sigma(t,T) = -\rho(r_t,t)\frac{\partial g}{\partial x}(r_t,t,T).$$

In addition, the initial forward rate curve $f(0,T)$ is given by

$$f(0,T) = \frac{\partial g}{\partial T}(r_0,0,T).$$

Proof.

$$f(t,T) = \frac{\partial}{\partial T}g(r_t,t,T),$$

$$\begin{aligned} df(t,T) &= \frac{\partial^2 g}{\partial x \partial T} dr_t + \frac{\partial^2 g}{\partial t \partial T} dt + \frac{1}{2} \frac{\partial^3 g}{\partial x^2 \partial T} \rho^2(r_t,t) dt \\ &= \frac{\partial^2 g}{\partial x \partial T} \Big(\nu(r_t,t) dt + \rho(r_t,t) d\widetilde{W}_t \Big) + \frac{\partial^2 g}{\partial t \partial T} dt + \frac{1}{2} \frac{\partial^3 g}{\partial x^2 \partial T} \rho^2(r_t,t) dt. \end{aligned}$$

Thus

$$\sigma(t,T) = \rho(r_t,t) \frac{\partial^2 g}{\partial x \partial T}(r_t,t,T),$$

and by theorem (10.7.1)

$$\begin{aligned} \Sigma(t,T) &= -\int_t^T \sigma(t,u) du \\ &= -\int_t^T \rho(r_t,t) \frac{\partial^2 g}{\partial x \partial u}(r_t,t,u) \\ &= -\rho(r_t,t) \frac{\partial g}{\partial x}(r_t,t,T). \end{aligned}$$

■

Note that once the function $\Sigma(t,T)$ has been specified, the risk-neutral processed for the $f(t,T)$'s and $P(t,T)$'s are known. The $\Sigma(t,T)$'s are, therefore, sufficient to define fully a one-factor interest model.

11.1.2 HJM from Bond Price Dynamics

Assume that the risk-neutral process for $P(t,T)$ has the form

$$dP(t,T) = P(t,T) \Big(r_t dt + \Sigma(t,T) d\widetilde{W}_t \Big).$$

Because a bond's price volatility declines to zero at maturity, we must have

$$\Sigma(t,t) = 0.$$

From (10.6.3) and Ito's lemma, we have

$$\begin{aligned} f(t,T_1,T_2) &= -\frac{\log P(t,T_2) - \log P(t,T_1)}{T_2 - T_1}, \\ d \log P(t,T_1) &= \left(r_t - \frac{1}{2} \Sigma^2(t,T_1) \right) dt + \sigma(t,T_1) d\widetilde{W}_t \\ d \log P(t,T_2) &= \left(r_t - \frac{1}{2} \Sigma^2(t,T_2) \right) dt + \sigma(t,T_2) d\widetilde{W}_t. \end{aligned}$$

Then we have

$$df(t,T_1,T_2) = \frac{1}{2} \frac{\Sigma^2(t,T_2) - \Sigma^2(t,T_1)}{T_2 - T_1} dt + \frac{\Sigma(t,T_1) - \Sigma(t,T_2)}{T_2 - T_1} d\widetilde{W}_t \qquad (11.1.10)$$

When we put $T_1 = T$ and $T_2 = T + \Delta T$ in (11.1.10) and then take limits as ΔT tends to zero, $f(t, T_1, T_2)$ becomes $f(t, T)$,

$$df(t,T) = \frac{1}{2}\Sigma(t,T)\frac{\partial \Sigma(t,T)}{\partial T}dt - \frac{\partial \Sigma(t,T)}{\partial T}d\widetilde{W}_t.$$

Once the function $\Sigma(t, T)$ bas been specified, the risk-neutral process for the $f(t, T)$'s are known. The $\Sigma(t, T)$'s are, therefore, sufficient to define fully a one factor interest model.

11.1.3 Ho and Lee Model in HJM

The short rate is given by the SDE:

$$dr_t = \theta_t dt + \sigma d\widetilde{W}_t, \tag{11.1.11}$$

for some θ_t deterministic and bounded, and σ constant.
To get $P(t,T)$, first, consider $\int_t^T r_s ds$.

$$\begin{aligned}
r_s &= r_t + \sigma(\widetilde{W}_s - \widetilde{W}_t) + \int_t^T \theta_u du, \\
\int_t^T r_s ds &= r_t(T-t) + \sigma \int_t^T (\widetilde{W}_s - \widetilde{W}_t) ds + \int_t^T \int_t^s \theta_u du ds \\
&= r_t(T-t) + \sigma \Big\{ (\widetilde{W}_T - \widetilde{W}_t)T - \int_t^T s d\widetilde{W}_s \Big\} + \int_t^T \int_u^T \theta_u ds du \\
&= r_t(T-t) + \sigma \int_t^T (T-s) d\widetilde{W}_s + \int_t^T (T-u)\theta_u du.
\end{aligned}$$

$$\exp\Big(- \int_t^T r_s ds \big| r_t \Big) = \exp\Big(- r_t(T-t) \Big) \cdot \exp\Big(- \int_t^T (T-u)\theta_u du \Big) \cdot \exp\Big(- \sigma \int_t^T (T-s) d\widetilde{W}_s \Big)$$

Thus we have

$$\begin{aligned}
P(t,T) &= E_{\widetilde{Q}}\Big[\exp\Big(- \int_t^T r_s ds \big| r_t \Big) \Big] \\
&= \exp\Big(- r_t(T-t) \Big) \cdot \exp\Big(- \int_t^T (T-u)\theta_u du \Big) \cdot E_{\widetilde{Q}}\Big[\exp\Big(- \sigma \int_t^T (T-s) d\widetilde{W}_s \Big) \Big| r_t \Big].
\end{aligned}$$

Since

$$\begin{aligned}
& E_{\widetilde{Q}}\Big[\exp\Big(- \sigma \int_t^T (T-s) d\widetilde{W}_s \Big) \Big| r_t \Big] \\
&= E_{\widetilde{Q}}\Big[\exp\Big(- \sigma \int_t^T (T-s) d\widetilde{W}_s - \frac{1}{2}\sigma^2 \int_t^T (T-s)^2 ds + \frac{1}{2}\sigma^2 \int_t^T (T-s)^2 ds \Big) \Big| r_t \Big] \\
&= \exp\Big(\frac{1}{2}\sigma^2 \int_t^T (T-s)^2 ds \Big) \\
&= \exp\Big(\frac{1}{6}\sigma^2 (T-t)^3 \Big), \\
P(t,T) &= \exp\Big(- r_t(T-t) \Big) \cdot \exp\Big(- \int_t^T (T-u)\theta_u du \Big) \cdot \exp\Big(\frac{1}{6}\sigma^2 (T-t)^3 \Big), \\
g(x,t,T) &= x(T-t) + \int_t^T (T-u)\theta_u du - \frac{1}{6}\sigma^2 (T-t)^3.
\end{aligned}$$

Theorem 11.1.6 (Ho and Lee model in HJM terms).

$$\begin{aligned}
df_t &= \sigma^2 (T-t) dt + \sigma d\widetilde{W}_t, \quad \text{with} \\
f(0,T) &= r_0 + \int_0^T \theta_s ds - \frac{1}{2}\sigma^2 T^2, \quad \text{i.e. } \theta_T (= \theta(T)) = \frac{\partial f}{\partial T}(0,T) + \sigma^2 T,
\end{aligned} \tag{11.1.12}$$

$$\begin{aligned}
dP(t,T) &= P(t,T)\Big(r_t dt - \sigma(T-t)d\widetilde{W}_t\Big), \\
P(t,T) &= \exp\left(-\sigma(T-t)\widetilde{W}_t - \int_t^T f(0,u)du - \frac{1}{2}\sigma^2 T(T-t)t\right) \\
&= \frac{P(0,T)}{P(0,t)}\exp\Big(-\frac{1}{2}\sigma^2 T(T-t)t - \sigma(T-t)\widetilde{W}_t\Big) \\
&= \frac{P(0,T)}{P(0,t)}\exp\left((T-t)f(0,t) - \frac{1}{2}\sigma^2 t(T-t)^2 - (T-t)r_t\right).
\end{aligned}$$

Proof. It follows from Theorem (11.1.5). Also, we can get that

$$\begin{aligned}
f(t,T) &= f(0,T) + \frac{1}{2}\sigma^2(2tT - t^2) + \sigma\widetilde{W}_t, \\
P(t,T) &= \exp\left(-\int_t^T f(t,u)du\right) \\
&= \exp\left(-\int_t^T f(0,u)du - \frac{1}{2}\sigma^2\int_t^T (2tu - t^2)du - \sigma(T-t)\widetilde{W}_t\right) \\
&= \exp\left(-\int_t^T f(0,u)du - \frac{1}{2}\sigma^2 T(T-t)t - \sigma(T-t)\widetilde{W}_t\right).
\end{aligned}$$

By (10.6.5), we get

$$P(t,T) = \frac{P(0,T)}{P(0,t)}\exp\Big(-\frac{1}{2}\sigma^2 T(T-t)t - \sigma(T-t)\widetilde{W}_t\Big)$$

In particular,

$$\begin{aligned}
r_t &= f(t,t) \\
&= f(0,t) + \frac{1}{2}\sigma^2 t^2 + \sigma\widetilde{W}_t. \qquad (11.1.13)
\end{aligned}$$

From this, $\sigma\widetilde{W}_t = r_t - f(0,t) - \frac{1}{2}\sigma^2 t^2$. Therefore

$$\begin{aligned}
P(t,T) &= \frac{P(0,T)}{P(0,t)}\exp\left(-\frac{1}{2}\sigma^2 T(T-t)t - (T-t)\Big(r_t - f(0,t) - \frac{1}{2}\sigma^2 t^2\Big)\right) \\
&= \frac{P(0,T)}{P(0,t)}\exp\left((T-t)f(0,t) - \frac{1}{2}\sigma^2 t(T-t)^2 - (T-t)r_t\right).
\end{aligned}$$

■

Note that we may obtain (11.1.13) from (11.1.12) [1].

[1]

$$\begin{aligned}
r_t &= r_0 + \int_0^t \theta(s)ds + \sigma\widetilde{W}_t \\
&= r_0\int_0^t \left[\frac{\partial f}{\partial T}(0,s) + \sigma^2 s\right]ds + \sigma\widetilde{W}_t \\
&= r_0 + f(0,t) - f(0,0) + \frac{1}{2}\sigma^2 t^2 + \sigma\widetilde{W}_t \\
&= f(0,t) + \frac{1}{2}\sigma^2 t^2 + \sigma\widetilde{W}_t.
\end{aligned}$$

Example 11.1.7. If the volatility process $\sigma(t,T)$ is constant, the drift term in Eq (11.1.5) is given by

$$\alpha(t,T) = \sigma \int_t^T \sigma ds = \sigma^2(T-t).$$

From Theorem (11.1.6),

$$\begin{aligned} r_t &= f(0,t) + \frac{1}{2}\sigma^2 t^2 + \sigma W_t, \\ dr_t &= \Big(f_T(0,t) + \sigma^2 t\Big)dt + \sigma dW_t. \end{aligned}$$

We thus find that an HJM model with constant σ coincides with a Ho & Lee model with calibrated drift [2].

The interpretation of such a model is that each increment dW_t moves all points on the forward curve $\{f(t,T); t \leq T\}$ by an equal amount σdW_t. Does the diffusion term introduce only parallel shifts in the forward curve. Since $f(t,T)$ is given by

$$f(t,T) = f(0,T) + \frac{1}{2}\sigma^2(2tT - t^2) + \sigma\widetilde{W}_t,$$

the drift will vary (slightly, because σ^2 is small) across maturities, T, keeping the forward curve form making exactly parallel shifts. □

[2] By Theorem (10.7.1), $dr_t = \Big(f_T(t,t) + \sigma^2(t-t)\Big)dt + \sigma dW_t = f_T(t,t)dt + \sigma dW_t$. Thus $f_T(t,t) = f_T(0,t) + \sigma^2 t$.

11.1.4 Hull-White Model in HJM(Extended Vasicek)

The short rate dynamics are given by

$$dr_t = (\Phi(t) - ar_t)dt + \sigma d\widetilde{W}_t.$$

From (10.11.10) and (10.11.11), we have

$$\begin{aligned} g(x,t,T) &= -\log P(t,T) \\ &= -\int_t^T \Big(\frac{1}{2}\sigma^2 B^2(s,T) - \Phi(s)B(s,T)\Big)ds + B(t,T)x. \end{aligned}$$

By Theorem (11.1.5),

$$\begin{aligned} df(t,T) &= \sigma^2 e^{-a(T-t)}B(t,T)dt + \sigma e^{-a(T-t)}d\widetilde{W}_t, \\ f(t,T) &= f(0,T) + \int_0^t \sigma^2 e^{-a(T-s)}B(s,T)ds + \int_0^t \sigma e^{-a(T-s)}d\widetilde{W}_s, \\ r_t &= f(t,t) \\ &= f(0,t) + \int_0^t \sigma^2 e^{-a(t-s)}B(s,t)ds + \int_0^t \sigma e^{-a(t-s)}d\widetilde{W}_s \\ &= f(0,t) - \frac{\sigma^2}{2}B(s,t)^2\Big|_0^t + \int_0^t \sigma e^{-a(t-s)}d\widetilde{W}_s \\ &= f(0,t) + \frac{\sigma^2}{2}B(0,t)^2 + \int_0^t \sigma e^{-a(t-s)}d\widetilde{W}_s, \end{aligned}$$

where $B(t,T) = \frac{1-e^{-a(T-t)}}{a}$. Now, we can calculate the bond price $P(t,T)$.

$$\begin{aligned} P(t,T) &= \exp\Big(-\int_t^T f(t,u)du\Big), \\ \log P(t,T) &= -\int_t^T f(t,u)du \\ &= -\int_t^T f(0,u)du - \int_t^T\int_0^t \sigma^2 e^{-a(u-s)}B(s,u)dsdu - \int_t^T\int_0^t \sigma e^{-a(u-s)}d\widetilde{W}_s du \\ &= \log\Big(\frac{P(0,T)}{P(0,t)}\Big) - \frac{\sigma^2}{2}\int_t^T \big(B(0,u)^2 - B(t,u)^2\big)du - \int_0^t\int_t^T \sigma e^{-a(u-s)}du\ d\widetilde{W}_s \\ &= \log\Big(\frac{P(0,T)}{P(0,t)}\Big) - \frac{\sigma^2}{2}\int_t^T \big(B(0,u)^2 - B(t,u)^2\big)du - \sigma\int_0^t \frac{e^{-a(t-s)}(1-e^{-a(T-t)})}{a}d\widetilde{W}_s \\ &= \log\Big(\frac{P(0,T)}{P(0,t)}\Big) - \frac{\sigma^2}{2}\int_t^T \big(B(0,u)^2 - B(t,u)^2\big)du - B(t,T)\sigma\int_0^t e^{-a(t-s)}d\widetilde{W}_s \\ &= \log\Big(\frac{P(0,T)}{P(0,t)}\Big) - \frac{\sigma^2}{2}\int_t^T \big(B(0,u)^2 - B(t,u)^2\big)du - B(t,T)\Big(r_t - f(0,t) - \frac{\sigma^2}{2}B(0,t)^2\Big) \\ &= \log\Big(\frac{P(0,T)}{P(0,t)}\Big) + B(t,T)f(0,t) - \frac{\sigma^2}{2}\Big(\int_t^T \big(B(0,u)^2 - B(t,u)^2\big)du - B(t,T)B(0,t)^2\Big) - B(t,T)r_t. \end{aligned}$$

$$-\frac{\sigma^2}{2}\Big\{\int_t^T \big(B(0,u)^2 - B(t,u)^2\big)du - B(t,T)B(0,t)^2\Big\}$$

$$
\begin{aligned}
&= -\frac{\sigma^2}{2}\Big\{\int_t^T \big(B(0,u)^2 - B(t,u)^2\big)du - B(t,T)B(0,t)^2\Big\} \\
&= -\frac{\sigma^2}{2}\Big\{\big(\frac{1}{a^2}+\frac{2}{a^3}\,e^{-au}-\frac{1}{2a^3}\,e^{-2au}\big)\Big|_t^T - \int_0^{B(t,T)}\frac{y^2}{1-ay}dy - B(t,T)B(0,t)^2\Big\} \\
&= -\frac{\sigma^2}{2}\Big\{\frac{1}{a^2}(T-t)+\frac{2e^{-at}}{a^2}B(t,T)-\frac{1}{2a^3}(e^{-2aT}-e^{-2at}) \\
&\qquad +\frac{1}{a^2}B(t,T)+\frac{1}{2a}B(t,T)^2-\frac{1}{a^2}(T-t)-B(t,T)\frac{1-2e^{-at}+e^{-2at}}{a^2}\Big\} \\
&= -\frac{\sigma^2}{2}\Big(\frac{1}{2a}B(t,T)^2-\frac{1}{2a^3}(e^{-2aT}-e^{-2at})-\frac{1}{a^2}B(t,T)e^{-2at}\Big) \\
&= -\frac{\sigma^2}{2}\Big(\frac{1}{2a}B(t,T)^2-\frac{1}{2a^3}(e^{-2aT}-e^{-2at})-\frac{1-e^{-a(T-t)}}{a^3}e^{-2at}\Big) \\
&= -\frac{\sigma^2}{2}\Big(\frac{1}{2a}B(t,T)^2-\frac{e^{-2aT}-e^{-2at}+2e^{-2at}-2e^{-a(T+t)}}{2a^3}\Big) \\
&= -\frac{\sigma^2}{2}\Big(\frac{1}{2a}B(t,T)^2-\frac{e^{-2aT}+e^{-2at}-2e^{-a(T+t)}}{2a^3}\Big) \\
&= -\frac{\sigma^2}{2}\Big(\frac{1}{2a}B(t,T)^2-\frac{e^{-2at}}{2a}B(t,T)^2\Big) \\
&= -\frac{\sigma^2}{4a}B(t,T)^2(1-e^{-2at}).
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\log P(t,T) &= \log\Big(\frac{P(0,T)}{P(0,t)}\Big)+B(t,T)f(0,t)-\frac{\sigma^2}{4a}B(t,T)^2(1-e^{-2at})-B(t,T)r_t, \\
P(t,T) &= \frac{P(0,T)}{P(0,t)}\exp\Big(B(t,T)f(0,t)-\frac{\sigma^2}{4a}B(t,T)^2(1-e^{-2at})-B(t,T)r_t\Big).
\end{aligned}
$$

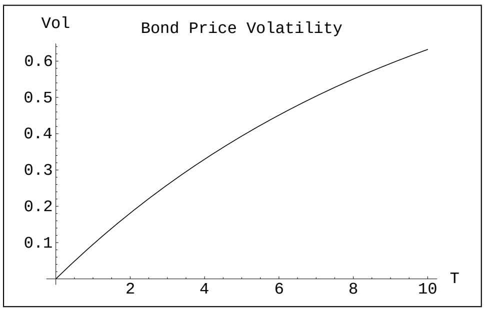
✎ **Volatility Structure in the Hull-White Model**

We can summarize the dynamics of bond price, forward rate and spot rate $R(t,T)$ as follows.

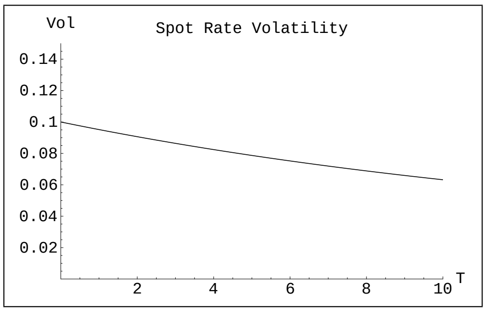
$$
\begin{aligned}
dP(t,T) &= P(t,T)\Big(r_t dt - \sigma B(t,T)d\widetilde{W}_t\Big), \\
df(t,T) &= \sigma^2 e^{-a(T-t)}B(t,T)dt + \sigma e^{-a(T-t)}d\widetilde{W}_t, \\
dR(t,T) &= (\cdots)dt + \frac{\sigma B(t,T)}{T-t}dW_t. \quad \Big(R(t,T) = -\frac{A(t,T)}{T-t}-\frac{B(t,T)}{T-t}r_t\Big)
\end{aligned}
$$

✎ **Summing Up(Vasicek & Hull-White(Extended Vasicek))**

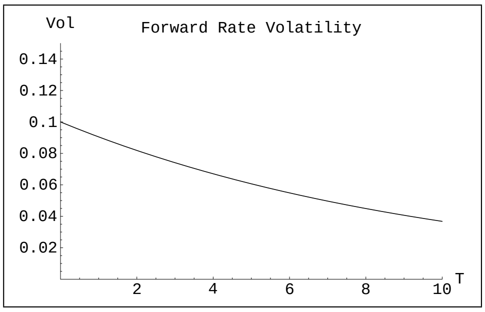
The summarized results are shown in Table (11.1).



(a) Bond Price Volatility



(b) Spot Rate Volatility



(c) Forward Rate Volatility

Figure 11.1: Volatility Structure in the Hull-White Model

Vasicek: $dr = (b - ar)dt + \sigma d\widetilde{W}$	Hull-White: $dr = (\Phi(t) - ar)dt + \sigma d\widetilde{W}$
1. $df = \sigma^2 e^{-a(T-t)} B(t,T)dt + \sigma e^{-a(T-t)} d\widetilde{W}_t$. 2. $dP = P(t,T)\Big(r_t dt - \sigma B(t,T) d\widetilde{W}_t\Big)$. 3. $f(0,T) = \frac{b}{a} + e^{-aT}(r_0 - \frac{b}{a}) - \frac{\sigma^2}{2a^2}(1-e^{-aT})^2$.	1. $df = \sigma^2 e^{-a(T-t)} B(t,T)dt + \sigma e^{-a(T-t)} d\widetilde{W}_t$. 2. $dP = P(t,T)\Big(r_t dt - \sigma B(t,T) d\widetilde{W}_t\Big)$. 3. $f(0,T) = e^{-aT} r_0 + \int_0^T \Phi(s) e^{-a(T-s)} ds - \frac{\sigma^2}{2a^2}(1 - e^{-aT})^2$.

Table 11.1: Vasicek vs Hull-White Model

11.2 Forward Measure

We specialize the theory to the case when the new numeraire chosen is a bond maturing at T_F. As can be expected, this choice of numeraire is particularly useful when dealing with interest rate derivatives.

For the given standard risk neutral measure, Q(with B as numeraire), The T_F-forward measure Q^{T_F} is defined by

$$dQ^{T_F} \quad = \quad L^{T_F}(t) dQ$$

on $\mathscr{F}_t$ for $0 \le t \le T_F$ where

$$L^{T_F}(t) \quad := \quad \frac{P(t,T_F)}{B_t P(0,T_F)}.$$

For a T_F-claim X,

$$X_t \quad = \quad P(t,T_F)E^T\big[X/P(T_F,T_F)|\mathscr{F}_t\big] \quad = \quad P(t,T_F)E^T\big[X|\mathscr{F}_t\big]. \tag{11.2.1}$$

Remark 11.2.1. Q and Q^{T_F} are coincide if and only if $L^{T_F}(T_F) = 1$, i.e.

$$1 \quad = \quad \frac{P(T_F,T_F)}{B_{T_F}P(0,T_F)} = \frac{e^{-\int_0^{T_F} r_s ds}}{E^Q\big[e^{-\int_0^{T_F} r_s ds}\big]}$$

i.e. if and only if r is deterministic.

Example 11.2.2. To get $L^{T_F}(t)$ explicitly, consider $d\log L^T(t)$.

$$\begin{aligned}
\log L^{T_F}(t) &= \log P(t,T_F) - \log B_t - \log P(0,T_F) \\
d\log L^{T_F}(t) &= \frac{1}{P(t,T_F)}dP(t,T_F) - \frac{1}{2}\frac{1}{P^2(t,T_F)}dP(t,T_F)^2 - \frac{1}{B_t}dB_t \\
&= r_t dt + \Sigma(t,T_F)d\widetilde{W}_t - \frac{1}{2}\Sigma^2(t,T_F)dt - r_t dt, \qquad \text{by (11.1.3)} \\
&= \Sigma(t,T_F)d\widetilde{W}_t - \frac{1}{2}\Sigma^2(t,T_F)dt.
\end{aligned}$$

Thus we have

$$L^{T_F}(t) \quad = \quad \exp\left(-\frac{1}{2}\int_0^t \Sigma^2(u,T_F)du + \int_0^t \Sigma(u,T_F)d\widetilde{W}_u\right).$$

By the Girsanov theorem, the process $W_t^{T_F}$ defined by

$$dW_t^{T_F} \quad = \quad -\Sigma(t,T_F)dt + d\widetilde{W}_t$$

is therefore a standard Brownian motion under Q^{T_F}. □

Example 11.2.3. We can find that under forward measure Q^{T_F} the forward dynamics (11.1.5) becomes

$$\begin{aligned}
df(t,T) &= -\sigma(t,T)\Sigma(t,T)dt + \sigma(t,T)d\widetilde{W}_t \\
&= -\sigma(t,T)\Sigma(t,T)dt + \sigma(t,T)\Big[\Sigma(t,T_F)dt + dW_t^{T_F}\Big] \\
&= -\sigma(t,T)\Big[\Sigma(t,T) - \Sigma(t,T_F)\Big]dt + \sigma(t,T)dW_t^{T_F} \\
&= -\sigma(t,T)\left(\int_T^{T_F} \sigma(t,u)du\right)dt + \sigma(t,T)dW_t^{T_F}
\end{aligned}$$

for $t \le T \le T_F$. Notice that $f(t,T_F)$ is a martingale under Q^{T_F}, though none of the forward rates is a martingale under risk neutral measure.

Theorem 11.2.4. Assume that, for all $T > 0$ we have $r_T/B_T \in L^1(Q)$. Then, for every fixed T, the forward rate process $f(t,T)$ is a Q^T-martingale for $0 \le t \le T$, and in particular we have

$$f(t,T) \;=\; E^T\big[r_T \big| \mathscr{F}_t\big].$$

Proof. Let $X := r$.

$$X_t \;=\; E^Q\Big[r_T e^{-\int_t^T r_s ds}\Big|\mathscr{F}_t\Big] = P(t,T)E^T\Big[r_T\Big|\mathscr{F}_t\Big].$$

This give us

$$\begin{aligned}
E^T\Big[r_T\Big|\mathscr{F}_t\Big] &= \frac{1}{P(t,T)}E^Q\Big[r_T e^{-\int_t^T r_s ds}\Big|\mathscr{F}_t\Big] \\
&= -\frac{1}{P(t,T)}E^Q\Big[\frac{\partial}{\partial T}e^{-\int_t^T r_s ds}\Big|\mathscr{F}_t\Big] \\
&= -\frac{1}{P(t,T)}\frac{\partial}{\partial T}E^Q\Big[e^{-\int_t^T r_s ds}\Big|\mathscr{F}_t\Big] \\
&= -\frac{P_T(t,T)}{P(t,T)} \\
&= -\frac{\partial}{\partial T}\log P(t,T) \\
&= f(t,T).
\end{aligned}$$

■

Remark 11.2.5. It is sometimes claimed that "the forward rate is an unbiased estimate of the future spot rate". We can see that in general this conjecture is false, not only under the physical measure, but also under the risk neutral measure Q. This conjecture is true under the forward measure.

Summing Up

1. $X_t = B_t E^Q\Big[\frac{X_T}{B_T}\Big|\mathscr{F}_t\Big] = P(t,T)E^T\Big[\frac{X_T}{P(T,T)}\Big|\mathscr{F}_t\Big] = P(t,T)E^T\Big[X_T\Big|\mathscr{F}_t\Big].$
2. $X_0 = E^Q\big[\frac{X}{B_T}\big] = P(0,T)E^T\big[X\big].$

11.2.1 Extended Black-Scholes Formula

Consider a fixed time T, and a European call on S with date of maturity T and strike K. We are thus considering the T-claim

$$
\begin{aligned}
X_T &= \max[S(T) - K, 0], \\
X_0 &= E^Q\Big[\frac{S(T) - K}{B_T}\ \mathbf{1}_{S(T)\geq K}\Big] \\
&= E^Q\Big[\frac{S(T)}{B_T}\ \mathbf{1}_{S(T)\geq K}\Big] \ - \ E^Q\Big[\frac{K}{B_T}\ \mathbf{1}_{S(T)\geq K}\Big] \\
&= S(0)E^S\Big[\frac{S(T)}{S(T)}\ \mathbf{1}_{S(T)\geq K}\Big] \ - \ P(0,T)E^T\Big[\frac{K}{P(T,T)}\ \mathbf{1}_{S(T)\geq K}\Big] \\
&= S(0)\mathrm{Prob}^S\big[S(T) \geq K\big] \ - \ KP(0,T)\mathrm{Prob}^T\big[S(T) \geq K\big].
\end{aligned}
$$

Assume that the process $Z_{S,T}$ defined by

$$Z_{S,T}(t) = \frac{S(t)}{P(t,T)}$$

has a stochastic differential of the form

$$dZ_{S,T}(t) = Z_{S,T}\big(m_{S,T}(t)dt + \sigma_{S,T}(t)dW\big). \tag{11.2.2}$$

where the volatility process $\sigma_{S,T}$ is <u>deterministic</u>. Note that the volatility process as always is unaffected by a change of measure, so it is not necessary to specify under which measure we check the condition.

In order to analyze the option formula we start with the second term which we write as

$$
\begin{aligned}
\mathrm{Prob}^T\big[S(T) \geq K\big] &= \mathrm{Prob}^T\big[S(T)/P(T,T) \geq K\big] \\
&= \mathrm{Prob}^T\big[Z_{S,T}(T) \geq K\big].
\end{aligned}
$$

By construction we know that $Z_{S,T}$ is a martingale under Q^T, so its Q^T-dynamics are given by

$$dZ_{S,T}(t) = Z_{S,T}(t)\sigma_{S,T}(t)dW^T,$$

with the solution

$$Z_{S,T}(T) = \frac{S(0)}{P(0,T)}\ \exp\left(-\frac{1}{2}\int_0^T \|\sigma_{S,T}\|^2(t)dt + \int_0^T \sigma_{S,T}(t)dW^T\right).$$

Since $\sigma_{S,T}(t)$ is deterministic, $\log\big\{Z_{S,T}(T)\frac{P(0,T)}{S(0)}\big\}$ is Gaussian with mean $-\frac{1}{2}\Sigma^2_{S,T}(T)$ and variance $\Sigma^2_{S,T}(T)$, where

$$\Sigma^2_{S,T}(T) = \int_0^T \|\sigma_{S,T}\|^2(t)dt.$$

Now we see that

$$\mathrm{Prob}^T\big[Z_{S,T}(T) \geq K\big] = N\left[-\frac{\log\big(\frac{P(0,T)K}{S(0)}\big) + \frac{1}{2}\Sigma^2_{S,T}(T)}{\sqrt{\Sigma^2_{S,T}(T)}}\right]$$

$$= \quad N\left[\frac{\log\left(\frac{S(0)}{P(0,T)K}\right) - \frac{1}{2}\Sigma^2_{S,T}(T)}{\sqrt{\Sigma^2_{S,T}(T)}}\right].$$

For the first term we write

$$\begin{aligned}
\mathrm{Prob}^S\big[S(T) \geq K\big] &= \mathrm{Prob}^S\big[Z_{S,T}(T) \geq K\big] \\
&= \mathrm{Prob}^S\big[\frac{1}{Z_{S,T}(T)} \leq \frac{1}{K}\big] \\
&= \mathrm{Prob}^S\big[Y_{S,T}(T) \leq \frac{1}{K}\big].
\end{aligned}$$

Where $Y_{S,T}$ is defined by

$$Y_{S,T}(t) = \frac{P(t,T)}{S(t)} = \frac{1}{Z_{S,T}(t)}.$$

Under the measure Q^S the process $Y_{S,T}$ is a martingale.

$$\begin{aligned}
dY_{S,T}(t) &= d\Big(\frac{1}{Z_{S,T}(t)}\Big) \\
&= -\frac{1}{Z^2_{S,T}(t)}\, Z_{S,T}\big(m_{S,T}(t)dt + \sigma_{S,T}(t)dW\big) + \frac{1}{2}\frac{2}{Z^3_{S,T}(t)} Z^2_{S,T}\sigma^2_{S,T}(t)dt \\
&= dY_{S,T}(t)\Big\{\big(-m_{S,T}(t) + \sigma^2_{S,T}(t)\big)dt - \sigma_{S,T}(t)dW\Big\}.
\end{aligned}$$

So Q^S-dynamics of $Y_{S,T}$ is of the form

$$\begin{aligned}
dY_{S,T}(t) &= -Y_{S,T}(t)\sigma_{S,T}(t)dW^S, \\
Y_{S,T}(t) &= \frac{P(0,T)}{S(0)}\ \exp\left(-\frac{1}{2}\int_0^T \|\sigma_{S,T}\|^2(t)dt - \int_0^T \sigma_{S,T}(t)dW^T\right).
\end{aligned}$$

From this we see that

$$\mathrm{Prob}^S\left[Y_{S,T}(T) \leq \frac{1}{K}\right] = N\left[\frac{\log\left(\frac{S(0)}{P(0,T)K}\right) + \frac{1}{2}\Sigma^2_{S,T}(T)}{\sqrt{\Sigma^2_{S,T}(T)}}\right].$$

We have thus proved the following result:

$$X_0 = S(0)N\left[\frac{\log\left(\frac{S(0)}{P(0,T)K}\right) + \frac{1}{2}\Sigma^2_{S,T}(T)}{\sqrt{\Sigma^2_{S,T}(T)}}\right] - KP(0,T)N\left[\frac{\log\left(\frac{S(0)}{P(0,T)K}\right) - \frac{1}{2}\Sigma^2_{S,T}(T)}{\sqrt{\Sigma^2_{S,T}(T)}}\right].$$

11.2.2 Valuing European Options on Zero-Coupon Bonds

We will now discuss pricing of interest rate derivatives in the Hull-White model. We start by pricing a European call option with date of maturity T_1 and strike price K, on an underlying bond with date of maturity T_2, where $T_1 < T_2$. In terms of preceding subsection this means that $T = T_1$ and that $S(t) = P(t, T_2)$, and first we have to check (11.2.2), i.e. if the volatility, σ_z, of the process

$$Z(t) = \frac{P(t,T_2)}{P(t,T_1)}$$

is deterministic. Note that $dP(t,T) = P(t,T)\big(r_t dt - \sigma B(t,T)d\widetilde{W}_t\big)$ and $B(t,T) = \frac{1-e^{-a(T-t)}}{a}$.

$$\begin{aligned}
dZ(t) &= \frac{dP(t,T_2)}{P(t,T_1)} - \frac{P(t,T_2)}{P(t,T_1)^2}dP(t,T_1) + \frac{1}{2}\frac{2P(t,T_2)}{P(t,T_1)^3}dP(t,T_1)^2 - \frac{1}{P(t,T_1)^2}dP(t,T_1)\cdot dP(t,T_2)\\
&= \frac{P(t,T_2)}{P(t,T_1)}\Big\{r_t dt - \sigma B(t,T_2)dW_t - r_t dt + \sigma B(t,T_1)dW_t + \sigma^2 B^2(t,T_1)dt - \sigma^2 B(t,T_1)B(t,T_2)dt\Big\}\\
&= \frac{P(t,T_2)}{P(t,T_1)}\Big\{\sigma\big(B(t,T_1) - B(t,T_2)\big)dW_t + \sigma^2 B(t,T_1)\big(B(t,T_1) - B(t,T_2)\big)dt\Big\}.
\end{aligned}$$

From this, we see that

$$\sigma_z = \frac{\sigma}{a}e^{at}\big(e^{-aT_2} - e^{-aT_1}\big).$$

Thus σ_z is in fact deterministic, so we may apply the preceding subsection technic.

Theorem 11.2.6 (Jamshidian, Hull-White Bond Option). In the Hull-White model, the price, at $t = 0$, of a European call with strike K, and time of maturity T_1, on a bond maturing at T_2 is given by the formula

$$P(0,T_2)N(d_1) - KP(0,T_1)N(d_2), \tag{11.2.3}$$

where

$$\begin{aligned}
d_1 &= \frac{\log\left(\frac{P(0,T_2)}{KP(0,T_1)}\right) + \frac{1}{2}\Sigma^2}{\sqrt{\Sigma^2}},\\
d_2 &= d_1 - \sqrt{\Sigma^2},\\
\Sigma^2 &= \frac{\sigma^2}{2a^3}\big(1 - e^{-2aT_1}\big)\left(1 - e^{-a(T_2-T_1)}\right)^2.
\end{aligned}$$

Proof.

$$\begin{aligned}
\Sigma^2 &= \int_0^{T_1}\sigma_z^2 dt\\
&= \frac{\sigma^2}{a^2}\big(e^{-aT_2} - e^{-aT_1}\big)^2\int_0^{T_1}e^{2at}dt\\
&= \frac{\sigma^2}{2a^3}\big(1 - e^{-2aT_1}\big)\left(1 - e^{-a(T_2-T_1)}\right)^2.
\end{aligned}$$

■

Jamshidian has shown that options on zero-coupon bonds can be valued by this formula using Vasicek's model.

Remark 11.2.7. In general, consider the forward model,

$$df(t,T) \quad = \quad \alpha(t,T)dt + \sigma(t,T)d\widetilde{W}_t.$$

From this, the dynamics of $P(t,T)$ is determined. The volatility of $P(t,T)$ is

$$\Sigma(t,T) \quad = \quad -\int_t^T \sigma(t,s)ds. \tag{11.2.4}$$

In order to computer the price of option, we first have to find the volatility σ_{T_1,T_2} of the process

$$Z(t) \quad = \quad \frac{P(t,T_2)}{P(t,T_1)} \quad (T_1 < T_2).$$

An easy calculation shows that in fact

$$\sigma_{T_1,T_2}(t) \quad = \quad \Sigma(t,T_2) - \Sigma(t,T_1) = -\int_{T_1}^{T_2} \sigma(t,s)ds \tag{11.2.5}$$

From this, we can generalize Theorem (11.2.6).

Theorem 11.2.8 (General Bond Option Pricing). Assume that the forward dynamics is given by

$$df(t,T) \quad = \quad \alpha(t,T)dt + \sigma(t,T)d\widetilde{W}_t.$$

If the volatility $\sigma(t,T)$ is deterministic, the bond option's value is given by

$$P(0,T_2)N(d_1) - KP(0,T_1)N(d_2), \tag{11.2.6}$$

where

$$\begin{aligned} d_1 &= \frac{\log\left(\frac{P(0,T_2)}{KP(0,T_1)}\right) + \frac{1}{2}\Sigma^2}{\sqrt{\Sigma^2}}, \\ d_2 &= d_1 - \sqrt{\Sigma^2}, \\ \Sigma^2 &= \int_0^{T_1} \|\sigma_{T_1,T_2}(t)\|^2 dt. \end{aligned}$$

Remark 11.2.9. In CIR model, the volatility of the forward dynamics is not deterministic. Hence Theorem (11.2.8) could not apply at getting the option price.

Corollary 11.2.10 (Bond Option With Vasicek Model). The zero coupon bond option price under Vasicek model is given by (11.2.3). i.e the zero coupon bond prices are same under Hull-White and Vasicek model.

Proof. The volatilities of forward dynamics in Vasicek and Hull-White model are same, so the option price is same by Theorem (11.2.8). ■

Corollary 11.2.11 (Bond Option With Ho-Lee Model). The zero coupon bond option price under Ho & Lee model is given by (11.2.6) with

$$\Sigma^2 = \sigma^2(T_2 - T_1)^2 T_1.$$

Proof. By Theorem 11.1.6, the forward dynamics of Ho-Lee model is

$$df_t = \sigma^2(T-t)dt + \sigma d\widetilde{W}_t.$$

By (11.2.5) and Theorem (11.2.8),

$$\begin{aligned}
\sigma_{T_1,T_2}(t) &= -\int_{T_1}^{T_2} \sigma ds = -\sigma(T_2 - T_1), \\
\Sigma^2 &= \int_0^{T_1} \sigma^2(T_2-T_1)^2 ds = \sigma^2(T_2-T_1)^2 T_1.
\end{aligned}$$

■

Note that θ_t in Eq (11.1.11) is not used in the price of interest derivatives.

Theorem 11.2.12 (Bond Option with CIR Model). The price at time zero, $X(0)$, of a call option that matures at time T_1 on a zero coupon bond maturing at time T_2, with strike price K is

$$\begin{aligned}
X(0) &= P(0,T_2)\ \chi^2\Big(2r^*\left(\phi(T_1) + B(T_1,T_2) + \psi\right); \frac{4b}{\sigma^2}, \frac{2\phi^2(T_1)\ e^{rT_1}\ r(0)}{\phi(T_1) + B(T_1,T_2) + \psi}\Big) \\
&\quad - K\ P(0,T_1)\ \chi^2\Big(2r^*\left(\phi(T_1) + \psi\right); \frac{4b}{\sigma^2}, \frac{2\phi^2(T_1)\ e^{rT_1}\ r(0)}{\phi(T_1) + \psi}\Big)
\end{aligned}$$

where

$$\begin{aligned}
r^* &= \frac{A(T_1,T_2) - \log K}{B(T_1,T_2)}, \\
\gamma &= \sqrt{a^2 + 2\sigma^2}, \\
\phi(T) &= \frac{2\gamma}{\sigma^2(e^{\gamma T} - 1)}, \\
\psi &= \frac{a+\gamma}{\sigma^2},
\end{aligned}$$

and $\chi^2(\cdot)$ is the non-central chi-square distribution. $A(t,T)$ and $B(t,T)$ are defined in Section (10.11.2).

Proof.

Step 1. Distribution of $r(t)$ under Q-measure:

$$\begin{aligned}
dr &= (b - ar)dt + \sigma\sqrt{r}dW, \\
P(t,T) &= \exp\left\{A(t,T) - B(t,T)r_t\right\}.
\end{aligned}$$

Let us calculate the volatility of $P(t,T)$.

$$dP = P\Big(\big(A_t(t,T) - B_t(t,T)r_t\big)dt - B(t,T)dr_t\Big) + P(\cdots).$$

From this we get the volatility of $P(t,T)$:

$$-\sigma\sqrt{r_t}\ B(t,T).$$

Then, by Theorem (1.15.8),

$$d\widetilde{W} = dW - \sigma\sqrt{r_t}\ B(t,T)dt$$

is a Brownian motion under T-measure. We have

$$dr_t = \Big(b - \big\{a + \sigma^2 B(t,T)\big\}r_t\Big)dt + \sigma\sqrt{r_t}d\widetilde{W}, \tag{11.2.7}$$

and note that Equation (10.11.9). Now consider the Ornstein Uhlenbeck process

$$dX_i(t) = -\frac{1}{2}\big\{a + \sigma^2 B(t,T)\big\}X_i(t)dt + \frac{1}{2}\sigma dW_i(t).$$

To get the solution to this SDE, consider the followings:

$$\begin{aligned} d\Big(e^{\frac{1}{2}\left(at+\frac{\sigma^2}{b}A(t,T)\right)}X_i(t)\Big) &= \frac{\sigma}{2}e^{\frac{1}{2}\left(at+\frac{\sigma^2}{b}A(t,T)\right)}dW_i(t), \\ e^{\frac{1}{2}\left(at+\frac{\sigma^2}{b}A(t,T)\right)}X_i(t) &= e^{\frac{\sigma^2}{2b}A(0,T)}X_i(0) + \frac{\sigma}{2}\int_0^t e^{\frac{1}{2}\left(as+\frac{\sigma^2}{b}A(s,T)\right)}dW_i(t), \end{aligned}$$

$$X_i(t) = e^{-\frac{1}{2}\left(at+\frac{\sigma^2}{b}A(t,T)-\frac{\sigma^2}{b}A(0,T)\right)}X_i(0) + \frac{\sigma}{2}e^{-\frac{1}{2}\left(at+\frac{\sigma^2}{b}A(t,T)\right)}\int_0^t e^{\frac{1}{2}\left(as+\frac{\sigma^2}{b}A(s,T)\right)}dW_i(s).$$

This solution is a Gaussian process with mean function

$$m(t,T) = e^{-\frac{1}{2}\left(at+\frac{\sigma^2}{b}A(t,T)-\frac{\sigma^2}{b}A(0,T)\right)}X_i(0)$$

and variance function

$$\begin{aligned} \rho(t,T) &= \frac{\sigma^2}{4}e^{-\left(at+\frac{\sigma^2}{b}A(t,T)\right)}\int_0^t e^{as+\frac{\sigma^2}{b}A(s,T)}ds \\ &= \frac{\sigma^2}{4}e^{-at}K^{-2}(t,T)\int_0^t e^{as}K^2(s,T)ds. \end{aligned}$$

where

$$K(t,T) = \frac{2\gamma e^{\frac{(a+\gamma)(T-t)}{2}}}{(a+\gamma)\big(e^{\gamma(T-t)}-1\big)+2\gamma}.$$

After a tedious calculation, we have that

$$\begin{aligned} m(t,T) &= e^{\frac{\gamma}{2}t}\ \frac{(a+\gamma)(e^{\gamma(T-t)}-1)+2\gamma}{(a+\gamma)(e^{\gamma T}-1)+2\gamma}X_i(0), \\ \rho(t,T) &= \frac{\sigma^2}{2}\ \frac{e^{\gamma t}-1}{\gamma(e^{\gamma t}+1)+(\sigma^2 B(t,T)+a)(e^{\gamma t}-1)} \\ &= \frac{1}{2}\ \frac{1}{\frac{\gamma(e^{\gamma t}+1)}{\sigma^2(e^{\gamma t}-1)}+B(t,T)+\frac{a}{\sigma^2}} \\ &= \frac{1}{2}\ \frac{1}{\frac{2\gamma}{\sigma^2(e^{\gamma t}-1)}+B(t,T)+\frac{a+\gamma}{\sigma^2}} \end{aligned}$$

$$= \frac{1}{2\big(\phi(t) + B(t,T) + \psi\big)}.$$

Define

$$r(t) := X_1^2(t) + X_2^2(t) + \cdots + X_n^2(t).$$

Ito's formula implies

$$dr(t) = \Big(\frac{n\sigma^2}{4} - \big\{a + \sigma^2 B(t,T)\big\} r_t\Big) dt + \sigma\sqrt{r_t}\, d\widetilde{W}.$$

If we define $b := \frac{n\sigma^2}{4}$, we obtain (11.2.7) and we can see that

$$\frac{r(t)}{\rho(t,T)}$$

is a non-central chi-square distribution, under T(-bond) measure, with $\frac{4b}{\sigma^2}$ degree of freedom and parameter

$$\begin{aligned}
\frac{\frac{4b}{\sigma^2}\ m^2(t,T)}{\rho(t,T)} &= \frac{8r(0)\gamma^2 e^{\gamma t}}{\sigma^2(e^{\gamma t}-1)\big[\gamma(e^{\gamma t}+1) + (\sigma^2 B(t,T) + a)(e^{\gamma t}-1)\big]}, \quad X_i^2(0)\cdot\frac{4b}{\sigma^2} = r(0) \\
&= \frac{1}{\sigma^4}\ \frac{8\ \frac{e^{\gamma t}\gamma^2}{(e^{\gamma t}-1)^2}\ r(0)}{\frac{\gamma(e^{\gamma t}+1)}{\sigma^2(e^{\gamma t}-1)} + B(t,T) + \frac{a}{\sigma^2}} \\
&= \frac{2\ \frac{4\gamma^2}{\sigma^4(e^{\gamma t}-1)^2}\ e^{\gamma t} r(0)}{\frac{2\gamma}{\sigma^2(e^{\gamma t}-1)} + B(t,T) + \frac{a+\gamma}{\sigma^2}} \\
&= \frac{2\phi^2(t)\ e^{rt}\ r(0)}{\phi(t) + B(t,T) + \psi}.
\end{aligned}$$

Note that n, in general, need not be integer.

Step 2 Option price:

$$X(0) = P(0,T_2)\mathrm{Prob}^{T_2}\big[Y(T_1) \le \frac{1}{K}\big] - KP(0,T_1)\mathrm{Prob}^{T_1}\big[Z(T_1) \ge K\big]$$

where $Z(t) = \frac{P(t,T_2)}{P(t,T_1)}, Y(t) = \frac{P(t,T_1)}{P(t,T_2)}$ are martingales under T_1-measure and T_2-measure, respectively If we define r^* such that r^* satisfies $P(r^*, T_1, T_2) = K$, then

$$Y(T_1) \le \frac{1}{K} \iff Z(T_1) \ge K \iff P\big(r(T_1), T_1, T_2\big) \ge K \iff r(T_1) \le r^*.$$

From this we have

$$\begin{aligned}
X(0) &= P(0,T_2)\mathrm{Prob}^{T_2}\left[r(T_1) \le r^*\right] - KP(0,T_1)\mathrm{Prob}^{T_1}\left[r(T_1) \le r^*\right] \\
&= P(0,T_2)\mathrm{Prob}^{T_2}\left[\frac{r(T_1)}{\rho(T_1,T_2)} \le \frac{r^*}{\rho(T_1,T_2)}\right] - KP(0,T_1)\mathrm{Prob}^{T_1}\left[\frac{r(T_1)}{\rho(T_1,T_1)} \le \frac{r^*}{\rho(T_1,T_1)}\right] \\
&= P(0,T_2)\ \chi^2\left(2r^*\Big(\phi(T_1) + B(T_1,T_2) + \psi\Big); \frac{4b}{\sigma^2}, \frac{2\phi^2(T_1)\ e^{rT_1}\ r(0)}{\phi(T_1) + B(T_1,T_2) + \psi}\right) \\
&\quad -\ K\ P(0,T_1)\ \chi^2\left(2r^*\big(\phi(T_1) + \psi\big); \frac{4b}{\sigma^2}, \frac{2\phi^2(T_1)\ e^{rT_1}\ r(0)}{\phi(T_1) + \psi}\right).
\end{aligned}$$

■

✎ Put-Call Parity of Bond Option

If we are given a option on T_2-bond with maturity T_1 and exercise price X, the put-call parity at time t is given by

$$p_t + P(t,T_2) \quad = \quad c_t + XP(t,T_1).$$

Option Pricing Formulae

1. $Z(t) = \frac{S(t)}{P(t,T_1)}, Y(t) = \frac{P(t,T_1)}{S(t)}$ are martingales under T_1-measure and S-measure, respectively.

$$X(0) \quad = \quad S(0)\mathrm{Prob}^S\left[Y(T_1) \le \frac{1}{K}\right] - KP(0,T_1)\mathrm{Prob}^{T_1}\left[Z(T_1) \ge K\right]$$

2. $Z(t) = \frac{P(t,T_2)}{P(t,T_1)}, Y(t) = \frac{P(t,T_1)}{P(t,T_2)}$ are martingales under T_1-measure and T_2-measure, respectively.

$$X(0) \quad = \quad P(0,T_2)\mathrm{Prob}^{T_2}\left[Y(T_1) \le \frac{1}{K}\right] - KP(0,T_1)\mathrm{Prob}^{T_1}\left[Z(T_1) \ge K\right]$$

11.2.3 Valuing European Options on Coupon Bearing Bonds

Consider a European call option with exercise price X and maturity T on a coupon-bearing bond. Suppose that the bond provides a total n cash flows after the option matures. Let the ith cash flow be c_i and occurs at time $s_i (1 \le i \le n; s_i > T)$.

	call option
underlying	coupon bearing bond
option maturity	T
exercise price	X
remainder cash flows	(c_i, s_i)

Table 11.2: Option on Coupon-Bearing Bond

Now we could show that option on coupon-bearing bond can be split into options on zero-coupon bonds.

Step 1. Find the short rate r_X such that

$$\sum_{i=1}^{n} c_i P(T, s_i; r_X) \quad = \quad X.$$

Step 2. Let X_i be $P(T, s_i; r_X)$ for each i. i.e $X = \sum_{i=1}^{n} c_i X_i$. We split the bond option into n bond options with exercise price $c_i X_i$ on zero-coupon bond which matures at s_i respectively.

Step 3. Let C_t be option value at time t. Since all bond prices, $P(T, s_i; r)$, are decreasing functions of r,

$$\begin{aligned} C_t &= E\left[\left(\sum_{i=1}^{n} c_i P(T, s_i) - X\right)^+ \Big| r_t\right] \\ &= E\left[\left(\sum_{i=1}^{n} c_i P(T, s_i) - \sum_{i=1}^{n} c_i X_i\right)^+ \Big| r_t\right] \\ &= E\left[\sum_{i=1}^{n} c_i \Big(P(T, s_i) - X_i\Big)^+ \Big| r_t\right] \\ &= \sum_{i=1}^{n} c_i E\left[\Big(P(T, s_i) - X_i\Big)^+ \Big| r_t\right]. \end{aligned}$$

This shows that the option on the coupon-bearing bond is the sum of n options on the underlying zero-coupon bonds.

11.2.4 Contingent Claim on $r(T)$ under Hull-White Model

We will analyze a T-claim of the form $X = H(r_T)$ where H is some given function. Using T-bond as numeraire, we have

$$X_t = P(t,T)E^T[H(r_T)|\mathscr{F}_t].$$

In Hull-White model,

$$\begin{aligned} dP(t,T) &= P(t,T)\big(r_t dt - \sigma B(t,T)d\widetilde{W}(t)\big) \\ dW^T(t) &= d\widetilde{W}(t) + \sigma B(t,T)dt, \quad \text{(by 1.15.8)}, \\ dr_t &= \Big(\Phi(t) - ar_t\Big)dt + \sigma d\widetilde{W}(t), \\ &= \Big(\Phi(t) - ar_t - \sigma^2 B(t,T)\Big)dt + \sigma dW^T(t). \end{aligned}$$

Now let us computer r_T:

$$\begin{aligned} d(e^{at}r_t) &= ae^{at}r_t dt + e^{at}dr_t \\ &= ae^{at}r_t dt + e^{at}\Big(\Phi(t) - ar_t - \sigma^2 B(t,T)\Big)dt + \sigma e^{at}dW^T(t) \\ &= e^{at}\Big(\Phi(t) - \sigma^2 B(t,T)\Big)dt + \sigma e^{at}dW^T(t), \\ e^{aT}r_T &= e^{at}r_t + \int_t^T e^{as}\Big(\Phi(s) - \sigma^2 B(s,T)\Big)ds + \sigma\int_t^T e^{as}dW^T(s), \\ r_T &= e^{-a(T-t)}r_t + \int_t^T e^{-a(T-s)}\Big(\Phi(s) - \sigma^2 B(s,T)\Big)ds + \sigma\int_t^T e^{-a(T-s)}dW^T(s). \end{aligned}$$

From this we can see that r_T is the normal distribution. Thus let us computer the conditional mean, $m_r(t,T)$, and variance, $\sigma_r^2(t,T)$, of r_T.

$$\begin{aligned} \sigma_r^2(t,T) &= \sigma^2\int_t^T e^{-2a(T-s)}ds = \frac{\sigma^2}{2a}\Big(1 - e^{-2a(T-t)}\Big), \\ m_r(t,T) &= E^T[r_T|\mathscr{F}_t] \;=\; f(t,T), \quad \text{(by Theorem 11.2.4)} . \end{aligned}$$

We see that under Q^T the conditional distribution of r_T has the normal distribution $\mathbb{N}\big(f(t,T), \sigma_r^2(t,T)\big)$.

Theorem 11.2.13. Given the assumption above, the price of the claim $X = H(r_T)$ is given by

$$\begin{aligned} X_t &= P(t,T)E^T[H(r_T)|\mathscr{F}_t] \\ &= P(t,T)\frac{1}{\sqrt{2\pi\sigma_r^2(t,T)}}\int_{-\infty}^{\infty} h(z)\exp\Big(-\frac{\big(z - f(t,T)\big)^2}{2\sigma_r^2(t,T)}\Big)dz. \end{aligned}$$

11.2.5 Caps and Floors

An interest rate cap is a contract that guarantees to its holder that floating rates will not exceed a specified rate(cap rate). The Caplet is one component of an interest rate cap, i.e caplet is defined as the contingent claim with the following payoff, paid at T_i,

$$X(T_i) := \delta\Big(L(T_{i-1}, T_i) - R\Big)^+,$$

where R is cap rate which is pre-determined. Note that $X(T_i)$ is payed at time T_i but it is determined at time T_{i-1}, i.e. $X(T_i)$ is $\mathscr{F}_{T_{i-1}}$-measurable.

Recall that

$$\begin{aligned} 1 &= \Big(1 + \delta L(T_{i-1}, T_i)\Big) P(T_{i-1}, T_i), \\ \delta L(T_{i-1}, T_i) &= \frac{1}{P(T_{i-1}, T_i)} - 1. \end{aligned}$$

Under the forward measure with respect to T_i-bond, we have

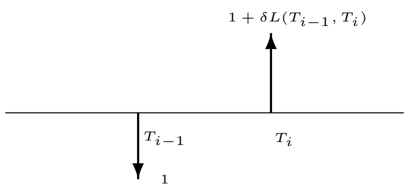


Figure 11.2: Floating Rate

1. $0 \le t \le T_{i-1}$

$$\begin{aligned} X(t) &= P(t, T_i) E_{T_i}\Big[X(T_i)\big|\mathscr{F}_t\Big] \\ &= P(t, T_i) E_{T_i}\Big[\delta\big(L(T_{i-1}, T_i) - R)^+\big|\mathscr{F}_t\Big]. \end{aligned} \tag{11.2.8}$$

2. $T_{i-1} \le t \le T_i$: Since $X(T_i)$ is $\mathscr{F}_t$-measurable, we have

$$X(t) = \delta P(t, T_i)\big(L(T_{i-1}, T_i) - R)^+.$$

Denoting $L = L(T_{i-1}, T_i)$ and $P = P(T_{i-1}, T_i)$, we have

$$\begin{aligned} X(T_i) &= \delta(L - R)^+ \\ &= \left(\frac{1}{P} - 1 - \delta R\right)^+ \\ &= \frac{1 + \delta R}{P}\left(\frac{1}{1 + \delta R} - P\right)^+, \\ X(T_{i-1}) &= (1 + \delta R)\left(\frac{1}{1 + \delta R} - P\right)^+. \end{aligned}$$

Consequently we see that a caplet is equivalent to $(1 + \delta R)$ put options on an underlying T_i-bond where the exercise date of the option is at T_{i-1} and the exercise price is $\dfrac{1}{1 + \delta R}$.

Example 11.2.14. When $T_{i-1} = 0, T_i = T$, the cap's value is given by

$$
\begin{aligned}
& P(0,T)\delta\Big(L(0,T) - R\Big)^+ \\
= \quad & P(0,T)\Big(1 + \delta L(0,T) - 1 - \delta R\Big)^+ \\
= \quad & \Big(1 + P(0,T)(1 + \delta R)\Big)^+ \\
= \quad & (1 + \delta R)\left(\frac{1}{1+\delta R} - P(0,T)\right)^+ .
\end{aligned}
$$

□

A floor is similar to a cap except that the floor ensures that the interest rate is bounded below, by R_f. A floor is made up of a sum of floorlets, each of which has a cash flow of

$$\delta\Big(R_f - L(T_{i-1}, T_i)\Big)^+ .$$

As the similar way of caplet, we can see that a floorlet is equivalent to $(1 + \delta R)$ call options on an underlying T_i-bond where the exercise date of the option is at T_{i-1} and the exercise price is $\dfrac{1}{1+\delta R}$.

A collar is an instrument designed to guarantee that the interest rate on the underlying FRN always lies between two levels. A collar is a combination of a long position in a cap and a short position in a floor.

✎ **Cap-Floor Parity**

A portfolio of a long caplet and a short floorlet (with $R_c = R_f$) has the cash flow

$$\Big(L(T_{i-1}, T_i) - R\Big)^+ - \Big(R_f - L(T_{i-1}, T_i)\Big)^+ \quad = \quad L(T_{i-1}, T_i) - R_c.$$

This is the same cashflow as one payment of a swap. There is a cap-floor parity relationship between the prices of caps and floors. This is

Summing Up :

1. The caplet can be characterized as a call option on a LIBOR, or a put option on a zero coupon bond.
2. The floorlet is a put option on a LIBOR, or a call option on a zero coupon bond.
3. cap price = floor price + value of swap.

11.2.6 Swaption

A swaption is an option on enter into a swap on a future date at a given rate. Suppose we have an option to receive fixed on a swap starting at date T_0. The swap payment dates are $T_i = T_0 + i\delta(i = 1, \cdots, n)$, and the fixed swap rate is k. Then the worth of the option at time T_0 is

$$\left(P(T_0, T_n) + k\delta \sum_{i=1}^{n} P(T_0, T_i) - 1\right)^+ .$$

This is exactly the same as a call option, struck 1, on a T_n-bond which pays a coupon at rate k at each time T_i. That is not entirely a conincidence as a swap is just a coupon bond less a floating bond (which always has par value). If you receive fixed on a swap, you have a long position in the bond market; a swap looks like a bond option.

Chapter 12

Brace-Gatarek-Musiela Model

1. **Gaussian Models**: In many popular interest models, either the short rate or the forward rates are modeled as Gaussian processes. This is the case with such models as Ho-Lee, Vasicek, and the Hull-White extension of the Visicek model. On the forward rate side, any model within the HJM framework possessing a deterministic volatility, will give rise to a Gaussian forward rate curve.

   i. Analytic tractability

   ii. Positive probability for negative interest rates. This would lead to theoretical arbitrage possibilities.

2. **Lognormal Models**: We can model the interest rates lognormally, thus avoiding negative rates, and this is the way taken Dothan and Black-Derman-Toy. Log-normal models are nice in the sense that interest rates stay positive, but the main drawback with a lognormal short rate is that the money account will have infinite expected value, i.e.

$$E[B_T] = \infty, \quad \text{for all } T.$$

3. **Lognormal for forward**: One may be tempted to model the forward rates in a lognormal fashion, by specifying the volatility structure as

$$df(t,T) = \alpha(t,T)dt + \sigma f(t,T)dW_t.$$

   It can be shown that in this case the forward rates will explode to plus infinity. This will force bond prices to zero. Under risk neutral measure, it follows that

$$df(t,T) = \sigma^2 f(t,T)\left(\int_t^T f(t,u)du\right)dt + \sigma f(t,T)d\widetilde{W}_t,$$

   and Heath, Jarrow and Morton show that solutions to this equation explode before T. To see what problem this cause, consider the similar deterministic ordinary differential equation

$$y' = y^2, \quad y(0) = c > 0$$

   The solution to this ODE is given by

$$y(t) = \frac{c}{1-ct}. \tag{12.0.1}$$

The difficulty can be avoided by working with simple rather than continuous compounded forward rates.

The models considered in this chapter are closely related to the HJM framework of the previous chapter in that they describe the arbitrage-free dynamics of the term structure of interest rates through the evolution of forward rates. But the models we turn to now are based on simple rather than continuous compounded forward rates. This seemingly minor shift in focus has surprisingly far-reaching practical and theoretical implications. This modeling approach has developed primarily through the work of Brace, Gatarek, and Musiela(1997); it is called LIBOR Market Model(LMM) or BGM model.

12.1 Musiela Parametrization

Let t be current time, T be bond maturity, and τ be time to maturity, i.e. $\tau = T - t$. Define $r(t,\tau)$ and $D(t,\tau)$ as follows:

$$\begin{aligned} r(t,\tau) &= f(t,t+\tau), && (12.1.1)\\ D(t,\tau) &= P(t,t+\tau). && (12.1.2)\end{aligned}$$

Note that

$$\begin{aligned} D(t,\tau) &= P(t,t+\tau)\\ &= \exp\left(-\int_t^{t+\tau} f(t,u)du\right)\\ &= \exp\left(-\int_t^{\tau} f(t,t+v)dv\right)\\ &= \exp\left(-\int_t^{\tau} r(t,v)dv\right),\\ \frac{\partial D}{\partial \tau} &= P_T(t,t+\tau) \;=\; -r(t,\tau)D(t,\tau).\end{aligned}$$

We now derive the differentials of $r(t,\tau)$ and $D(t,\tau)$. We will now write $\bar{\sigma}(t,\tau) := \sigma(t,t+\tau)$, i.e. we will change the meaning of the second parameter from the 'maturity' to the 'time to maturity'. In this notation, the risk-neutral HJM model is

$$\begin{aligned} df(t,T)|_{T=t+\tau} &= -\sigma(t,t+\tau)\Sigma(t,t+\tau)dt + \sigma(t,t+\tau)d\widetilde{W}_t\\ &= -\bar{\sigma}(t,\tau)\bar{\Sigma}(t,\tau)dt + \bar{\sigma}(t,\tau)d\widetilde{W}_t,\\ dP(t,T)|_{T=t+\tau} &= P(t,T)\left(r_t dt + \Sigma(t,t+\tau)d\widetilde{W}_t\right)\\ &= P(t,T)\left(r_t dt + \bar{\Sigma}(t,\tau)d\widetilde{W}_t\right), && (12.1.3)\end{aligned}$$

where

$$\begin{aligned} \Sigma(t,T) &:= -\int_t^T \sigma(t,s)ds,\\ \bar{\Sigma}(t,\tau) &:= \Sigma(t,t+\tau)\end{aligned}$$

$$
\begin{aligned}
&= -\int_t^{t+\tau} \sigma(t,s)ds \\
&= -\int_0^{\tau} \sigma(t,t+u)du \\
&= -\int_0^{\tau} \bar{\sigma}(t,u)du
\end{aligned}
$$

We now derive the differential of $r(t,\tau)$ and $D(t,\tau)$. We have

$$
\begin{aligned}
dr(t,\tau) &= df(t,t+\tau) + f_T(t,t+\tau)dt \\
&= -\bar{\sigma}(t,\tau)\bar{\Sigma}(t,\tau)dt + \bar{\sigma}(t,\tau)d\widetilde{W}_t + \frac{\partial}{\partial\tau}r(t,\tau)dt \\
&= \frac{\partial}{\partial\tau}\Big[r(t,\tau) + \frac{1}{2}\bar{\Sigma}^2(t,\tau)\Big]dt + \bar{\sigma}(t,\tau)d\widetilde{W}_t. \qquad (12.1.4)
\end{aligned}
$$

Also,

$$
\begin{aligned}
dD(t,\tau) &= dP(t,t+\tau) + P_T(t,t+\tau)dt \\
&= P(t,t+\tau)\Big(r(t)dt + \bar{\Sigma}(t,\tau)d\widetilde{W}_t\Big) - r(t,\tau)D(t,\tau)dt \\
&= D(t,\tau)\Big[r(t,0) - r(t,\tau)\Big]dt + D(t,\tau)\bar{\Sigma}(t,\tau)d\widetilde{W}_t.
\end{aligned}
$$

12.2 Single-Factor Brace-Gatarek-Musiela Model

✎ **LIBOR**

Fix δ(say,$\delta = \frac{1}{2}$ or $\frac{1}{4}$). $D(t,\delta)$ invested at time t in $(t+\delta)$-maturity bond grow to 1 at time $t+\delta$. $L(t,0)$ is defined to be the corresponding rate of simple interest:

$$\begin{aligned}
1 &= D(t,\delta)(1+\delta L(t,0)), \\
L(t,0) &:= \frac{\frac{1}{D(t,\delta)} - 1}{\delta} \\
&= \frac{\exp\left(\int_t^{t+\delta} f(t,v)dv\right) - 1}{\delta} \\
&= \frac{\exp\left(\int_0^{\delta} r(t,u)du\right) - 1}{\delta},
\end{aligned}$$

where $r(t,u)$ is defined in (12.1.1).

✎ **Forward LIBOR Rate**

The forward LIBOR $L(t,\tau)$ is defined to be simple forward interest rate for period from $t+\tau$ to $t+\tau+\delta$.

Figure 12.1: Forward LIBOR

$$\begin{aligned}
1 &= \frac{D(t,\tau+\delta)}{D(t,\tau)}(1+\delta L(t,\tau)), \\
L(t,\tau) &:= \frac{\frac{D(t,\tau)}{D(t,\tau+\delta)} - 1}{\delta} \\
&= \frac{\exp\left(\int_{t+\tau}^{t+\tau+\delta} f(t,v)dv\right) - 1}{\delta} \\
&= \frac{\exp\left(\int_{\tau}^{\tau+\delta} f(t,t+u)du\right) - 1}{\delta} \\
&= \frac{\exp\left(\int_{\tau}^{\tau+\delta} r(t,u)du\right) - 1}{\delta}, \quad \delta > 0 \text{ fixed.}
\end{aligned}$$

$r(t,\tau)$ is the continuously compounded rate. $L(t,\tau)$ is the simple rate over a period of duration δ. We cannot have a log-normal model for $r(t,\tau)$ because the solutions explode as we saw in Eq. (12.0.1). For a fixed positive δ, we can have a lognormal model for $L(t,\tau)$.

✎ **The Dynamics of** $L(t,\tau)$

To get $dL(t,\tau)$, first computer the following:

$$
\begin{aligned}
& d\left(\int_\tau^{\tau+\delta} r(t,u)du\right) \\
= \quad & \int_\tau^{\tau+\delta} dr(t,u)du \\
\overset{(12.1.4)}{=} \quad & \int_\tau^{\tau+\delta} \frac{\partial}{\partial u}\Big[r(t,u)+\frac{1}{2}\bar{\Sigma}^2(t,u)\Big]dudt + \int_\tau^{\tau+\delta}\bar{\sigma}(t,u)dud\widetilde{W}_t \\
= \quad & \left(r(t,\tau+\delta)-r(t,\tau)+\frac{1}{2}\bar{\Sigma}^2(t,\tau+\delta)-\frac{1}{2}\bar{\Sigma}^2(t,\tau)\right)dt - \big(\bar{\Sigma}(t,\tau+\delta)-\bar{\Sigma}(t,\tau)\big)\,d\widetilde{W}.
\end{aligned}
$$

From this we obtain

$$
\begin{aligned}
& dL(t,\tau) \\
= \quad & d\left(\frac{\exp\left(\int_\tau^{\tau+\delta} r(t,u)du\right)-1}{\delta}\right) \\
= \quad & \frac{1}{\delta}\exp\left(\int_\tau^{\tau+\delta} r(t,u)du\right)\; d\left(\int_\tau^{\tau+\delta} r(t,u)du\right) + \frac{1}{2\delta}\exp\left(\int_\tau^{\tau+\delta} r(t,u)du\right)\; d\left(\int_\tau^{\tau+\delta} r(t,u)du\right)^2 \\
= \quad & \frac{1}{\delta}\left(1+\delta L(t,\tau)\right)\Bigg\{\left(r(t,\tau+\delta)-r(t,\tau)+\frac{1}{2}\bar{\Sigma}^2(t,\tau+\delta)-\frac{1}{2}\bar{\Sigma}^2(t,\tau)\right)dt \\
& \quad - \big(\bar{\Sigma}(t,\tau+\delta)-\bar{\Sigma}(t,\tau)\big)\,d\widetilde{W} + \frac{1}{2}\left(\bar{\Sigma}(t,\tau+\delta)-\bar{\Sigma}(t,\tau)\right)^2 dt\Bigg\} \\
= \quad & \frac{1}{\delta}\left(1+\delta L(t,\tau)\right)\Bigg\{\big(r(t,\tau+\delta)-r(t,\tau)+\bar{\Sigma}^2(t,\tau+\delta)-\bar{\Sigma}(t,\tau+\delta)\bar{\Sigma}(t,\tau)\big)\,dt \\
& \quad - \big(\bar{\Sigma}(t,\tau+\delta)-\bar{\Sigma}(t,\tau)\big)\,d\widetilde{W}\Bigg\}.
\end{aligned}
$$

But

$$
\begin{aligned}
\frac{\partial}{\partial\tau}L(t,\tau) \quad &= \quad \frac{\partial}{\partial\tau}\left(\frac{\exp\left(\int_\tau^{\tau+\delta} r(t,u)du\right)-1}{\delta}\right) \\
&= \quad \frac{1}{\delta}\exp\left(\int_\tau^{\tau+\delta} r(t,u)du\right)\left(r(t,\tau+\delta)-r(t,\tau)\right).
\end{aligned}
$$

Therefore we have

$$
dL(t,\tau) \quad = \quad \frac{\partial}{\partial\tau}L(t,\tau)dt + \frac{1}{\delta}\Big(1+\delta L(t,\tau)\Big)\Big(\bar{\Sigma}(t,\tau+\delta)-\bar{\Sigma}(t,\tau)\Big)\,\Big(\bar{\Sigma}(t,\tau+\delta)dt - d\widetilde{W}_t\Big)
$$

Define $\gamma(t,\tau)$ as it satisfies

$$
\gamma(t,\tau)L(t,\tau) \quad = \quad -\frac{1}{\delta}\Big(1+\delta L(t,\tau)\Big)\Big(\bar{\Sigma}(t,\tau+\delta)-\bar{\Sigma}(t,\tau)\Big). \tag{12.2.1}
$$

Then

$$
dL(t,\tau) \quad = \quad \left\{\frac{\partial}{\partial\tau}L(t,\tau)-\gamma(t,\tau)L(t,\tau)\bar{\Sigma}(t,\tau+\delta)\right\}dt + \gamma(t,\tau)L(t,\tau)d\widetilde{W}_t. \tag{12.2.2}
$$

Since (12.2.1) is equivalent to

$$\bar{\Sigma}(t,\tau+\delta) = \bar{\Sigma}(t,\tau) - \frac{\delta L(t,\tau)\gamma(t,\tau)}{1+\delta L(t,\tau)},$$

(12.2.2) yields

$$dL(t,\tau) = \left\{\frac{\partial}{\partial\tau}L(t,\tau) - \gamma(t,\tau)L(t,\tau)\bar{\Sigma}(t,\tau) + \frac{\delta L^2(t,\tau)\gamma^2(t,\tau)}{1+\delta L(t,\tau)}\right\}dt + \gamma(t,\tau)L(t,\tau)d\widetilde{W}_t. \quad (12.2.3)$$

✎ **Forward Measures**

Let $W^T(t)$ be a Brownian motion under forward measure with respect to T-maturity zero coupon bond. Then from (12.1.3) $W^T(t)$ is given by

$$\begin{aligned} W^T(t) &= \widetilde{W}(t) - \int_0^t \Sigma(u,T)du \\ &= \widetilde{W}(t) - \int_0^t \bar{\Sigma}(u,T-u)du, \quad 0 \le t \le T. \end{aligned} \quad (12.2.4)$$

✎ **Returning to HJM**

Set

$$K(t,T) = L(t,T-t).$$

Then

$$\begin{aligned} dK(t,T) &= dL(t,T-t) - \frac{\partial}{\partial\tau}L(t,T-t)dt \\ &\overset{(12.2.2)}{=} \gamma(t,T-t)K(t,T)\Big(-\bar{\Sigma}(t,T-t+\delta)dt + d\widetilde{W}_t\Big) \quad (12.2.5) \\ &\overset{(12.2.3)}{=} \gamma(t,T-t)K(t,T)\left(-\bar{\Sigma}(t,T-t+\delta)dt + \frac{\delta K(t,T)\gamma(t,T-t)}{1+\delta K(t,T)}dt + d\widetilde{W}_t\right). \quad (12.2.6) \end{aligned}$$

Remark 12.2.1. Although the dt term in (12.2.6) has the term

$$\frac{\delta K(t,T)\gamma(t,T-t)}{1+\delta K(t,T)}$$

involving K^2, solutions to this equation do not explode because

$$\frac{\delta K(t,T)\gamma(t,T-t)}{1+\delta K(t,T)} \le \frac{\delta K(t,T)\gamma(t,T-t)}{\delta K(t,T)} \le \gamma^2(t,T-t)K(t,T).$$

□

From (12.2.5), (12.2.4) and (C.0.1) we have

$$\begin{aligned} dK(t,T) &= \gamma(t,T-t)K(t,T)\Big(-\bar{\Sigma}(t,T-t+\delta)dt + d\widetilde{W}_t\Big) \\ &= \gamma(t,T-t)K(t,T)dW^{T+\delta}(t), \end{aligned}$$

so

$$K(t,T) = K(0,T)\exp\left(\int_0^t \gamma(u,T-u)dW^{T+\delta}(u) - \frac{1}{2}\int_0^t \gamma^2(u,T-u)du\right)$$

and

$$\begin{aligned} K(T,T) &= K(0,T)\exp\left(\int_0^T \gamma(u,T-u)dW^{T+\delta}(u) - \frac{1}{2}\int_0^T \gamma^2(u,T-u)du\right) \\ &= K(t,T)\exp\left(\int_t^T \gamma(u,T-u)dW^{T+\delta}(u) - \frac{1}{2}\int_t^T \gamma^2(u,T-u)du\right). \end{aligned}$$

We assume that γ is non-random. Then

$$X(t) := \int_t^T \gamma(u,T-u)dW^{T+\delta}(u) - \frac{1}{2}\int_t^T \gamma^2(u,T-u)du$$

is normal with variance

$$\rho^2(t) := \int_t^T \gamma^2(u,T-u)du,$$

and mean $-\frac{1}{2}\rho^2(t)$. Thus we have

$$K(T,T) = K(t,T)\exp\Big(X(t)\Big). \tag{12.2.7}$$

✎ **Pricing Interest Rate Caplet**

Refer to Subsection (11.2.5). Consider a floating rate interest payment settled in arrears. At time $T+\delta$, the floating rate interest payment due is $\delta L(T,0) = \delta K(T,T)$, the LIBOR at time T. The value of caplet at time $T+\delta$ is

$$\delta\Big(K(T,T) - R\Big)^+.$$

Let $C_{T+\delta}(t)$ be the value of caplet at time t. From (11.2.8) we have

$$\begin{aligned} C_{T+\delta}(t) &= P(t,T+\delta)E^{T+\delta}\left[\delta\Big(K(T,T)-R\Big)^+ |\mathscr{F}_t\right] \\ &\overset{(12.2.7)}{=} \delta P(t,T+\delta)E^{T+\delta}\left[\Big(K(t,T)\exp\big(X(t)\big) - R\Big)^+ |\mathscr{F}_t\right] \\ &= \delta P(t,T+\delta)g\big(K(t,T)\big), \end{aligned}$$

where

$$\begin{aligned} g(y) &:= E^{T+\delta}\left[\Big(y\exp\big(X(t)\big) - R\Big)^+ |\mathscr{F}_t\right] \\ &= yN\left(\frac{\log\frac{y}{R} + \frac{1}{2}\rho^2(t)}{\rho(t)}\right) - RN\left(\frac{\log\frac{y}{R} - \frac{1}{2}\rho^2(t)}{\rho(t)}\right), \quad \text{by Theorem (1.10.4).} \end{aligned}$$

We used the fact

$$E\left[\exp\big(X(t)\big)\right] = 1.$$

This is called Black caplet formula.

✎ **Calibration for γ**

Let us suppose γ is a deterministic function of its second argument, i.e. time-homogeneous:

$$\gamma(t,\tau) \quad = \quad \gamma(\tau).$$

Then g depends on

$$\int_0^T \gamma^2(T-u)du \quad = \quad \int_0^T \gamma^2(v)dv.$$

Since

$$C_{T+\delta}(0) \quad = \quad \delta P(0,T+\delta)\, g\big(K(0,T)\big),$$

if we know the caplet price $C_{T+\delta}(0)$ for all T, we can "back out" the squared volatility $\int_0^T \gamma^2(v)dv$.

For example, if we know caplet prices

$$C_{T_0+\delta}(0), \quad C_{T_1+\delta}(0), \quad \cdots, \quad C_{T_n+\delta}(0),$$

where $T_0 < T_1 < \cdots < T_n$, we can "back out"

$$\int_0^{T_0} \gamma^2(v)dv, \quad \int_{T_0}^{T_1} \gamma^2(v)dv = \int_0^{T_1} \gamma^2(v)dv - \int_0^{T_0} \gamma^2(v)dv, \quad \cdots, \quad \int_{T_{n-1}}^{T_n} \gamma^2(v)dv.$$

12.3 Multi-Factor Brace-Gatarek-Musiela Model

The main theory in this section is taken from Glasserman(2004). Suppose that a finite set of maturities or tenor dates

$$0 = T_0 < T_1 < \cdots < T_M < T_{M+1}$$

are fixed in advance. Let

$$\delta_i \quad = \quad T_{i+1} - T_i, \quad i = 0, \cdots, M,$$

denote the lengths of the intervals between tenor dates.

For each date T_n we let

$$P_n(t) \quad := \quad P(t, T_n), \quad 0 \le t \le T_n.$$

Define the forward LIBOR rate by

$$L_n(t) \quad := \quad L(t, T_n) \quad := \quad \frac{P(t,T_n) - P(t,T_{n+1})}{\delta_n P(t,T_{n+1})}, \quad 0 \le t \le T_n, \quad n = 0, \cdots, M. \quad (12.3.1)$$

After T_n, the forward rate L_n becomes meaningless; it sometimes simplifies notation to extend the definition of $L_n(t)$ beyond T_n by setting $L_n(t) = L_n(T_n)$ for all $t \ge T_n$.

1. From (12.3.1) we know that bond prices determine the forward rates. At a tenor date $t = T_i$, the relation can be inverted to produce

$$\begin{aligned} P_n(T_i) &= \frac{P_{i+1}(T_i)}{P_i(T_i)} \times \frac{P_{i+2}(T_i)}{P_{i+1}(T_i)} \times \cdots \times \frac{P_n(T_i)}{P_{n-1}(T_i)} \\ &= \frac{1}{1+\delta_i L_i(T_i)} \times \frac{1}{1+\delta_{i+1} L_{i+1}(T_i)} \times \cdots \times \frac{1}{1+\delta_{n-1} L_{n-1}(T_i)} \\ &= \prod_{j=i}^{n-1} \frac{1}{1+\delta_j L_j(T_i)}, \quad n = i+1, \cdots, M+1. \end{aligned}$$

2. For, at an arbitrary date t, the forward LIBOR rates do not determine the bond prices because they do not determine the discount factor for interval short than the accrual periods.

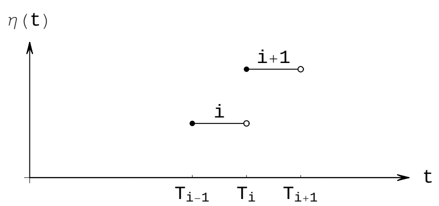


Figure 12.2: $\eta(t)$ is right continuous.

Suppose that $T_i \le t < T_{i+1}$. Define a function $\eta : [0, T_{M+1}) \to \{1, \cdots, M+1\}$ by taking $\eta(t)$ to be the unique integer satisfying

$$T_{\eta(t)-1} \le t < T_{\eta(t)}.$$

Thus $\eta(t)$ gives the index of the next tenor date at time t. With this notation, we have

$$P_n(t) \quad = \quad P_{\eta(t)}(t) \prod_{j=\eta(t)}^{n-1} \frac{1}{1+\delta_j L_j(t)}, \quad 0 \le t < T_n. \tag{12.3.2}$$

Example 12.3.1. As the compounding period δ_n decrease to zero, the forward LIBOR rate can be written as

$$\begin{aligned}
\lim_{\delta_n \to 0} L_n(t) &= \lim_{\delta_n \to 0} \frac{1}{\delta_n} \left\{ \exp\left(\int_{T_n}^{T_n+\delta_n} f(t,u)du \right) - 1 \right\} \\
&= \lim_{\delta_n \to 0} \exp\left(\int_{T_n}^{T_n+\delta_n} f(t,u)du \right) f(t, T_n+\delta_n) \\
&= f(t, T_n).
\end{aligned}$$

□

✎ Spot Measure

We seek a model in which the evolution of the forward LIBOR rates is described by a system of SDEs of the form

$$\frac{dL_n(t)}{L_n(t)} \quad = \quad \mu_n(t)dt + \sigma_n(t)^\top dW(t), \quad 0 \le t \le T_n, \quad n = 1, \cdots, M, \tag{12.3.3}$$

where $W(t)$ is d-dimensional standard Brownian motion. The coefficients μ_n and σ_n may depend on the current vector of rates $\big(L_1(t), \cdots, L_M(t)\big)$ as well as the current time t. Notice that in (12.3.3) σ_n is the proportional volatility because we have divided by L_n on the left, whereas in the HJM setting (11.1.7) we took $\sigma(t,T)$ to be the absolute level of volatility.

Recall that in the HJM setting we derived the form of the drift of the forward rate from the absence of arbitrage. More specifically, we derived the drift from the condition that bond prices be martingales when divided by the numeraire asset. The numeraire we used is the usual one associated with the risk-neutral measure $B(t) = \exp\left(\int_0^t r_u du\right)$. But introducing a short-rate process r_t would undermine our objective of developing a model based on the simple rates $L_n(t)$. We therefore avoid the usual risk-neutral measure and instead use a numeraire asset better suited to the tenor dates T_i.

A simple compounded counterpart of $B(t)$ works as follows. Star with 1 unit of account at time 0 and buy $\frac{1}{P_1(0)}$ bonds maturing at T_1. A time T_1, reinvest the funds in bonds maturing at time T_2 and proceed this way, at each T_i putting all funds in bonds maturing at time T_{i+1}. This strategy earns simple interest at rate $L_i(T_i)$ over each interval $[T_i, T_{i+1}]$ just as in the continuously compounded case a saving account earns interest at rate r_t at time t. The initial investment of 1 at time 0 grows to a value of

$$P^*(t) \quad := \quad P_{\eta(t)}(t) \prod_{j=0}^{\eta(t)-1} \Big(1 + \delta_j L_j(T_j)\Big) \tag{12.3.4}$$

at time t. We take this as numeraire asset and call the associated measure the spot measure. Suppose, then, that (12.3.3) holds under the spot measure, meaning that $W(t)$ is a d-dimensional standard Brownian motion under that measure. The absence of arbitrage restricts the dynamics of the forward

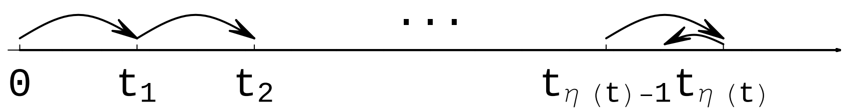


Figure 12.3: numeraire associated with spot measure

LIBOR rates through the condition that bond prices be martingales when *deflated*[1] by the numeraire asset. From (12.3.2) and (12.3.4), we find that the deflated bond price $D_n(t) := P_n(t)/P^*(t)$ is given by

$$D_n(t) = \left(\prod_{j=0}^{\eta(t)-1} \frac{1}{1+\delta_j L_j(T_j)}\right) \prod_{j=\eta(t)}^{n-1} \frac{1}{1+\delta_j L_j(t)}, \quad 0 \le t \le T_n. \tag{12.3.5}$$

Notice that the spot measure numeraire P^* cancels the factor $P_{\eta(t)}(t)$ used in (12.3.2) to discount between tenor dates. We are thus left in (12.3.5) with an expression defined purely in terms of the LIBOR rates. This would not have been the case had we divided by the risk-neutral numeraire $B(t)$.

We require that the deflated bond prices D_n be positive martingales and proceed to derive the restrictions this imposes on the LIBOR dynamics (12.3.3). If the deflated bonds are indeed positive martingales, we may write

$$\frac{dD_{n+1}(t)}{D_{n+1}(t)} = \nu_{n+1}^\top \, dW(t), n = 1, \cdots, M, \tag{12.3.6}$$

for some $\mathbb{R}^d$-valued processes ν_{n+1} which may depend on the current level of $(D_2, \cdots, D_{M+1})$ (equivalently, of $(L_1, \cdots, L_M)$). By Ito's formula and (12.3.6)

$$d\log D_{n+1}(t) = -\frac{1}{2}\|\nu_{n+1}(t)\|^2 dt + \nu_{n+1}^\top(t) dW(t). \tag{12.3.7}$$

Since the first factor in (12.3.5) is constant, we have

$$\begin{aligned}
d\log D_{n+1}(t) &= -\sum_{j=\eta(t)}^{n} d\log\Big(1+\delta_j L_j(t)\Big) \\
&= -\sum_{j=\eta(t)}^{n} \left\{\frac{\delta_j}{1+\delta_j L_j(t)} dL_j(t) + (\cdots)dt\right\} \\
&= -\sum_{j=\eta(t)}^{n} \left\{\frac{\delta_j L_j(t)}{1+\delta_j L_j(t)} \Big(\mu_j(t)dt + \sigma_j(t)^\top dW(t)\Big) + (\cdots)dt\right\} \\
&= -\sum_{j=\eta(t)}^{n} \left\{\frac{\delta_j L_j(t)}{1+\delta_j L_j(t)} \; \sigma_j(t)^\top dW(t) + (\cdots)dt\right\}. \qquad (12.3.8)
\end{aligned}$$

Comparing the coefficients of $dW(t)$ in (12.3.6) and (12.3.8), we find that

$$\nu_{n+1}(t) = -\sum_{j=\eta(t)}^{n} \frac{\delta_j L_j(t)}{1+\delta_j L_j(t)} \; \sigma_j(t). \tag{12.3.9}$$

[1] We use the term "deflated" rather than "discounted" to emphasize that we are dividing by the numeraire asset and not discounting at a continuously compounded rate.

We now proceed by induction to fine the μ_n in (12.3.3). Setting $D_1(t) \equiv P_1(0)$, we make D_1 constant and hence a martingale without restrictions on any of LIBOR rates. Suppose now that $\mu_1, \cdots, \mu_{n-1}$ have been chosen consistent with the martingale condition on D_n. Form the identity

$$D_n(t) = D_{n+1}(t)\Big(1 + \delta_n L_n(t)\Big),$$

we find that

$$\delta_n L_n(t) D_{n+1}(t) = D_n(t) - D_{n+1}(t),$$

so D_{n+1} is martingale if and only if $L_n D_{n+1}$ is a martingale. Applying Ito's formula, we get

$$\begin{aligned} d\big(L_n D_{n+1}\big) &= D_{n+1} dL_n + L_n dD_{n+1} + L_n D_{n+1} \nu_{n+1}^\top \sigma_n dt \\ &= \Big(D_{n+1}\mu_n L_n + L_n D_{n+1}\nu_{n+1}^\top \sigma_n\Big) dt + L_n D_{n+1}\sigma_n^\top dW + L_n dD_{n+1}. \end{aligned}$$

(We have suppressed the time argument to lighten the notation.) To be consistent with the martingale restriction on D_{n+1} and $L_n D_{n+1}$, the dt coefficient must be zero, and thus

$$\mu_n(t) = -\sigma_n^\top \nu_{n+1}.$$

Combining this with (12.3.9), we arrive at

$$\mu_n(t) = \sum_{j=\eta(t)}^{n} \frac{\delta_j L_j(t)}{1 + \delta_j L_j(t)} \, \sigma_n^\top \sigma_j(t)$$

as the required drift parameter in (12.3.3), so

$$\frac{dL_n(t)}{L_n(t)} = \sum_{j=\eta(t)}^{n} \frac{\delta_j L_j(t) \sigma_n^\top \sigma_j(t)}{1 + \delta_j L_j(t)} dt + \sigma_n(t)^\top dW(t), \quad 0 \le t \le T_n, \tag{12.3.10}$$

$n = 1, \cdots, M$, describes the arbitrage-free dynamics of forward LIBOR rates under the spot measure.

✎ **Forward Measure**

We formulate a LIBOR market model under forward measure P^{M+1} for maturity T_{M+1} and take the bond $P(\cdot, T_{M+1})$ as numeraire.

The dynamics of the forward LIBOR rates is described by a system of SDEs of the form

$$\frac{dL_n(t)}{L_n(t)} = \mu_n(t)dt + \sigma_n(t)^\top dW^{M+1}(t), \quad 0 \le t \le T_n, \quad n = 1, \cdots, M,$$

with W^{M+1} a standard d-dimensional Brownian motion under P^{M+1}. The coefficients $\mu_n \in \mathbb{R}$ and $\sigma_n \in \mathbb{R}^d$ my depend on the current vector of rates $(L_1(t), \cdots, L_M(t))$ as well as the current time t. The variables $\sigma_n(t)$ are the primary determinants of both the level of volatility in forward rate and the correlations between forward rate.

Theorem 12.3.2. The arbitrage-free dynamics of the $L_n, n = 1, \cdots, M$, under the forward measure P_{M+1} are given by

$$\frac{dL_n(t)}{L_n(t)} = -\sum_{j=n+1}^{M} \frac{\delta_j L_j(t) \sigma_n(t)^\top \sigma_j(t)}{1 + \delta_j L_j(t)} \, dt + \sigma_n(t)^\top dW^{M+1}(t), \quad 0 \le t \le T_n. \tag{12.3.11}$$

Proof. By the definition of the forward LIBOR rate, we have

$$L_n(t)P(t,T_{n+1}) = \frac{P(t,T_n)-P(t,T_{n+1})}{\delta_n}.$$

Notice that $L_n(t)P(t,T_{n+1})$ is a tradable asset and it can be divided with the numeraire $P(t,T_{n+1})$ which, by definition of a martingale measure, gives a martingale under P^{n+1}. In this case, we define the deflated bond prices to the ratio $D_n(t) = P(t,T_n)/P(t,T_{M+1})$, which simplify to

$$D_n(t) = \prod_{j=n}^{M}\big(1+\delta_j L_j(t)\big). \tag{12.3.12}$$

The diffusion process above can, in the martingale sense, be written as

$$\frac{dL_n(t)}{L_n(t)} = \sigma_n(t)^\top dW^{n+1}(t), \tag{12.3.13}$$

where $W^{n+1}(t)$ is a Brownian motion under P^{n+1}. However applying the reasoning above to the forward LIBOR rate $L_m(t), m \neq n$ one notices that $L_m(t)P(t,T_{n+1})$ is not a tradable asset and hence under the measure P^{n+1} only $L_n(t)$ is a martingale.

Note that a change of measure only changes the drift term whether the diffusion term remains unaffected. In order to determine the drift term under P^{n+2} for the P^{n+1}-martingale $L_n(t)$ consider the change of measure dP^{n+1}/dP^{n+2}. A Radon-Nikodym derivative $\rho(t)$ is given by

$$\rho(t) := \frac{dP^{n+1}}{dP^{n+2}} = \frac{P(t,T_{n+1})/P(0,T_{n+1})}{P(t,T_{n+2})/P(0,T_{n+2})} = \frac{P(0,T_{n+2})}{P(0,T_{n+1})}\Big(1+\delta_{n+1}L_{n+1}(t)\Big). \tag{12.3.14}$$

An application of Ito's lemma on (12.3.14) shows that

$$\begin{aligned} d\rho(t) &= \frac{P(0,T_{n+2})}{P(0,T_{n+1})}\,\delta_{n+1}\; dL_{n+1}(t) \\ &= \frac{P(0,T_{n+2})}{P(0,T_{n+1})}\,\delta_{n+1}\; L_{n+1}(t)\sigma_{n+1}(t)^\top dW^{n+1}(t), \quad \text{by (12.3.13)} \\ &= \rho(t)\,\frac{\delta_{n+1}L_{n+1}(t)}{1+\delta_{n+1}L_{n+1}(t)}\;\sigma_{n+1}(t)^\top dW^{n+1}(t). \end{aligned}$$

Through Girsanov theorem, we have that

$$dW^{n+1}(t) = dW^{n+2}(t) - \frac{\delta_{n+1}L_{n+1}(t)}{1+\delta_{n+1}L_{n+1}(t)}\,\sigma_{n+1}(t)\; dt.$$

Hence we have

$$\begin{aligned} \frac{dL_n(t)}{L_n(t)} &= \sigma_n(t)^\top dW^{n+1}(t) \\ &= -\frac{\delta_{n+1}L_{n+1}(t)}{1+\delta_{n+1}L_{n+1}(t)}\,\sigma_n(t)^\top\sigma_{n+1}(t)\; dt + \sigma_n(t)^\top dW^{n+2}(t) \\ &\;\;\vdots \end{aligned}$$

$$= \quad -\sum_{j=n+1}^{M} \frac{\delta_j L_j(t)}{1+\delta_j L_j(t)} \; \sigma_n(t)^\top \sigma_j(t)\; dt + \sigma_n(t)^\top dW^{M+1}(t).$$

■

Example 12.3.3. If we take $n = M$ in (12.3.11), we find that

$$\frac{dL_M(t)}{L_M(t)} \quad = \quad \sigma_M(t)^\top dW^{M+1}(t),$$

so that, subject to regularity conditions on its volatility, L_M is a martingale under the forward measure for maturity T_{M+1}. Moreover, if σ_M is deterministic then $L_M(t)$ has lognormal distribution, i.e.

$$\log \frac{L_M(t)}{L_M(0)} \quad \sim \quad N\left(-\frac{1}{2}\overline{\sigma}^2_M(t), \overline{\sigma}^2_M(t)\right)$$

with

$$\overline{\sigma}_M(t) \quad := \quad \sqrt{\frac{1}{t}\int_0^t \|\sigma_M(u)\|^2 \; du}.$$

In fact, since the choice of M is arbitrary, each L_n is a martingale (lognormal if σ_n is deterministic)

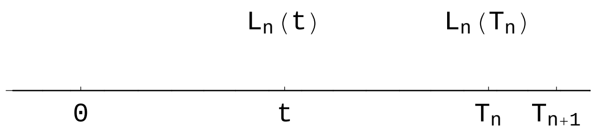


Figure 12.4: Lognormal Distribution of $L_n(t)$

under the forward measure P_{n+1} associated with T_{n+1}. □

Remark 12.3.4. These observations raise the question of whether we may in fact take the coefficients σ_n to be deterministic in (12.3.10) and (12.3.11). Recall form Example(11.1.4) that this choice(i.e. deterministic proportional volatility) is inadmissible in the HJM setting, essentially because it makes the HJM drift quadratic in the current level of ratios. To see what happens with simple compounding, rewrite (12.3.10) as

$$dL_n(t) \quad = \quad \sum_{j=\eta(t)}^{n} \frac{\delta_j L_j(t) L_n(t) \sigma_n^\top \sigma_j(t)}{1+\delta_j L_j(t)} dt + L_n(t)\sigma_n(t)^\top dW(t) \tag{12.3.15}$$

and consider the case of deterministic σ_i. The numerators in the drift are quadratic in the forward LIBOR rates, but they are stabilized by the terms $1+\delta_j L_j(t)$ in the denominators; indeed, because $L_j(t) \geq 0$ implies

$$\left|\frac{\delta_j L_j(t)}{1+\delta_j L_j(t)}\right| \leq 1,$$

the drift is linearly bounded in $L_n(t)$, making deterministic σ_i admissible. This feature is lost in the limit as the compounding period δ_j decreases to zero. Thus the distinction between continuous and simple forward rates turns out to have important mathematical as well as practical implications. □

12.3.1 Pricing Derivatives

✎ **Caplet**

Consider a caplet for the accrual period $[T_n, T_{n+1}]$. the underlying rate is $L_n(t)$ and the value $L_n(T_n)$ is fixed at T_n. With strike of K, the caplet's payoff is $\delta_n\big(L_n(T_n) - K\big)^+$. This payoff is made at T_{n+1}.

Let $C_n(t)$ denote the price of this caplet at time t; we know the terminal value $C_n(T_{n+1}) = \delta_n\big(L_n(T_n) - K\big)^+$. Under the spot measure, the deflated price $C_n(t)/P^*(t)$ must be a martingale so

$$C_n(t) \quad = \quad P^*(t)E^*\left[\frac{\delta_n\big(L_n(T_n) - K\big)^+}{P^*(T_{n+1})}\right]$$

where we have written E^* for expectation under the spot measure. Through $P^*(T_{n+1})$, this expectation involves the joint distribution of $L_1(T_1), \cdots, L_n(T_n)$, making its value difficult to discern. In contrast, under the forward measure P^{n+1} associated with maturity T_{n+1}, the martingale property applies to $C_n(t)/P_{n+1}(t)$. We may therefore also write

$$C_n(t) \quad = \quad P_{n+1}(t)E^{n+1}\left[\frac{\delta_n\big(L_n(T_n) - K\big)^+}{P_{n+1}(T_{n+1})}\right],$$

with E^{n+1} denoting expectation under P^{n+1}. Conveniently, $P_{n+1}(T_{n+1}) \equiv 1$, so this expectation depends only on the marginal distribution of $L_n(T_n)$.

If we take σ_n to be deterministic, then $L_n(T_n)$ has the lognormal distribution or

$$\log\frac{L_n(T_n)}{L_n(t)} \quad \sim \quad N\left(-\frac{1}{2}\overline{\sigma}_n^2(t, T_n), \overline{\sigma}_n^2(t, T_n)\right)$$

with

$$\overline{\sigma}_n(t, s) \quad := \quad \sqrt{\frac{1}{s-t}\int_t^s \|\sigma_n(u)\|^2\, du}.$$

In this case, the caplet price is given by the Black formula,

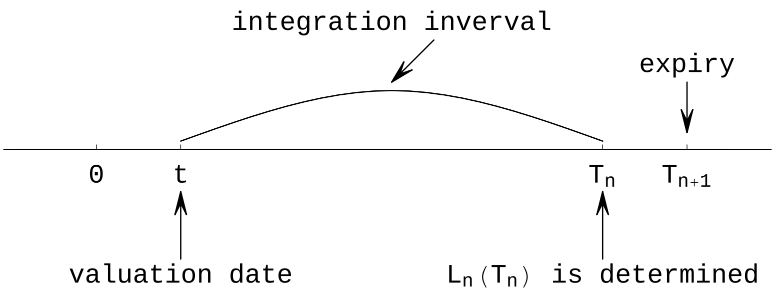


Figure 12.5: Time Horizon

$$C_n(t) \tag{12.3.16}$$

$$= \quad P_{n+1}(t)\left[L_n(t)N\left(\frac{\log\frac{L_n(t)}{K}+\overline{\sigma}_n^2(t,T_n)(T_n-t)}{\overline{\sigma}_n(t,T_n)\sqrt{T_n-t}}\right)-KN\left(\frac{\log\frac{L_n(t)}{K}-\overline{\sigma}_n^2(t,T_n)(T_n-t)}{\overline{\sigma}_n(t,T_n)\sqrt{T_n-t}}\right)\right]$$

since L_n is a martingale with respect to the forward measure P^{n+1} and $E^{n+1}\Big[L_n(T_n)\,|\,\mathscr{F}_t\Big] = L_n(t)$.

This formula is frequently used in the reverse direction. Given the market price of a caplet, one can solve for the "implied volatility" that makes the formula match the market price. This is useful in calibrating a model to market data.

✎ **Swaption**

Suppose the underlying swap begins at T_n with fixed- and floating-rate payments exchanged at $T_{n+1}, \cdots, T_{M+1}$. From (10.6.8) we find that the forward swap rate at time t is given by

$$S_n(t) \quad = \quad \frac{P_n(t)-P_{M+1}(t)}{\sum_{j=n+1}^{M+1}\delta_j P_j(t)}. \tag{12.3.17}$$

However the value of swaption cannot be evaluated explicitly and simulation is required.

Remark 12.3.5.

1. By applying Ito's formula to the swap rate (12.3.17), it is not difficult to conclude that if the forward LIBOR rates have deterministic volatilities, then the forward swap rate cannot also have a deterministic volatility.
2. In particular, then, the forward swap rate cannot be geometric Brownian motion under any equivalent measure.
3. One may choose the volatilities of the forward swap rates to be deterministic, and that in this case swaption prices are given by a variant of the Black formula. However the LIBOR rates cannot also have deterministic volatilities, so caplets are no longer priced by the Black formula. One must therefore choose between the two pricing formulas. The volatility structure is of a lower-triangular matrix of the following form:

□

12.3.2 Discretization For Simulation

An exact simulation is generally infeasible and some discretization error is inevitable. Because the models of this subsection deal with a finite set of maturities from the outset, we need only discretize the time argument, whereas in the HJM setting both time and maturity required discretization.

We fix a time grid $0 = t_0 < t_1 < \cdots < t_m < t_{m+1}$ over which to simulate. It is sensible to include the tenor dates $T_1, \cdots, T_{M+1}$ among the simulation dates. In practice, one would often even take $t_i = T_i$ so that the simulation evolves directly from one tenor date to the next.

Simulation of forward LIBOR rates is a special case of the general problem of simulating a system of SDEs. One can apply an Euler scheme or a higher-order method. However, even if we restrict ourselves to Euler schemes, there are countless alternatives. We have many choices of variables to discretize and many choices of probability measure under which to simulate. In this case, we use (12.3.15) to calculate $L_1(t_1), L_2(t_1), \cdots, L_M(t_1)$. We then use equation (12.3.15) again to calculate $L_2(t_2), L_3(t_2), \cdots, L_M(t_2)$; and so on until $L_M(t_M)$ is obtained. Note that as we move through time the zero curve gets shorter and shorter.

1. The most immediate application of the Euler scheme under the spot measure discretize the SED (12.3.15), producing
$$\begin{aligned}\widehat{L}_n(t_{i+1}) &= \widehat{L}_n(t_i) + \mu_n\big(\widehat{L}_n(t_i), t_i\big)\widehat{L}_n(t_i)(t_{i+1} - t_i) \\ &\quad + \widehat{L}_n(t_i)\sqrt{t_{i+1} - t_i}\;\sigma_n(t_i)^\top Z_{i+1}\end{aligned} \tag{12.3.18}$$
with
$$\mu_n\big(\widehat{L}_n(t_i), t_i\big) = \sum_{j=\eta(t_i)}^{n} \frac{\delta_j \widehat{L}_j(t_i)\sigma_n(t_i)^\top \sigma_j(t_i)}{1 + \delta_j \widehat{L}_j(t_i)}$$
and $Z_1, Z_2, \cdots$ independent $N(0, I)$ random vectors in $\mathbb{R}^d$. Here, we use hats to identify discretized variables. We assume that we are given an initial set of bond prices $P_1(0), \cdots, P_{M+1}(0)$ and initialize the simulation by setting
$$\widehat{L}_n(0) = \frac{P_n(0) - P_{n+1}(0)}{\delta_n P_{n+1}(0)}, \quad n = 1, \cdots, M,$$
in accordance with (12.3.1).

2. An alternative to (12.3.18) approximates the LIBOR rates under the spot measure using
$$\widehat{L}_n(t_{i+1}) = \widehat{L}_n(t_i)\exp\left\{\Big(\mu_n\big(\widehat{L}_n(t_i), t_i\big) - \frac{1}{2}\|\sigma_n(t_i)\|^2\Big)(t_{i+1} - t_i) + \sqrt{t_{i+1} - t_i}\;\sigma_n(t_i)^\top Z_{i+1}\right\}. \tag{12.3.19}$$
This is equivalent to applying an Euler scheme to $\log L_n$; it may also be viewed as approximating L_n by geometric Brownian motion over $[t_i, t_{t+1}]$, with drift and volatility parameters fixed at t_i. This method seems particularly attractive in the case of deterministic σ_n, since then L_n is close to lognormal. A further property of (12.3.19) is that it keeps all $\widehat{L}_n$ positive, whereas (12.3.18) can produce negative rates.

3. Under the forward measure P_{M+1}, the only modification necessary is to replace $\mu_n\big(\widehat{L}_n(t_i), t_i\big)$ with
$$\mu_n\big(\widehat{L}_n(t_i), t_i\big) = -\sum_{j=n+1}^{M} \frac{\delta_j \widehat{L}_j(t_i)\sigma_n(t_i)^\top \sigma_j(t_i)}{1 + \delta_j \widehat{L}_j(t_i)}.$$

✎ **Martingale Discretization**

In our discussion of simulation in the LMM setting, we will devote substantial attention to the issue of choosing the discrete drift to keep the model arbitrage-free even after discretization. It is therefore natural to examine whether an analogous choice of drift can be made in the LIBOR rate dynamics. In the LMM, the requirement is that

$$\widehat{D}_n(t_i) = \prod_{j=0}^{n-1} \frac{1}{1+\delta_j \widehat{L}_j(t_i \wedge T_j)} \tag{12.3.20}$$

be a martingale in i for each n under the spot measure; see (12.3.5). Under the forward measure, the martingale condition applies to

$$\widehat{D}_n(t_i) = \prod_{j=n}^{M} 1+\delta_j \widehat{L}_j(t_i);$$

see (12.3.12).

1. Consider the spot measure first. We would like, as a special case of (12.3.20), for $1/\big(1+\delta_1\widehat{L}_1\big)$ to be a martingale. Under the Euler scheme (12.3.18), this require

$$E\left[\frac{1}{1+\delta_1\widehat{L}_1(t_1)}\right] = \frac{1}{1+\delta_1\widehat{L}_1(0)}, \quad (n=2, i=1, t_1 \le T_1)$$

or

$$E\left[\frac{1}{1+\delta_1\widehat{L}_1(0)\big[1+\mu_1 t_1+\sqrt{t_1}\sigma_1^\top Z_1\big]}\right] = \frac{1}{1+\delta_1\widehat{L}_1(0)},$$

the expectation taken with respect to $Z_1 \sim N(0,I)$. However, because the denominator inside the expectation has a normal distribution, the expectation is infinite[2] no matter how we choose μ_1. There is no discrete drift that preserves the martingale property. If, instead, we use the method in (12.3.19), the condition becomes

$$E\left[\frac{1}{1+\delta_1\widehat{L}_1(0)\exp\Big\{\big(\mu_1-\|\sigma_1\|^2/2\big)t_1+\sqrt{t_1}\sigma_1^\top Z_1\Big\}}\right] = \frac{1}{1+\delta_1\widehat{L}_1(0)}.$$

In this case, there is a value of μ_1 for which this equation holds, but there is no explicit expression for it. The root of the difficulty lies in evaluating an expression of the form

$$E\left[\frac{1}{1+\exp(a+bZ)}\right], \quad Z \sim N(0,1),$$

which is effectively intractable.

[2]Notice that

$$\lim_{t\to 0}\int_t^\infty \frac{1}{x}e^{-\frac{x^2}{2}} = \infty.$$

2. An alternative strategy is to discretize and simulate the deflated bond prices themselves, rather than the forward LIBOR rates. For example, under the spot measure, the deflated bond prices satisfy

$$\begin{aligned}\frac{dD_{n+1}(t)}{D_{n+1}(t)} &= -\sum_{j=\eta(t)}^{n}\left(\frac{\delta_j L_j(t)}{1+\delta_j L_j(t)}\right)\ \sigma_j^\top(t)dW(t)\\ &= \sum_{j=\eta(t)}^{n}\left(\frac{D_{j+1}(t)}{D_j(t)}-1\right)\ \sigma_j^\top(t)dW(t).\end{aligned}$$

An Euler scheme for $\log D_{n+1}$ therefore evolves according to

$$\widehat{D}_{n+1}(t_{i+1}) = \widehat{D}_{n+1}(t_i)\exp\left(-\frac{1}{2}\|\widehat{\nu}_{n+1}(t_i)\|^2(t_{i+1}-t_i)+\sqrt{t_{i+1}-t_i}\,\widehat{\nu}_{n+1}(t_i)^\top Z_{i+1}\right) \tag{12.3.21}$$

with

$$\widehat{\nu}_{n+1}(t_i) = \sum_{j=\eta(t)}^{n}\left(\frac{\widehat{D}_{j+1}(t_i)}{\widehat{D}_j(t_i)}-1\right)\ \sigma_j(t_i).$$

in either case, the discretized deflated bond prices are automatically martingale; in (12.3.21) they are positive martingale and in this sense the discretization is arbitrage-free. From the simulated $\widehat{D}_n(t_i)$ we can then define the discretized forward LIBOR rates by setting

$$\widehat{L}_n(t_i) = \frac{1}{\delta_n}\left(\frac{\widehat{D}_n(t_i)-\widehat{D}_{n+1}(t_i)}{\widehat{D}_{n+1}(t_i)}\right), \quad n=1,\cdots,M.$$

Any other term structure variables (e.g. swap rates) required in the simulation can then be defined from $\widehat{L}_n$.

Glasserman and Zhao(2000) recommend replacing

$$\left(\frac{\widehat{D}_{j+1}(t)}{\widehat{D}_j(t)}-1\right) \quad \text{with} \quad \min\left\{\frac{\widehat{D}_{j+1}(t)}{\widehat{D}_j(t)}-1,0\right\}.$$

This modification has no effect in the continuous-time limit because $0\le D_{j+1}(t)\le D_j(t)$ (if $L_j(t)\ge 0$). But in the discretized process the ratio $\widehat{D}_{j+1}/\widehat{D}_j$ could potentially exceed 1.

12.3.3 Volatility Structure and Calibration

The volatility factors $\sigma_n(t)$ are chosen to calibrate a model to market prices of actively traded derivatives, especially caps and swaptions.

The variables $\sigma_n(t)$ are the primary determinants of both:

- the level of volatility in forward rates.
- the correlations between forward rates.

✎ **The level of volatility**

Suppose we are given the market price of caplet for the interval $[T_i, T_{i+1}]$ and from this price we calculate an impled volatility v_n by inverting the Black formula (12.3.16). If we choose σ_n to be any deterministic $\mathbb{R}^d$-valued function satisfying

$$\frac{1}{T_n}\int_0^{T_n} \|\sigma_n(t)\|^2\, dt \quad = \quad v_n^2. \tag{12.3.22}$$

By imposing this constraint on all of the σ_j, we ensure that the model is calibrated to all caplet prices.

Because LIBOR market models do not specify interest rates over accrual periods shorter than the interval $[T_i, T_{i+1}]$, it is natural and customary to restrict attention to functions $\sigma_n(t)$ that are constant between tenor dates. We take each σ_n to be right-continuous and thus denote by $\sigma_n(T_i)$ its value over the interval $[T_i, T_{i+1})$. It is convenient to think of the volatility structure as specified through a lower-triangular matrix of the following form:

$$\begin{pmatrix} \sigma_1(T_0) & & & \\ \sigma_2(T_0) & \sigma_2(T_1) & & \\ \vdots & \vdots & \ddots & \\ \sigma_M(T_0) & \sigma_M(T_1) & \cdots & \sigma_M(T_{M-1}) \end{pmatrix}, \quad \sigma_j(T_i) \in \mathbb{R}^d.$$

In a multifactor model, since the $\sigma_n(T_i)$ are now vectors, we find that

$$\frac{1}{T_n}\int_0^{T_n} \|\sigma_n(t)\|^2\, dt \quad = \quad \|\sigma_n(T_0)\|^2\delta_0 + \|\sigma_n(T_1)\|^2\delta_1 + \cdots + \|\sigma_n(T_{n-1})\|^2\delta_{n-1}.$$

The volatility structure is *stationary* if $\sigma_n(t)$ depends on n and t only through the difference $T_n - t$. For a stationary, single-factor, piecewise constant volatility structure, the volatility matrix takes the form

$$\begin{pmatrix} \sigma(1) & & & \\ \sigma(2) & \sigma(1) & & \\ \vdots & \vdots & \ddots & \\ \sigma(M) & \sigma(M-1) & \cdots & \sigma(1) \end{pmatrix}$$

for some values $\sigma(1), \cdots, \sigma(M)$. Think of $\sigma(i)$ as the volatility of a forward rate i period away from maturity. In this case, we find that

$$\frac{1}{T_n}\int_0^{T_n} \|\sigma_n(t)\|^2\, dt$$

$$
\begin{aligned}
&= \frac{1}{T_n}\int_0^{T_n} \|\sigma(T_n - t)\|^2 \, dt \\
&= \frac{1}{T_n}\int_0^{T_n} \|\sigma(t)\|^2 \, dt. \qquad (12.3.23)
\end{aligned}
$$

✎ **The correlations between forward rate**

The potential value of a multifactor model lies in capturing correlations between forward rates of different maturities. For example, we see that over a short time interval the correlation between the increment of $\log L_j(t)$ and $\log L_k(t)$ is approximately

$$
\frac{\sigma_k(t)^\top \sigma_j(t)}{\|\sigma_k(t)\| \, \|\sigma_j(t)\|}.
$$

These correlation are often chosen to match market prices of swaption (which, unlike caps, are sensitive to rate correlations) or to match historical correlations.

The hypothetical factor in Figure 12.6 are the first three principal components of the matrix

$$
0.12^2 \exp\Big(-0.8\sqrt{|i-j|}\Big), \quad i, j = 1, \cdots, 15.
$$

In a three-factor model, each $\sigma(i)$ has three components. The three curves in Figure 12.6 are graphs of

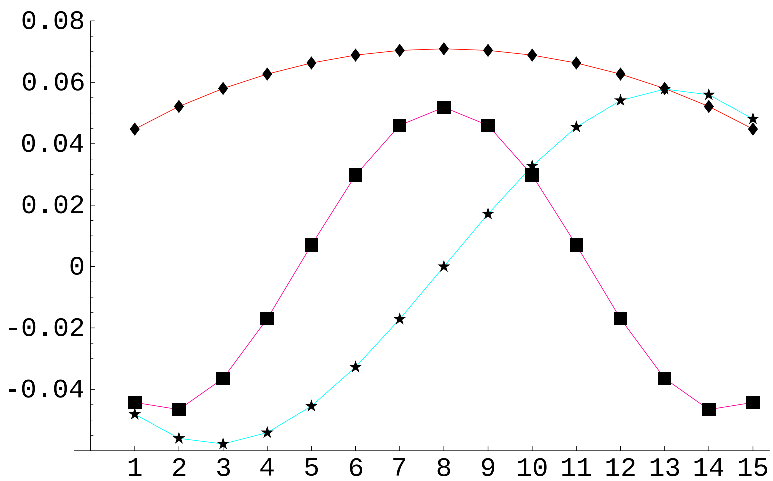


Figure 12.6: Volatility factors

the three components as functions of time to maturity. If we fix a time to maturity on the horizontal axis, the total volatility[3] at that point is given by the sum of squares of the three components: the inner products of these three-dimensional vectors at different times determine the covariances between the forward rates. More precisely, they are the first three eigenvectors of covariance matrix as ranked by their eigenvalues, scaled to have square of length equal to their eigenvalues.

It is common in practice to use the principal components of either the covariance matrix or the correlation matrix of changes in forward rates in choosing a factor structure.

[3]The total volatility is different to the implied volatility. The total volatility means $\|\sigma(t)\|$.

✎ A Case Study

Suppose we are given 3-month forward LIBOR rates in Table 12.1. From Table 12.1, we have annualized log return for 3-month forward LIBOR rates in Table 12.2. Through a principal component analysis, we have Table 12.3, Table 12.4, Table 12.5, and Table 12.6.

Date	1m	3m	6m	1y	2y	3y	5y	7y	10y	20y
2001-07-31	0.0367	0.0354	0.0347	0.0353	0.0379	0.0406	0.0457	0.0486	0.0507	0.0561
2001-08-01	0.0365	0.0353	0.0347	0.0356	0.0383	0.0409	0.0462	0.0490	0.0511	0.0563
2001-08-02	0.0365	0.0353	0.0346	0.0357	0.0389	0.0417	0.0469	0.0497	0.0517	0.0568
2001-08-03	0.0363	0.0352	0.0347	0.0357	0.0391	0.0422	0.0472	0.0499	0.0520	0.0570
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 12.1: 3-month forward LIBOR rates from 2001-07-31 to 2007-1-23.

Date	tcm1m	tcm3m	tcm6m	tcm1y	tcm2y	tcm3y	tcm5y	tcm7y	tcm10y	tcm20y
2001-08-01	-0.086	-0.045	0.000	0.134	0.166	0.116	0.172	0.130	0.124	0.056
2001-08-02	0.000	0.000	-0.046	0.044	0.246	0.306	0.238	0.224	0.185	0.140
2001-08-03	-0.087	-0.045	0.046	0.000	0.081	0.188	0.101	0.063	0.091	0.056
2001-08-06	-0.044	0.000	0.000	-0.044	-0.122	-0.188	-0.034	0.000	-0.030	0.000
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 12.2: Annualized log return for 3-month forward LIBOR: $\sqrt{250}\log\frac{L_n(0)^{i+1}}{L_n(0)^i}$

The procedure to get the volatility structure is given as follows:

1. Get zero curve by interpolating zero rate data.
2. Get Forward LIBOR rate from zero curve and via (12.3.1). $\longrightarrow$ Table 12.1.
3. Get annualized log-return and covariance matrix. $\longrightarrow$ Table 12.2.
4. Principal component analysis to data in Table 12.2. $\longrightarrow$ Table 12.3, Table 12.4, Table 12.5, and Table 12.6.
5. Estimate vector-valued function $\sigma(t)$. $\longrightarrow$ Table 12.5.
6. Monte carlo simulation.

	1m	3m	6m	1y	2y	3y	5y	7y	10y	20y
Mean	0.0036	0.0043	0.0046	0.0043	0.0031	0.0021	0.0006	-0.0001	-0.0006	-0.0013
StD	0.3609	0.2616	0.2634	0.3471	0.4193	0.3748	0.2937	0.2496	0.2134	0.1649
Var	0.1303	0.0684	0.0694	0.1205	0.1758	0.1405	0.0863	0.0623	0.0456	0.0272

Table 12.3: Simple Statistics.

	tcm1m	tcm3m	tcm6m	tcm1y	tcm2y	tcm3y	tcm5y	tcm7y	tcm10y	tcm20y
tcm1m	0.1303									
tcm3m	0.0534	0.0684								
tcm6m	0.0374	0.0527	0.0694							
tcm1y	0.0333	0.0485	0.0770	0.1205						
tcm2y	0.0311	0.0477	0.0764	0.1279	0.1758					
tcm3y	0.0269	0.0412	0.0655	0.1093	0.1526	0.1405				
tcm5y	0.0195	0.0299	0.0481	0.0804	0.1127	0.1054	0.0863			
tcm7y	0.0152	0.0229	0.0382	0.0652	0.0918	0.0866	0.0715	0.0623		
tcm10y	0.0119	0.0182	0.0306	0.0527	0.0742	0.0708	0.0597	0.0521	0.0456	
tcm20y	0.0068	0.0114	0.0203	0.0363	0.0514	0.0497	0.0429	0.0385	0.0338	0.0272

Table 12.4: Covariance Matrix

	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8	Prin9	Prin10
tcm1m	0.1145	0.3183	0.1193	-0.0386	0.0113	-0.0031	0.0003	0.0001	0.0002	0.0001
tcm3m	0.1446	0.1624	-0.0868	0.0985	-0.0544	0.0302	-0.0016	0.0000	-0.0005	0.0001
tcm6m	0.2068	0.0876	-0.1167	0.0226	0.0371	-0.0582	0.0086	0.0005	0.0004	0.0000
tcm1y	0.3171	0.0174	-0.0978	-0.0651	0.0636	0.0399	-0.0123	0.0011	-0.0002	0.0000
tcm2y	0.4056	-0.0520	-0.0043	-0.0614	-0.0596	0.0010	0.0338	-0.0112	0.0005	-0.0005
tcm3y	0.3638	-0.0578	0.0299	-0.0133	-0.0436	-0.0161	-0.0273	0.0286	-0.0007	-0.0003
tcm5y	0.2788	-0.0545	0.0517	0.0351	0.0053	-0.0078	-0.0247	-0.0276	0.0142	0.0032
tcm7y	0.2303	-0.0520	0.0574	0.0438	0.0249	0.0005	0.0006	-0.0075	-0.0245	0.0087
tcm10y	0.1893	-0.0467	0.0562	0.0506	0.0339	0.0064	0.0104	0.0036	0.0001	-0.0230
tcm20y	0.1335	-0.0407	0.0496	0.0483	0.0387	0.0123	0.0249	0.0181	0.0130	0.0130

Table 12.5: The eigenvectors scaled to have the square of length equal to their eigenvalues

Example 12.3.6. From Table 12.5, for example, we can find that the sum of squares of first elements of eigenvectors equals to the variance of the first bucket, or

$$0.1145^2 + 0.3183^2 + 0.1193^2 + \cdots + 0.0001^2 \quad = \quad 0.1303.$$

□

	Eigenvalue	Cumulative		Eigenvalue	Cumulative
1	0.65857467	0.7111	6	0.00641055	0.9921
2	0.15121104	0.8744	7	0.00345758	0.9958
3	0.05743289	0.9364	8	0.00209741	0.9981
4	0.02793828	0.9666	9	0.00097115	0.9992
5	0.01725305	0.9852	10	0.0007833	1

Table 12.6: Eigenvalues

Example 12.3.7. The three month volatility can be calculated as

$$\overline{\sigma}(0.25) \quad = \quad \sqrt{\frac{1}{0.25}\int_0^{0.25} \|\sigma(t)\|^2 dt}$$

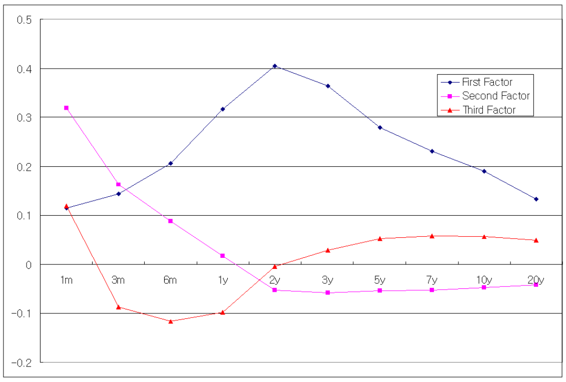


Figure 12.7: The first three eigenvectors

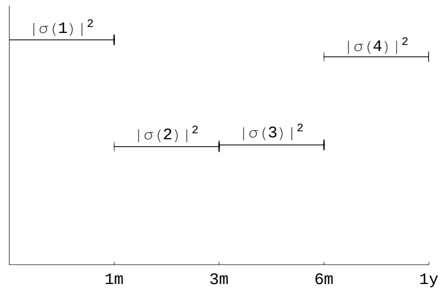


Figure 12.8: $\|\sigma(t)\|^2$ with $\|\sigma(i)\| =$ standard deviation of time bucket

$$
\begin{aligned}
&= \sqrt{\frac{1}{0.25}\left(\|\sigma(1)\|^2 \times \frac{1}{12} + \|\sigma(2)\|^2 \times \frac{2}{12}\right)} \\
&= \sqrt{\frac{1}{0.25}\left(0.1303 \times \frac{1}{12} + 0.0684 \times \frac{2}{12}\right)} \\
&= 0.298376,
\end{aligned}
$$

where $\|\sigma(i)\|$'s are given in Table 12.3. This volatility can be used for Black's formula with time to maturity of 3-month. Figure (12.9) shows $\overline{\sigma}(t)$ obtained by

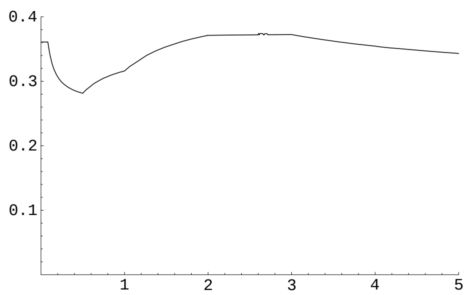


Figure 12.9: $\overline{\sigma}(t)$

$$\overline{\sigma}(t) = \sqrt{\frac{1}{t}\int_0^t \|\sigma(s)\|^2 ds}.$$

□

Example 12.3.8. The volatility fitting procedure can be summarized as follows:

1. Get implied volatilities from market prices of caps. i.e. v_n in (12.3.22).
2. Compute the forward volatility $\|\sigma(t)\|$ via (12.3.22) and (12.3.23).
3. Apply a PCA to the market data.
4. Rescale eigenvectors to have length equal to $\|\sigma(t)\|$ which is obtained from the second step. We then have a modified correlation structure.

Notice that if we cannot obtain caplet prices from market, we could alternatively have the length of eigenvecotrs equal to the square roots of eigenvalues (i.e. standard deviation of market data of each time bucket). □

Example 12.3.9. Suppose we are given eivenvector $\mathbf{e}_1, \cdots, \mathbf{e}_p$ and eigenvalues $\lambda_1, \cdots, \lambda_p$ by a principal component analysis. Also assume that

$$\begin{aligned} \mathbf{e}_i &= (e_{i1}, e_{i2}, \cdots, e_{ip}), \quad i = 1, 2, \cdots, p, \\ \|\mathbf{e}_i\| &= 1, \\ \mathbf{e}_i \cdot \mathbf{e}_j &= \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \end{aligned}$$

1. If the number of factors used in LIBOR market model, d, is equal to the total number of factors, p, it is correct to set

   $$\sigma_{ij} = e_{ij} \times \sqrt{\lambda_i},$$

   where σ_{ij} is i-th component of $\sigma(j) \in \mathbb{R}^d$. Table 12.5 can be obtained from this way.

2. When $d < p$, the σ_{ij} must be scaled so that

   $$\|\sigma(j)\| = \sqrt{\sum_{k=1}^{d} \sigma_{kj}^2},$$

   where $\|\sigma(j)\|$ is obtained from market prices of caplets. This evolves setting

   $$\sigma_{ij} = \frac{\|\sigma(j)\| \times e_{ij} \times \sqrt{\lambda_i}}{\sqrt{\sum_{k=1}^{d} e_{kj}^2 \lambda_k}}.$$

   See Figure 12.10.

To define forward volatility function, assume that the time bucket $\widetilde{t}_1 < \widetilde{t}_2 < \cdots < \widetilde{t}_p$ are given. Now define forward volatilitity function as follows:

$$\sigma_i(t) := \begin{cases} \sigma_{i1}, & t < \widetilde{t}_1 \\ \sigma_{ij}, & \widetilde{t}_{j-1} \leq t < \widetilde{t}_j, \quad j = 2, \cdots, p-1 \\ \sigma_{ip}, & t \geq \widetilde{t}_{p-1} \end{cases}$$

$$
\begin{array}{c}
\begin{array}{ccccccc}
\|\sigma(1)\| & \|\sigma(2)\| & & \|\sigma(d)\| & \cdots & & \\
\nearrow & \nearrow & & \nearrow & \nearrow & &
\end{array}\\
\left(\begin{array}{ccccccc}
\boxed{e_{11}} & \boxed{e_{12}} & \cdots & \boxed{e_{1j}} & \boxed{e_{1j+1}} & \cdots & \boxed{e_{1p}} \\
\boxed{e_{21}} & \boxed{e_{22}} & \cdots & \boxed{e_{2j}} & \boxed{e_{2j+1}} & \cdots & \boxed{e_{2p}} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\boxed{e_{d1}} & \boxed{e_{d2}} & \cdots & \boxed{e_{dj}} & \boxed{e_{dj+1}} & \cdots & \boxed{e_{dp}} \\
e_{d+11} & e_{d+12} & \cdots & e_{d+1j} & e_{d+1j+1} & \cdots & e_{d+1p} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
e_{p1} & e_{p2} & \cdots & e_{pj} & e_{pj+1} & \cdots & e_{pp}
\end{array}\right)
\begin{array}{ccc}
\leftarrow & \times & \sqrt{\lambda_1} \\
\leftarrow & \times & \sqrt{\lambda_2} \\
\vdots & & \vdots \\
\leftarrow & \times & \sqrt{\lambda_d} \\
\leftarrow & \times & \sqrt{\lambda_{d+1}} \\
\vdots & & \vdots \\
\leftarrow & \times & \sqrt{\lambda_p}
\end{array}
\end{array}
$$

Figure 12.10: Volatility matrix

for $i = 1, 2, \cdots, d$. Then we have a right continuous forward volatility function $\sigma_i(t)$ and

$$\|\sigma(t)\| \quad = \quad \sqrt{\sum_{k=1}^{d} \sigma_k^2(t)}.$$

Notice that the implied volatility function is given by

$$\overline{\sigma}(t) \quad = \quad \sqrt{\frac{1}{t}\int_0^t \|\sigma(s)\|^2 ds}.$$

See Figure 12.11. □

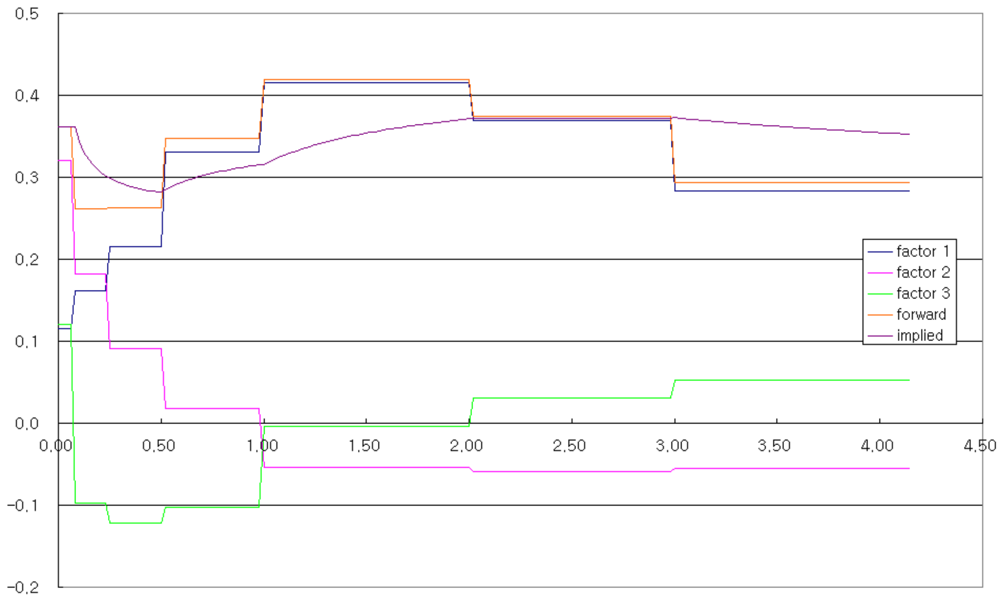


Figure 12.11: Volatility factors

Chapter 13

Interest Rate Trees

If the time step on the interest tree is Δt, the rates on the tree are the continuously compounded Δt-period rates. The usual assumption when a tree is constructed is that the Δt-period zero rate, R, follows the same stochastic process as the short rate, r, in the corresponding continuous time model.

13.1 Hull-White Two-Stage Procedure

13.1.1 First Stage

The Hull-White model for the instantaneous short rate r_t is

$$dr_t = \Big[\Phi(t) - ar_t\Big]dt + \sigma dW_t.$$

We suppose that the time step on the tree is constant and equal to Δt. We assume that the Δt-rate, $R(t) = R(t, t + \Delta t)$, follows the same process as r_t.

$$dR(t) = \Big[\Phi(t) - aR(t)\Big]dt + \sigma dW_t.$$

The first stage is to construct a tree for a variable $R^*(t)$ that follows the process

$$dR^*(t) = -aR^*(t)dt + \sigma dW_t, \quad R^*(0) = 0.$$

Note that $R^*(t)$ follows the Ornstein-Uhlenbeck process. Hence, by Example (2.2.6),

$$\begin{aligned}
e^{a(t+\Delta t)}R^*(t+\Delta t) &= e^{at}R^*(t) + \sigma\int_t^{t+\Delta t} e^{as}dW_s, \\
R^*(t+\Delta t) &= e^{-a\Delta t}R^*(t) + \sigma e^{-a(t+\Delta t)}\int_t^{t+\Delta t} e^{as}dW_s, \\
R^*(t+\Delta t) - R^*(t) &= (e^{-a\Delta t} - 1)R^*(t) + \sigma e^{-a(t+\Delta t)}\int_t^{t+\Delta t} e^{as}dW_s.
\end{aligned}$$

If terms of higher order than Δt are ignored,

$$\begin{aligned} E\Big[R^*(t+\Delta t) - R^*(t)\Big|\mathscr{F}_t\Big] &= (e^{-a\Delta t}-1)R^*(t) \\ &\approx -a\Delta t R^*(t), \qquad (13.1.1) \\ \mathrm{Var}\Big[R^*(t+\Delta t) - R^*(t)\Big|\mathscr{F}_t\Big] &= \sigma^2 e^{-2a(t+\Delta t)}\int_t^{t+\Delta t} e^{2as}ds \\ &= \frac{1}{2a}\sigma^2(1-e^{-2a\Delta t}) \\ &\approx \sigma^2\Delta t. \qquad (13.1.2) \end{aligned}$$

We denote ΔR as the spacing between interest rate on the tree and set

$$\sigma^2\frac{\Delta t}{\Delta R^2} = \frac{1}{3}, \quad \text{i.e. } \Delta R = \sigma\sqrt{3\Delta t}.$$

This proves to be a good choice of ΔR from the viewpoint of error minimization.

Denote (i, j) as the node where $t = i\Delta t$ and $R^* = j\Delta R$. The value i is positive integer and j is a positive or negative integer.

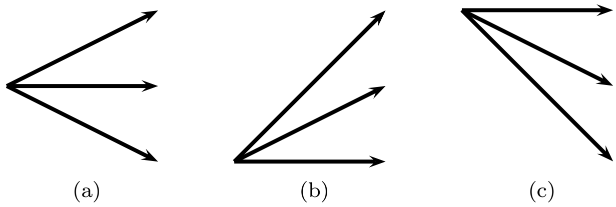


Figure 13.1: Alternative branching methods in a trinomial tree.

If the branching pattern from node (i, j) is as in Figure 13.1(a), the equations for p_u, p_m and p_d are

$$\begin{aligned} p_u\Delta R - p_d\Delta R &= -aj\Delta R\Delta t, \quad \text{by (13.1.1)} \\ p_u\Delta R^2 + p_d\Delta R^2 &= \sigma^2\Delta t + (aj\Delta R\Delta t)^2, \quad \text{by (13.1.2)} \\ p_u + p_m + p_d &= 1. \end{aligned}$$

Using $\Delta R^2 = 3\sigma^2\Delta t$, these equations become

$$\begin{aligned} p_u - p_d &= -\rho, \\ p_u + p_d &= \frac{1}{3} + \rho^2 \end{aligned}$$

where $\rho = aj\Delta t$. The solution to these equations is

$$\begin{aligned} p_u &= \frac{1}{6} + \frac{\rho^2 - \rho}{2} \\ p_m &= \frac{2}{3} - \rho^2 \\ p_d &= \frac{1}{6} + \frac{\rho^2 + \rho}{2}. \end{aligned}$$

Similarly, the probabilities in another branching patterns are given by Table (13.1). Since the probabilities must be positive for any value of j,

$$\frac{0.184}{a\Delta t} \approx \frac{1-\sqrt{2/3}}{a\Delta t} \le j_{\max} \le \frac{\sqrt{2/3}}{a\Delta t} \approx \frac{0.816}{a\Delta t}.$$

branching pattern 13.1(b)	branching pattern 13.1(c)
$p_u(2\Delta R) + p_m\Delta R = -aj\Delta R\Delta t$ $p_u(4\Delta R^2) + p_m\Delta R^2 = \sigma^2\Delta t + (aj\Delta R\Delta t)^2$	$p_d(-2\Delta R) + p_m(-\Delta R) = -aj\Delta R\Delta t$ $p_d(4\Delta R^2) + p_m\Delta R^2 = \sigma^2\Delta t + (aj\Delta R\Delta t)^2$
$2p_u + p_m = -\rho$ $4p_u + p_m = \frac{1}{3} + \rho^2$	$2p_u + p_m = \rho$ $4p_u + p_m = \frac{1}{3} + \rho^2$
$p_u = \frac{1}{6} + \frac{\rho^2+\rho}{2}$ $p_m = -\frac{1}{3} - \rho^2 - 2\rho$ $p_d = \frac{7}{6} + \frac{\rho^2+3\rho}{2}$	$p_u = \frac{7}{6} + \frac{\rho^2-3\rho}{2}$ $p_m = -\frac{1}{3} - \rho^2 + 2\rho$ $p_d = \frac{1}{6} + \frac{\rho^2-\rho}{2}$

Table 13.1: Probabilities for each branching pattern.

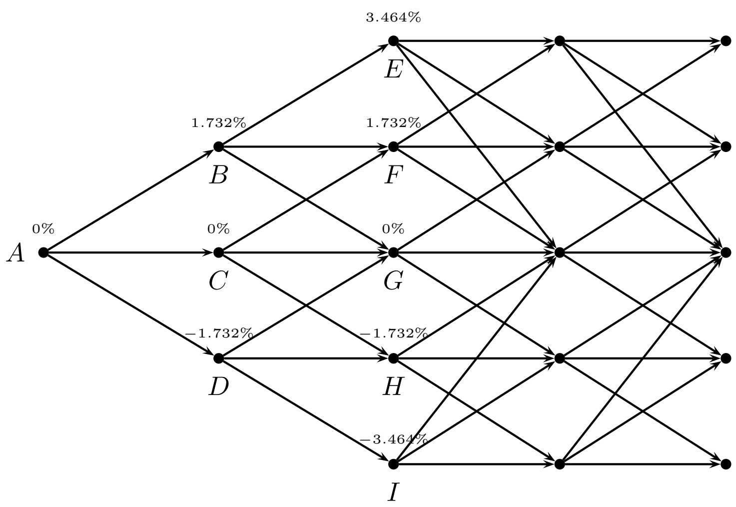


Figure 13.2: Tree for R^* in Hull-White model($\sigma = 0.01, a = 0.1, \Delta t = 1$)

13.1.2 Second Stage

The second stage in the tree construction is to convert the tree for R^* into a tree for R Define

$$\alpha(t) = R(t) - R^*(t).$$

Because

$$dR(t) = \Big[\Phi(t) - aR(t)\Big]dt + \sigma dW_t,$$

$$dR^*(t) = -aR^*(t)dt + \sigma dW_t$$

it follow that

$$\begin{aligned} d\alpha &= dR - dR^* \\ &= \Big[\Phi(t) - a\alpha(t)\Big]dt. \end{aligned}$$

Define α_i as $\alpha(i\Delta t)$, the value of $R - R^*$ at time $i\Delta t$. Define Q_{ij} as the present value of a security that pays off 1 if node (i, j) is reached and zero otherwise.

$$Q_{0,0} \to \alpha_0 \to Q_{1,\cdot} \to \alpha_1 \to Q_{2,\cdot} \to \alpha_2 \to \cdots .$$

1. By definition $Q_{00} = 1$.
2. $e^{-(R^*(0)+\alpha_0)} = P(0, \Delta t)$
3. $Q_{1,j} = \sum_k Q_{0,k} p_{kj} e^{-(k+\Delta R+\alpha_0)\Delta t}$.
4. ...

To express the approach more formally, we suppose that the Q_{ij} have been determined for $i \leq m(m \geq 0)$. The next step is to determine α_m so that the tree correctly prices a zero-coupon bond maturing at $(m+1)\Delta t$. The interest rate at node (m, j) is $\alpha_m + j\Delta R$, so that the price of a zero-coupon bond maturing at time $(m+1)\Delta t$ is given by

$$P\big(0, (m+1)\Delta t\big) = \sum_j Q_{m,j}\ e^{-(\alpha_m + j\Delta R)}.$$

Once α_m has been determined, the $Q_{m+1,j}$ can be calculated using

$$Q_{m+1,j} = \sum_k Q_{m,k}\ p_{kj}\ e^{-(\alpha_m + k\Delta R)\Delta t}.$$

Part V

Implements

Chapter 14

Tree Approaches

1. The advantage of tree approach is that it considerably reduces the number of computation.
2. The value given by a tree could be converges slowly to the exact value as the number of time steps is increased. For example, in barrier options, the barrier being assumed by the tree is different from the barrier true barrier.
3.

14.1 Path-Dependent Derivatives(Hull-White)

14.2 Lookback Options(Reiner)

This approach was proposed by E. Reiner in a lecture at Berkeley.

14.3 Barrier Options

✎ **Positioning nodes on the barriers**

✎ **Adjusting for nodes not lying on barriers**
The first value is obtained by assuming that the inner barrier is correct. The second value is obtained by assuming that the outer barrier is correct. A final estimate for the value of the derivative is then obtained by interpolating between these two values. For details, see Derman et al.(1995).

✎ **The Adaptive mesh modes**

Chapter 15

Finite Difference Methods

Two of the most popular procedures for valuing derivative securities are the lattice(or tree) approach and the finite difference approach(Refer to Hull & White(1990)).

1. The lattice approach was suggested by Cox, Ross, and Rubinstein(1979).
2. The finite difference approach was suggested by Schwartz(1977) and Brennan and Schwartz(1978). There are two alternative ways of implementing the finite difference approach. The first is the explicit finite difference method and the second is the implicit finite difference method.
   (a) Explicit: Brennan and Schwartz show that the explicit finite difference method is equivalent to a trinomial lattice approach.
   (b) Implicit: Brennan and Schwartz also show that the implicit finite difference method corresponds to a multinomial lattice approach where, in the limit, the underlying variable can move from its value at time t to an infinity of possible values at time $t + \Delta t$.

Geske and Shastri(1985) provide an interesting comparison of lattice and finite difference approaches. They conclude that the explicit finite difference method, with logarithmic transformations, is the most efficient approach when large numbers of stock options are being evaluated. The explicit finite difference methods is also attractive for a number of other reasons. It is computationally much simpler than the implicit method since it does not require the inversion of matrices. It is conceptually simpler than the implicit method since it is, in effect, nothing more than an application of the trinomial lattice approach. The method's only disadvantages is that the numerical solution does not necessarily converge to the solution of the differential equations as Δt tends to zero.

There are various FDMs to deal with several spatial variables. All suffer from the fact that the computational burden grows exponentially with the number of spatial variables. Due to this *curse of dimension*, problems involving more than three variables are usually solved with Monte Carlo simulation.

15.1 The Explicit FDM & Trinomial Tree

Consider a derivative security with price f that depends on a single asset S. Suppose that the stochastic process followed by S is

$$dS = S\mu(S,t)dt + S\sigma(S,t)dW,$$

where dW is a Wiener process under a physical measure. The variable μ and σ, which may be function of S and t, are the instantaneous proportional drift rate and volatility of S.

If λ is the market price of risk of S, then f must satisfy the differential equation

$$\frac{\partial f}{\partial t} + (\mu - \lambda\sigma)S\frac{\partial f}{\partial S} + \frac{1}{2}S^2\sigma^2\frac{\partial^2 f}{\partial S^2} = rf, \tag{15.1.1}$$

where r is risk-free interest rate.

To implement the explicit finite difference method, a small time interval Δt and a small change ΔS in S are chosen. A grid is then constructed for considering values of f when S is equal to

$$S_0, \quad S_0 + \Delta S, \quad S_0 + 2\Delta S, \quad \cdots, \quad S_0 + N\Delta S = S_N$$

and time is equal to

$$t_0, \quad t_0 + \Delta t, \quad t_0 + 2\Delta t, \quad \cdots, \quad t_0 + M\Delta t = t_M = T.$$

We will denote $t_0 + i\Delta t$ by t_i, $S_0 + j\Delta S$ by S_j, and the value of the derivative security at the (i,j) point on the grid by f_{ij}.

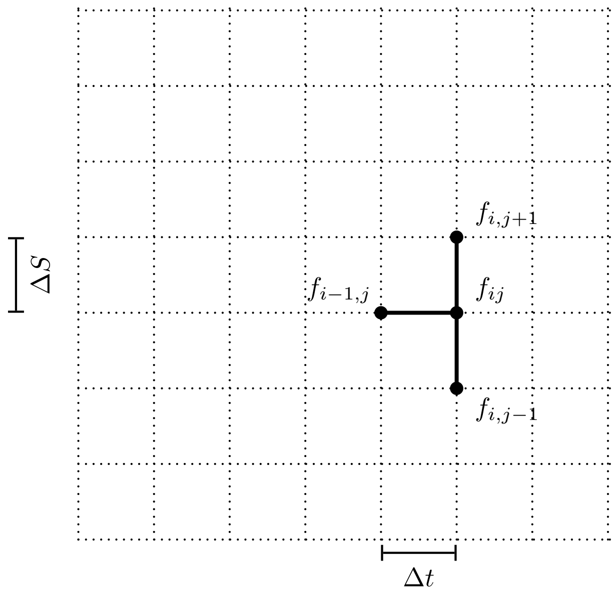


Figure 15.1: Explicit Finite-Difference Discretization.

The partial derivatives of f with respect to S at node (i,j) are approximated as

$$\frac{\partial f}{\partial S} = \frac{f_{i,j+1} - f_{i,j-1}}{2\Delta S} \tag{15.1.2}$$

$$\frac{\partial^2 f}{\partial S^2} = \frac{f_{i,j+1} - 2f_{ij} + f_{i,j-1}}{(\Delta S)^2}, \tag{15.1.3}$$

and the time derivative is approximated as

$$\frac{\partial f}{\partial t} = \frac{f_{ij} - f_{i-1,j}}{\Delta t}.$$

Substituting (15.1.2),(15.1.3), and (15.1.4) into (15.1.1)gives

$$f_{i-1,j} = a_{j-1} f_{i,j-1} + a_j f_{ij} + a_{j+1} f_{i,j+1}, \tag{15.1.4}$$

where

$$\begin{aligned} a_{j-1} &= \frac{1}{1+r\Delta t}\left(-\frac{(\mu-\lambda\sigma)S_j\Delta t}{2\Delta S} + \frac{S_j^2\sigma^2\Delta t}{2(\Delta S)^2}\right), \\ a_j &= \frac{1}{1+r\Delta t}\left(1 - \frac{S_j^2\sigma^2\Delta t}{(\Delta S)^2}\right), \\ a_{j+1} &= \frac{1}{1+r\Delta t}\left(\frac{(\mu-\lambda\sigma)S_j\Delta t}{2\Delta S} + \frac{S_j^2\sigma^2\Delta t}{2(\Delta S)^2}\right). \end{aligned}$$

Define

$$\begin{aligned} p_{j,j-1} &= -\frac{(\mu-\lambda\sigma)S_j\Delta t}{2\Delta S} + \frac{S_j^2\sigma^2\Delta t}{2(\Delta S)^2}, \\ p_{jj} &= 1 - \frac{S_j^2\sigma^2\Delta t}{(\Delta S)^2}, \\ p_{j,j+1} &= \frac{(\mu-\lambda\sigma)S_j\Delta t}{2\Delta S} + \frac{S_j^2\sigma^2\Delta t}{2(\Delta S)^2}. \end{aligned} \tag{15.1.5}$$

so that equation (15.1.4) becomes

$$f_{i-1,j} = \frac{1}{1+r\Delta t}\Big(p_{j,j-1} f_{i,j-1} + p_{jj} f_{ij} + p_{j,j+1} f_{i,j+1}\Big). \tag{15.1.6}$$

The variable $p_{j,j-1}$, $p_{j,j}$, and $p_{j,j+1}$ can be interpreted as the probabilities of moving from S_j to S_{j-1}, S_j, and S_{j+1}, respectively, during time Δt, in a risk-neutral world.

This is because

$$\begin{aligned} p_{j,j+1} + p_{jj} + p_{j,j-1} &= 1, \\ p_{j,j+1}(S_j + \Delta S) + p_{jj}S_j + p_{j,j-1}(S_j - \Delta S) &= S_j\Big(1 + (\mu - \lambda\sigma)\Delta t\Big) \\ &= S_j(1 + r\Delta t), \end{aligned}$$

and

$$\begin{aligned} & p_{j,j+1}(S_j + \Delta S)^2 + p_{jj}S_j^2 + p_{j,j-1}(S_j - \Delta S)^2 \\ = \;& p_{j,j+1}\Big(S_j^2 + 2S_j\Delta S + (\Delta S)^2\Big) + p_{jj}S_j^2 + \Big(S_j^2 - 2S_j\Delta S + (\Delta S)^2\Big) \\ = \;& S_j^2 + 2S_j\Delta S(p_{j+1} - p_{j-1}) + (\Delta S)^2(p_{j+1} + p_{j-1}) \end{aligned}$$

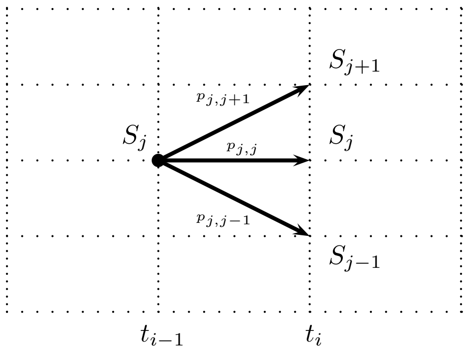


Figure 15.2: Probabilistic Interpretation.

$$
\begin{aligned}
&= S_j^2 + 2S_j^2 r\Delta t + S_j^2\sigma^2\Delta t \\
&= S_j^2(1+2r\Delta t) + S_j^2\sigma^2\Delta t \\
&\approx S_j^2(1+r\Delta t)^2 + S_j^2\sigma^2\Delta t, \quad \text{(if terms of } O(\Delta t^2) \text{ are ignored).}
\end{aligned}
$$

When p's are interpreted in this way, (15.1.6) gives the value of f at time t_{i-1} as its expected value at time t_i, in a risk-neutral world, discounted to time t_{i-1} at the risk-free rate of interest.

We can conclude from this that the explicit FDM is equivalent to a trinomial lattice approach.

15.1.1 The Transformation of Variables

The probabilities p_{ij}, in Eq. (15.1.5) are dependent on stock price S_j. It is efficient to use $z = \log S$ rather than S. After transformation, a grid is constructed for values of z equal to

$$z_0, \quad z_0 + \Delta z, \quad z_0 + 2\Delta z, \quad \cdots, \quad z_N,$$

where $z_j = z_0 + j\Delta z$, and PDE becomes

$$\frac{\partial f}{\partial t} + (r - \frac{1}{2}\sigma^2)\frac{\partial f}{\partial z} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial z^2} \quad = \quad rf,$$

and the probabilities in (15.1.5) become

$$p_{j,j-1} = -\frac{1}{2}\,(r - \frac{1}{2}\sigma^2)\frac{\Delta t}{\Delta z} + \frac{\sigma^2}{2}\,\frac{\Delta t}{(\Delta z)^2}, \tag{15.1.7}$$

$$p_{jj} = 1 - \sigma^2\,\frac{\Delta t}{(\Delta z)^2}, \tag{15.1.8}$$

$$p_{j,j+1} = \frac{1}{2}\,(r - \frac{1}{2}\sigma^2)\frac{\Delta t}{\Delta z} + \frac{\sigma^2}{2}\,\frac{\Delta t}{(\Delta z)^2}. \tag{15.1.9}$$

Note that $p_{j,j-1}, p_{jj}$ and $p_{j,j+1}$ are independent of j. If the grid is selected so that

$$\frac{\Delta t}{(\Delta z)^2} = \frac{1}{\sigma^2}, \quad \text{i.e. } p_{jj} = 0$$

then a binomial lattice results.

15.1.2 Stability of the Explicit FDM

We begin by taking the discrete Fourier transform of

$$p_{j,j-1}f_{i,j-1} + p_{jj}f_{ij} + p_{j,j+1}f_{i,j+1}.$$

Put

$$A = 2\sigma^2\,\frac{\Delta t}{(\Delta z)^2}, \qquad B = (r - \frac{1}{2}\sigma^2)\frac{\Delta t}{\Delta z}.$$

Then

$$p_{j,j-1} = -\frac{1}{2}B + \frac{1}{4}A, \quad p_{jj} = 1 - \frac{1}{2}A, \quad p_{j,j+1} = \frac{1}{2}B + \frac{1}{4}A,$$

and the discrete Fourier transform is given by

$$\begin{aligned} & \left[\left(-\frac{B}{2} + \frac{1}{4}A\right)e^{-i\xi} + \left(1 - \frac{1}{2}A\right) + \left(\frac{B}{2} + \frac{1}{4}A\right)e^{i\xi}\right]\widehat{f}(\xi) \\ = & \left[\frac{B}{2}(e^{i\xi} - e^{-i\xi}) + \left(1 - A\sin^2\frac{\xi}{2}\right)\right]\widehat{f}(\xi) \\ = & \left[iB\sin\xi + \left(1 - A\sin^2\frac{\xi}{2}\right)\right]\widehat{f}(\xi). \end{aligned}$$

If

$$B^2\sin^2\xi + \left(1 - A\sin^2\frac{\xi}{2}\right)^2 < 1, \tag{15.1.10}$$

the explicit scheme converges.

$$\begin{aligned} & B^2\sin^2\xi + \left(1 - A\sin^2\frac{\xi}{2}\right)^2 \\ = \; & 4B^2(\sin^2\frac{\xi}{2} - \sin^4\frac{\xi}{2}) + \left(1 - A\sin^2\frac{\xi}{2}\right)^2 \\ = \; & (A^2 - 4B^2)\sin^4\frac{\xi}{2} - (2A - 4B^2)\sin^2\frac{\xi}{2} + 1 \end{aligned}$$

From this we can see that the condition (15.1.10) is satisfied if

$$A^2 - 4B^2 > 0 \quad \text{and} \quad A^2 - 4B^2 < 2A - 4B^2,$$
$$\text{i.e.} \quad A^2 - 4B^2 > 0 \quad \text{and} \quad A^2 < 2A.$$

Now we have a sufficient condition of convergence:

$$\sigma^2\,\frac{\Delta t}{(\Delta z)^2} < 1 \tag{15.1.11}$$

and

$$|r - \frac{1}{2}\sigma^2|\frac{\Delta t}{\Delta z} < \sigma^2\,\frac{\Delta t}{(\Delta z)^2}, \quad \text{i.e.} \quad |r - \frac{1}{2}\sigma^2| < \frac{\sigma^2}{\Delta z}. \tag{15.1.12}$$

Note that condition (15.1.11) and (15.1.12) are equivalent to the conditions that probabilities $p_{j,j-1}, p_{jj}$ and $p_{j,j+1}$ are positive.

15.2 The Crank-Nicolson Scheme

✎ **Preparations**

The Black-Scholes equation

$$\frac{\partial f}{\partial t} + (r-q)S\frac{\partial f}{\partial S} + \frac{1}{2}S^2\sigma^2\frac{\partial^2 f}{\partial S^2} = rf, \quad S>0, 0 \le t \le T$$

is equivalent to the heat equation

$$\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial x^2} \tag{15.2.1}$$

by means of the transformations

$$\begin{aligned} S &= Ke^x, \quad t = T - \frac{2\tau}{\sigma^2}, \quad 0 \le t \le T, 0 \le \tau \le \frac{\sigma^2}{2}T, \\ f(S,t) &= K\exp\left\{-\frac{1}{2}(k_2-1)x - \left(\frac{1}{4}(k_2-1)^2 + k_1\right)\tau\right\}v(x,\tau), \end{aligned} \tag{15.2.2}$$

where

$$k_1 := \frac{2r}{\sigma^2}, \quad k_2 := \frac{2(r-q)}{\sigma^2},$$

and K is an arbitrary positive constant. For details, refer to Subsection 2.4.3.

The original domain $S>0, 0 \le t \le T$ becomes the strip

$$-\infty < x < \infty, \quad 0 \le \tau \le \frac{1}{2}\sigma^2 T.$$

The step in τ is

$$\Delta\tau = \frac{\tau_{\max}}{M}, \quad \tau_{\max} = \frac{1}{2}\sigma^2 T.$$

The infinite interval $-\infty < x < \infty$ must be replaced by a finite interval $x_{\min} \le x \le x_{\max}$. For a suitable integer N the step length is defined by

$$\Delta x = \frac{x_{\max} - x_{\min}}{N}.$$

	Explicit	Implicit	Crank-Nicolson
convergence rate	$O(\Delta\tau)$	$O(\Delta\tau)$	$O\big((\Delta\tau)^2\big)$
stable condition	$\frac{\Delta\tau}{\Delta x^2} \le \frac{1}{2}$	unconditionally stable	unconditionally stable

Table 15.1: Convergence rate and stable condition of finite difference methods

The advantage of the Crank-Nicolson method is that it has faster convergence than either the explicit and implicit method. The Crank-Nicolson has $O((\Delta\tau)^2)$ rates of convergence to the solution of the partial differential equation. Recall that the rate of convergence of the implicit and explicit method is $O(\Delta\tau)$ [1]. See Table 15.1.

[1] See Wilmott, Dewynne and Dowison(1993).

15.2.1 Implementation(Crank-Nicolson)

Let

$$v_n^m \quad := \quad v(x_{\min} + n\Delta x, m\Delta\tau).$$

The Crank-Nicolson scheme for (15.2.1) is given by

$$\frac{v_n^{m+1} - v_n^m}{\Delta\tau} + O(\Delta\tau) \quad = \quad \frac{v_{n+1}^m - 2v_n^m + v_{n-1}^m}{2\Delta x^2} + \frac{v_{n+1}^{m+1} - 2v_n^{m+1} + v_{n-1}^{m+1}}{2\Delta x^2} + O(\Delta x^2).$$

Ignoring terms of $O(\Delta\tau)$ and $O(\Delta x^2)$, we can approximate this by

$$\begin{aligned} \left(1 - \frac{\lambda}{2}\delta^2\right) u_n^{m+1} \quad &= \quad \left(1 + \frac{\lambda}{2}\delta^2\right) u_n^m \qquad (15.2.3) \\ \text{where} \quad \lambda \quad &= \quad \frac{\Delta\tau}{\Delta x^2}, \quad \delta^2 u_n \quad = \quad u_{n+1} - 2u_n + u_{n-1}. \end{aligned}$$

We use u_n^m to emphasize that the solution of (15.2.3) is only an approximate solution of the differential equation, not the exact solution, since we have derived (15.2.3) by neglecting terms of $O(\Delta\tau)$ and $O(\Delta x^2)$.

Let

$$Q_1 \quad = \quad \begin{bmatrix} 1+\lambda & -\frac{\lambda}{2} & 0 & \cdots & 0 & 0 & 0 \\ -\frac{\lambda}{2} & 1+\lambda & -\frac{\lambda}{2} & \cdots & 0 & 0 & 0 \\ 0 & -\frac{\lambda}{2} & 1+\lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\frac{\lambda}{2} & 1+\lambda & -\frac{\lambda}{2} \\ 0 & 0 & 0 & \cdots & 0 & -\frac{\lambda}{2} & 1+\lambda \end{bmatrix},$$

$$Q \quad = \quad \begin{bmatrix} 1-\lambda & \frac{\lambda}{2} & 0 & \cdots & 0 & 0 & 0 \\ \frac{\lambda}{2} & 1-\lambda & \frac{\lambda}{2} & \cdots & 0 & 0 & 0 \\ 0 & \frac{\lambda}{2} & 1-\lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{\lambda}{2} & 1-\lambda & \frac{\lambda}{2} \\ 0 & 0 & 0 & 0 & 0 & \frac{\lambda}{2} & 1-\lambda \end{bmatrix}$$

and

$$\begin{aligned} \mathbf{u}^{m+1} = (u_0^{m+1}, \cdots, u_N^{m+1})^T, \quad \mathbf{u}^m = (u_0^m, \cdots, u_N^m)^T, \\ \tilde{\mathbf{u}}^{m+1} = (u_1^{m+1}, \cdots, u_{N-1}^{m+1})^T, \quad \tilde{\mathbf{u}}^m = (u_1^m, \cdots, u_{N-1}^m)^T. \end{aligned}$$

Then (15.2.3) can be written in a matrix form as

$$Q_1\tilde{\mathbf{u}}^{m+1} \quad = \quad Q\tilde{\mathbf{u}}^m + \frac{\lambda}{2}\begin{pmatrix} u_0^m + u_0^{m+1} \\ 0 \\ \vdots \\ 0 \\ u_N^m + u_N^{m+1} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\lambda}{2}u_0^m + (1-\lambda)u_1^m + \frac{\lambda}{2}u_2^m + \frac{\lambda}{2}u_0^{m+1} \\ \frac{\lambda}{2}u_1^m + (1-\lambda)u_2^m + \frac{\lambda}{2}u_3^m \\ \vdots \\ \frac{\lambda}{2}u_{N-3}^m + (1-\lambda)u_{N-2}^m + \frac{\lambda}{2}u_{N-1}^m \\ \frac{\lambda}{2}u_{N-2}^m + (1-\lambda)u_{N-1}^m + \frac{\lambda}{2}u_N^m + \frac{\lambda}{2}u_N^{m+1} \end{pmatrix}$$

To sum up, we have to get $\mathbf{u}^{m+1}$ from $\mathbf{u}^m$. The elements u_0^{m+1} and u_N^{m+1} in $\mathbf{u}^{m+1}$ can be obtained from the boundary condition. Hence we only have to get $\mathbf{v}^{m+1} = (u_1^{m+1}, \cdots, u_{N-1}^{m+1})^T$.

✎ **Initial Conditions**

A call with exercise price K, for example, satisfies

$$f(S,T) = (S-K)^+ = K(e^x - 1)^+.$$

We see from (15.2.2) that

$$f(S,T) = K\exp\left\{-\frac{1}{2}(k_2-1)x\right\} v(x,0)$$

and thus

$$\begin{aligned} v(x,0) &= \exp\left\{\frac{1}{2}(k_2-1)x\right\}(e^x-1)^+ \\ &= \left\{\exp\left(\frac{1}{2}(k_2+1)x\right) - \exp\left(\frac{1}{2}(k_2-1)x\right)\right\}^+. \end{aligned}$$

Analogously for the put we have

$$v(x,0) = \left\{\exp\left(\frac{1}{2}(k_2-1)x\right) - \exp\left(\frac{1}{2}(k_2+1)x\right)\right\}^+.$$

✎ **The Boundary Conditions**

When we consider a finite difference solution to the simple call option in the Black-Scholes model, usually we would refine the range of the underlying, S, i.e. we only want to get the solution in the region

$$S_{\min} \le S \le S_{\max}.$$

At the $S_{\min}$, we can simply set the option value equal to zero. If $S_{\min}$ is small enough, the error is negligible. At the $S_{\max}$, we could use

$$f(S_{\max}, t) = S_{\max}e^{-q(T-t)} - Ke^{-r(T-t)}.$$

Example 15.2.1 (Boundary Conditions for Vanilla Options). Assume dividend rate $q = 0$.

$$\begin{aligned} \text{Call}(S_t,t) &= S_t - Ke^{-r(T-t)} && \text{for } S_t \to \infty \\ \text{Put}(S_t,t) &= Ke^{-r(T-t)} - S_t && \text{for } S \approx 0. \end{aligned}$$

For a heat equation the boundary conditions are given by

$$v(x,\tau) = z_1(x,\tau) \text{ for } x \to -\infty, \quad v(x,\tau) = z_2(x,\tau) \text{ for } x \to \infty,$$

$$
\begin{aligned}
\text{call: } z_1(x,\tau) &= 0, \quad z_2(x,\tau) = \exp\left\{\frac{1}{2}(k_2+1)x+\frac{1}{4}(k_2+1)^2\tau\right\} - \exp\left\{\frac{1}{2}(k_2-1)x+\frac{1}{4}(k_2-1)^2\tau\right\},\\
\text{put: } z_1(x,\tau) &= \exp\left\{\frac{1}{2}(k_2-1)x+\frac{1}{4}(k_2-1)^2\tau\right\} - \exp\left\{\frac{1}{2}(k_2+1)x+\frac{1}{4}(k_2+1)^2\tau\right\}, \quad z_2(x,\tau) = 0.
\end{aligned}
$$

□

As another approach, there is the linear boundary condition:

$$\frac{\partial^2 f}{\partial S^2}(S_{\max},t) = 0.$$

This linear boundary condition is given by

$$
\begin{aligned}
u_0^{m+1} &= 2u_1^{m+1} - u_2^{m+1},\\
u_N^{m+1} &= 2u_{N-1}^{m+1} - u_{N-2}^{m+1}.
\end{aligned}
$$

In the Crank-Nicolson scheme of the heat equation, the equations are given by

$$
\begin{pmatrix}
-\frac{\lambda}{2}u_0^{m+1} + (1+\lambda)u_1^{m+1} - \frac{\lambda}{2}u_2^{m+1}\\
-\frac{\lambda}{2}u_1^{m+1} + (1+\lambda)u_2^{m+1} - \frac{\lambda}{2}u_3^{m+1}\\
\vdots\\
-\frac{\lambda}{2}u_{N-3}^{m+1} + (1+\lambda)u_{N-2}^{m+1} - \frac{\lambda}{2}u_{N-1}^{m+1}\\
-\frac{\lambda}{2}u_{N-2}^{m+1} + (1+\lambda)u_{N-1}^{m+1} - \frac{\lambda}{2}u_N^{m+1}
\end{pmatrix}
=
\begin{pmatrix}
\frac{\lambda}{2}u_0^{m} + (1-\lambda)u_1^{m} + \frac{\lambda}{2}u_2^{m}\\
\frac{\lambda}{2}u_1^{m} + (1-\lambda)u_2^{m} + \frac{\lambda}{2}u_3^{m}\\
\vdots\\
\frac{\lambda}{2}u_{N-3}^{m} + (1-\lambda)u_{N-2}^{m} + \frac{\lambda}{2}u_{N-1}^{m}\\
\frac{\lambda}{2}u_{N-2}^{m} + (1-\lambda)u_{N-1}^{m} + \frac{\lambda}{2}u_N^{m}
\end{pmatrix},
$$

$$
\begin{pmatrix}
-\frac{\lambda}{2}(2u_1^{m+1} - u_2^{m+1}) + (1+\lambda)u_1^{m+1} - \frac{\lambda}{2}u_2^{m+1}\\
-\frac{\lambda}{2}u_1^{m+1} + (1+\lambda)u_2^{m+1} - \frac{\lambda}{2}u_3^{m+1}\\
\vdots\\
-\frac{\lambda}{2}u_{N-3}^{m+1} + (1+\lambda)u_{N-2}^{m+1} - \frac{\lambda}{2}u_{N-1}^{m+1}\\
-\frac{\lambda}{2}u_{N-2}^{m+1} + (1+\lambda)u_{N-1}^{m+1} - \frac{\lambda}{2}(2u_{N-1}^{m+1} - u_{N-2}^{m+1})
\end{pmatrix}
=
\begin{pmatrix}
\frac{\lambda}{2}u_0^{m} + (1-\lambda)u_1^{m} + \frac{\lambda}{2}u_2^{m}\\
\frac{\lambda}{2}u_1^{m} + (1-\lambda)u_2^{m} + \frac{\lambda}{2}u_3^{m}\\
\vdots\\
\frac{\lambda}{2}u_{N-3}^{m} + (1-\lambda)u_{N-2}^{m} + \frac{\lambda}{2}u_{N-1}^{m}\\
\frac{\lambda}{2}u_{N-2}^{m} + (1-\lambda)u_{N-1}^{m} + \frac{\lambda}{2}u_N^{m}
\end{pmatrix},
$$

$$
\begin{pmatrix}
u_1^{m+1}\\
-\frac{\lambda}{2}u_1^{m+1} + (1+\lambda)u_2^{m+1} - \frac{\lambda}{2}u_3^{m+1}\\
\vdots\\
-\frac{\lambda}{2}u_{N-3}^{m+1} + (1+\lambda)u_{N-2}^{m+1} - \frac{\lambda}{2}u_{N-1}^{m+1}\\
u_{N-1}^{m+1}
\end{pmatrix}
=
\begin{pmatrix}
\frac{\lambda}{2}u_0^{m} + (1-\lambda)u_1^{m} + \frac{\lambda}{2}u_2^{m}\\
\frac{\lambda}{2}u_1^{m} + (1-\lambda)u_2^{m} + \frac{\lambda}{2}u_3^{m}\\
\vdots\\
\frac{\lambda}{2}u_{N-3}^{m} + (1-\lambda)u_{N-2}^{m} + \frac{\lambda}{2}u_{N-1}^{m}\\
\frac{\lambda}{2}u_{N-2}^{m} + (1-\lambda)u_{N-1}^{m} + \frac{\lambda}{2}u_N^{m}
\end{pmatrix}.
$$

This system of equations can be written as the following matrix form:

$$
\begin{bmatrix}
1 & 0 & 0 & \cdots & 0 & 0 & 0\\
-\frac{\lambda}{2} & 1+\lambda & -\frac{\lambda}{2} & \cdots & 0 & 0 & 0\\
0 & -\frac{\lambda}{2} & 1+\lambda & \cdots & 0 & 0 & 0\\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots\\
0 & 0 & 0 & \cdots & -\frac{\lambda}{2} & 1+\lambda & -\frac{\lambda}{2}\\
0 & 0 & 0 & \cdots & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
u_1^{m+1}\\ u_2^{m+1}\\ u_3^{m+1}\\ \vdots\\ u_{N-2}^{m+1}\\ u_{N-1}^{m+1}
\end{bmatrix}
$$

$$
=
\begin{bmatrix}
1-\lambda & \frac{\lambda}{2} & 0 & \cdots & 0 & 0 & 0\\
\frac{\lambda}{2} & 1-\lambda & \frac{\lambda}{2} & \cdots & 0 & 0 & 0\\
0 & \frac{\lambda}{2} & 1-\lambda & \cdots & 0 & 0 & 0\\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots\\
0 & 0 & 0 & \cdots & \frac{\lambda}{2} & 1-\lambda & \frac{\lambda}{2}\\
0 & 0 & 0 & 0 & 0 & \frac{\lambda}{2} & 1-\lambda
\end{bmatrix}
\begin{bmatrix}
u_1^{m}\\ u_2^{m}\\ u_3^{m}\\ \vdots\\ u_{N-2}^{m}\\ u_{N-1}^{m}
\end{bmatrix}
+
\begin{bmatrix}
\frac{\lambda}{2}u_0^{m}\\ 0\\ 0\\ \vdots\\ 0\\ \frac{\lambda}{2}u_N^{m}
\end{bmatrix}
$$

15.2.2 Implementation(θ-Method)

There is a generalization of the Crank-Nicolson method, usually called θ-method. It takes the form

$$\frac{v_n^{m+1}-v_n^m}{\Delta\tau}+O(\Delta\tau) = (1-\theta)\frac{v_{n+1}^m-2v_n^m+v_{n-1}^m}{\Delta x^2}+\theta\frac{v_{n+1}^{m+1}-2v_n^{m+1}+v_{n-1}^{m+1}}{\Delta x^2}+O(\Delta x^2).$$

Ignoring terms of $O(\Delta\tau)$ and $O(\Delta x^2)$, we can approximate this by

$$\begin{aligned}(1-\theta\lambda\delta^2)u_n^{m+1} &= \big(1+(1-\theta)\lambda\delta^2\big)u_n^m,\\ \text{where}\quad \lambda &= \frac{\Delta\tau}{\Delta x^2},\quad \delta^2 u_n = u_{n+1}-2u_n+u_{n-1}.\end{aligned}$$

The cases $\theta=0$, $\theta=\frac{1}{2}$ and $\theta=1$ give, respectively, the explicit, the Crank-Nicolson and the implicit finite difference approximation.

Let

$$Q_1 = \begin{bmatrix} 1+2\theta\lambda & -\theta\lambda & 0 & \cdots & 0\\ -\theta\lambda & 1+2\theta\lambda & -\theta\lambda & \cdots & 0\\ 0 & -\theta\lambda & 1+2\theta\lambda & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & -\theta\lambda\\ 0 & 0 & 0 & -\theta\lambda & 1+2\theta\lambda \end{bmatrix},$$

$$Q = \begin{bmatrix} 1-2(1-\theta)\lambda & (1-\theta)\lambda & 0 & \cdots & 0\\ (1-\theta)\lambda & 1-2(1-\theta)\lambda & (1-\theta)\lambda & \cdots & 0\\ 0 & (1-\theta)\lambda & 1-2(1-\theta)\lambda & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & (1-\theta)\lambda\\ 0 & 0 & 0 & (1-\theta)\lambda & 1-2(1-\theta)\lambda \end{bmatrix}$$

The the θ-method scheme can be expressed in a matrix form:

$$\begin{aligned} Q_1\tilde{\mathbf{u}}^{m+1} &= Q\tilde{\mathbf{u}}^m + \begin{pmatrix} (1-\theta)\lambda u_0^m+\theta\lambda u_0^{m+1}\\ 0\\ \vdots\\ 0\\ (1-\theta)\lambda u_N^m+\theta\lambda u_N^{m+1} \end{pmatrix}\\ &= \begin{pmatrix} (1-\theta)\lambda u_0^m+\big(1-2(1-\theta)\lambda\big)u_1^m+(1-\theta)\lambda u_2^m+\theta\lambda u_0^{m+1}\\ (1-\theta)\lambda u_1^m+\big(1-2(1-\theta)\lambda\big)u_2^m+(1-\theta)\lambda u_3^m\\ \vdots\\ (1-\theta)\lambda u_{N-3}^m+\big(1-2(1-\theta)\lambda\big)u_{N-2}^m+(1-\theta)\lambda u_{N-1}^m\\ (1-\theta)\lambda u_{N-2}^m+\big(1-2(1-\theta)\lambda\big)u_{N-1}^m+(1-\theta)\lambda u_N^m+\theta\lambda u_N^{m+1} \end{pmatrix}\end{aligned}$$

When the linear boundary condition is given, i,e

$$\begin{aligned} u_0^{m+1} &= 2u_1^{m+1}-u_2^{m+1},\\ u_N^{m+1} &= 2u_{N-1}^{m+1}-u_{N-2}^{m+1}, \end{aligned}$$

Q_1 is given by

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\theta\lambda & 1+2\theta\lambda & -\theta\lambda & \cdots & 0 \\ 0 & -\theta\lambda & 1+2\theta\lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & -\theta\lambda \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

15.3 Alternative Approach to Explicit FDM

In this section, we will study the explicit difference scheme without any transformation.

1. To quote Wilmott(2000):

   > The main reason that I rarely do any transforming of the equation when I am solving it numerically is that I like to solve in terms of the real financial variables since terms of the contract are specified using real variables: transforming to the heat equation could cause problems for contracts such as barrier options.

2. In practice, the upper bound for underlying asset does not have to be too large. Typically it should be three of four times the value of the exercise price or, more generally, three or four times the values of the asset at which there is some important behavior.

   For an up-and-out barrier option there is no need to make the grid extend beyond the barrier.

3. Approximation Δ:

   $$\frac{f_{k,i+1} - f_{k,i}}{\Delta S}, \quad \frac{f_{k,i} - f_{k,i-1}}{\Delta S}, \quad \frac{f_{k,i+1} - f_{k,i-1}}{2\Delta S},$$

   where $S = S_0 + i\Delta S$ and $t = T - k\Delta t$.

   One side approximation for Δ with error of $O(\Delta S^2)$ is given by

   $$\frac{\partial f}{\partial S} = \frac{-3f_{k,i} + 4f_{k,i+1} - f_{k,i+2}}{2\Delta S} + O(\Delta S^2).$$

4. Boundary condition: On $S = 0$, the Black-Scholes PDE is

   $$\frac{\partial f}{\partial t}(t, 0) - rf(0, t) = 0, \quad \text{or } f_{k,0} = (1 - r\Delta t)f_{k-1,0}.$$

   When the option has a payoff that at most linear in the underlying for large values of S then you can use the upper boundary condition

   $$\frac{\partial^2 f}{\partial S^2}(t, S) \to 0, \quad \text{as } S \to \infty.$$

   The finite-difference representation is

   $$\begin{aligned} f_{k,N} &= 2f_{k,N-1} - f_{k,N-2}, \\ f_{k,0} &= 2f_{k,1} - f_{k,2}. \end{aligned}$$

5. The explicit FDM: The Black-Scholes equation is

   $$\frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} + (r - \delta)S\frac{\partial f}{\partial S} - rf = 0.$$

   We going to write this as

   $$\frac{\partial f}{\partial t} + a(S, t)\frac{\partial^2 f}{\partial S^2} + b(S, t)\frac{\partial f}{\partial S} + c(S, t)f = 0.$$

The explicit difference scheme is given by

$$\frac{V_{k,i}-V_{k+1,i}}{\Delta t}+a_{k,i}\left(\frac{V_{k,i+1}-2V_{k,i}+V_{k,i-1}}{\Delta S^2}\right)+b_{k,i}\left(\frac{V_{k,i+1}-V_{k,i-1}}{2\Delta S}\right)+c_{k,i}V_{k,i}=O(\Delta t,\Delta S^2)$$

where $V_{k,i}=f(T-k\Delta t,S_0+i\Delta S)$.

This is equivalent to the followings:

$$\begin{aligned}
\Delta_i &= \frac{V_{k,i+1}-V_{k,i-1}}{2\Delta S},\\
\Gamma_i &= \frac{V_{k,i+1}-2V_{k,i}+V_{k,i-1}}{\Delta S^2},\\
V_{k+1,i} &= V_{k,i}+\Delta t\left(\frac{1}{2}\sigma^2 S_i^2\Gamma_i+(r-\delta)S_i\Delta_i-rV_{k,i}\right)\\
&= A_{k,i}V_{k,i-1}+(1+B_{k,i})V_{k,i}+C_{k,i}V_{k,i+1}, \qquad (15.3.1)
\end{aligned}$$

where

$$\begin{aligned}
A_{k,i} &= \nu_1 a_{k,i}-\frac{1}{2}\nu_2 b_{k,i},\\
B_{k,i} &= -2\nu_1 a_{k,i}+\Delta t c_{k,i},\\
C_{k,i} &= \nu_1 a_{k,i}+\frac{1}{2}\nu_2 b_{k,i},
\end{aligned}$$

$\nu_1=\frac{\Delta t}{\Delta S^2}$ and $\nu_2=\frac{\Delta t}{\Delta S}$.

Note that we use $rV_{k+1,i}$ for $rV_{k,i}$ in Eq(15.1.4).

6. Conditions for stability:

   (a) The first condition is

   $$c_{k,i}<0.$$

   In financial problems, we always have a negative c, often it is simply $-r$ where r is interest rate.

   (b) The second condition is

   $$\Delta t\le\frac{\Delta S^2}{2a}=\frac{\Delta S^2}{\sigma^2 S^2}=\frac{1}{\sigma^2}\left(\frac{\Delta S}{S}\right)^2$$

   If

   $$\Delta t\le\frac{1}{\sigma^2}\left(\frac{\Delta S}{S_{\max}}\right)^2,$$

   the second condition is satisfied.

   (c) The third condition is

   $$\Delta S\le\frac{2a}{|b|}=\frac{\sigma^2 S}{|r-\delta|}.$$

   If

   $$\Delta S\le\frac{\sigma^2}{|r-\delta|}\,S_{\min}$$

   the third condition is satisfied.

Theorem 15.3.1. The explicit finite difference scheme (15.3.1) is convergent(stable) if

$$c_{k,i} \le 0, \quad 2\nu_1\, a_{k,i} - \Delta t c_{k,i} \le 1, \quad \text{and} \quad \frac{1}{2}\nu_2 |b_{k,i}| \le \nu_1 a_{k,i}.$$

Proof. (This proof is not rigorous.) Taking the discrete Fourier transformation for (15.3.1) , we get

$$\begin{aligned}
\widehat{V}_{k+1}(\xi) &= Ae^{-i\xi}\widehat{V}_k + (1+B)\widehat{V}_k + Ce^{i\xi}\widehat{V}_k \\
&= \big[(A+C)\cos\xi + (1+B) + i(C-A)\sin\xi\big]\widehat{V}_k \\
&= \big[2\nu_1 a\cos\xi + 1 - 2\nu_1 a + \Delta t c + i\nu_2 b\sin\xi\big]\widehat{V}_k \\
&= \big[2\nu_1 a(\cos\xi - 1) + 1 + \Delta t c + i\nu_2 b\sin\xi\big]\widehat{V}_k \\
&= \big[X(\cos\xi - 1) + 1 + Y + iZ\sin\xi\big]\widehat{V}_k,
\end{aligned}$$

where $X = 2\nu_1 a$, $Y = \Delta t c$ and $Z = \nu_2 b$. If

$$\epsilon := \big|X(\cos\xi - 1) + 1 + Y + iZ\sin\xi\big| \le 1$$

the scheme (15.3.1) is convergent.

$$\begin{aligned}
\epsilon^2 &= (1+Y-X+X\cos\xi)^2 + Z^2\cos^2\xi \\
&= (1+Y-X)^2 + 2X(1+Y-X)\cos\xi + X^2\cos^2\xi + Z^2\sin^2\xi.
\end{aligned}$$

Since $|Z| \le X$,

$$\epsilon^2 \le (1+Y-X)^2 + 2X(1+Y-X)\cos\xi + X^2.$$

Since $(1+Y-X) \ge 0$ if

$$\frac{1 - X^2 - (1+Y-X)^2}{2X(1+Y-X)} \ge \cos\xi,$$

then $|\epsilon| \le 1$. Thus

$$\frac{1 - X^2 - (1+Y-X)^2}{2X(1+Y-X)} \ge 1$$

is a sufficient condition for stability. Also

$$1 \ge (X + 1 + Y - X)^2 = (1+Y)^2 \quad \Big(\iff \quad Y \le 0\Big)$$

is a sufficient condition for stability[2]. ■

[2]In fact, to get this result, we have to assume that all coefficients are slowly varying over the ΔS length-scales.

15.4 Fully Implicit Scheme

In this section we follow Wilmott(2000). The fully implicit method is given as follows:

$$\frac{V_{k,i}-V_{k+1,i}}{\Delta t}+a_{k+1,i}\left(\frac{V_{k+1,i+1}-2V_{k+1,i}+V_{k+1,i-1}}{\Delta S^2}\right)+b_{k+1,i}\left(\frac{V_{k+1,i+1}-V_{k+1,i-1}}{2\Delta S}\right)+c_{k+1,i}V_{k+1,i}=0.$$

This can be written as

$$\begin{aligned}
\Delta_i &= \frac{V_{k+1,i+1}-V_{k+1,i-1}}{2\Delta S},\\
\Gamma_i &= \frac{V_{k+1,i+1}-2V_{k+1,i}+V_{k+1,i-1}}{\Delta S^2},\\
V_{k+1,i} &- \Delta t\left(\frac{1}{2}\sigma^2 S_i^2\Gamma_i+(r-\delta)S_i\Delta_i-rV_{k+1,i}\right)=V_{k,i}
\end{aligned}$$

or

$$A_{k+1,i}V_{k+1,i-1}+(1+B_{k+1,i})V_{k+1,i}+C_{k+1,i}V_{k+1,i+1}=V_{k,i}, \tag{15.4.1}$$

where

$$\begin{aligned}
A_{k+1,i} &= -\nu_1 a_{k,i}+\frac{1}{2}\nu_2 b_{k,i},\\
B_{k+1,i} &= 2\nu_1 a_{k,i}-\Delta t c_{k,i},\\
C_{k+1,i} &= -\nu_1 a_{k,i}-\frac{1}{2}\nu_2 b_{k,i},
\end{aligned}$$

$\nu_1=\frac{\Delta t}{\Delta S^2}$ and $\nu_2=\frac{\Delta t}{\Delta S}$.

Theorem 15.4.1. The implicit difference scheme (15.4.1) is unconditionally stable.

Proof. The discrete Fourier transformation is given as

$$\left[Ae^{-i\xi}+(1+B)+Ce^{i\xi}\right]\widehat{V}_{k+1} \quad = \quad \widehat{V}_k.$$

The left hand side can be written as

$$\begin{aligned}
\text{LHS} &= \left[(A+C)\cos\xi+(1+B)+i(C-A)\sin\xi\right]\widehat{V}_{k+1}\\
&= \left[-2\nu_1 a\cos\xi+1+2\nu_1 a-\Delta tc+i\nu_2 b\sin\xi\right]\widehat{V}_{k+1}\\
&= \left[2\nu_1 a(1-\cos\xi)+1-\Delta tc+i\nu_2 b\sin\xi\right]\widehat{V}_{k+1}\\
&= \left[X(1-\cos\xi)+1-Y+iZ\sin\xi\right]\widehat{V}_{k+1},
\end{aligned}$$

where $X=2\nu_1 a$, $Y=\Delta tc$ and $Z=\nu_2 b$. The norm of RHS is

$$\begin{aligned}
&\Big(X(1-\cos\xi)+1-Y\Big)^2+Z^2\sin^2\xi\\
=\ &\Big(2X\sin^2\frac{\xi}{2}+1-Y\Big)^2+Z^2\sin^2\xi\\
\geq\ &(1-Y)^2 \quad \geq \quad 1
\end{aligned}$$

since $Y\leq 0$. ■

15.5 Alternative Approach to Crank-Nicolson

In this section we follow Wilmott(2000). The Black-Scholes equation is

$$\frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} + (r-\delta) S \frac{\partial f}{\partial S} - rf = 0. \tag{15.5.1}$$

Substituting $x = \log S$ gives us

$$\frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2} + \left(r - \delta - \frac{1}{2}\sigma^2\right) \frac{\partial f}{\partial x} - rf = 0. \tag{15.5.2}$$

We are going to write Eq(15.5.1) and Eq(15.5.2) as

$$\frac{\partial f}{\partial t} + a(x,t)\frac{\partial^2 f}{\partial S^2} + b(x,t)\frac{\partial f}{\partial x} + c(x,t) f = 0. \tag{15.5.3}$$

The Crank-Nicolson method can be thought of as an average of the explicit method and the implicit method. The Crank- Nicolson scheme is given by as follows:

$$\begin{aligned}
&\frac{V_{k,i} - V_{k+1,i}}{\Delta t} + \frac{a_{k+1,i}}{2}\left(\frac{V_{k+1,i+1} - 2V_{k+1,i} + V_{k+1,i-1}}{\Delta S^2}\right) + \frac{a_{k,i}}{2}\left(\frac{V_{k,i+1} - 2V_{k,i} + V_{k,i-1}}{\Delta S^2}\right) \\
&+ \frac{b_{k+1,i}}{2}\left(\frac{V_{k+1,i+1} - V_{k+1,i-1}}{2\Delta S}\right) + \frac{b_{k,i}}{2}\left(\frac{V_{k,i+1} - V_{k,i-1}}{2\Delta S}\right) + \frac{c_{k+1,i}}{2} V_{k+1,i} + \frac{c_{k,i}}{2} V_{k,i} = O(\Delta t^2, \Delta S^2),
\end{aligned}$$

where $V_{k,i} = f(T - k\Delta t, S_0 + i\Delta x)$. This can be written as

$$\begin{aligned}
& -A_{k+1,i}V_{k+1,i-1} + (1 - B_{k+1,i})V_{k+1,i} - C_{k+1,i}V_{k+1,i+1} \\
= \quad & A_{k,i}V_{k,i-1} + (1 + B_{k,i})V_{k,i} + C_{k,i}V_{k,i+1},
\end{aligned} \tag{15.5.4}$$

where

$$\begin{aligned}
A_{k,i} &= \frac{1}{2}\nu_1 a_{k,i} - \frac{1}{4}\nu_2 b_{k,i}, \\
B_{k,i} &= -\nu_1 a_{k,i} + \frac{1}{2}\Delta t c_{k,i}, \\
C_{k,i} &= \frac{1}{2}\nu_1 a_{k,i} + \frac{1}{4}\nu_2 b_{k,i},
\end{aligned}$$

$\nu_1 = \frac{\Delta t}{\Delta x^2}$ and $\nu_2 = \frac{\Delta t}{\Delta x}$.

Example 15.5.1. If the parameter are given as

$$\begin{aligned}
a &= \frac{1}{2}\sigma^2, \\
b &= r - \delta - \frac{1}{2}\sigma^2, \\
c &= -r.
\end{aligned}$$

then we have

$$A = \frac{1}{4}\sigma^2 \frac{\Delta t}{\Delta x^2} - \frac{\Delta t}{4\Delta x}\left(r - \delta - \frac{1}{2}\sigma^2\right),$$

$$
\begin{aligned}
B &= -\frac{1}{2}\sigma^2 \frac{\Delta t}{\Delta x^2} - \frac{1}{2} r\,\Delta t, \\
C &= \frac{1}{4}\sigma^2 \frac{\Delta t}{\Delta x^2} + \frac{\Delta t}{4\Delta x}\left(r - \delta - \frac{1}{2}\sigma^2\right).
\end{aligned}
$$

□

The following theorem (15.5.2) gives the stability condition.

Theorem 15.5.2. If $a_{k,i}, b_{k,i}$ and $c_{k,i}$ are independent of time parameter k and $c_{k,i} \le 0$, the Crank-Nicolson scheme is unconditionally stable.

Proof. After taking the discrete Fourier transform, we have

$$
\begin{aligned}
\Big(-A_{k+1}e^{-i\xi} + (1 - B_{k+1}) - C_{k+1}e^{i\xi}\Big)\widehat{V}_{k+1} &= \Big(A_k e^{-i\xi} + (1 + B_k) + C_k e^{i\xi}\Big)\widehat{V}_k, \\
\Big((-A - C)\cos\xi + (1 - B) + i(A - C)\sin\xi\Big)\widehat{V}_{k+1} &= \Big((A + C)\cos\xi + (1 + B) + i(-A + C)\sin\xi\Big)\widehat{V}_k, \\
\Big(-\nu_1 a\cos\xi + (1 + \nu_1 a - \frac{1}{2}\Delta t c) + i(-\frac{1}{2}\nu_2 b)\sin\xi\Big)\widehat{V}_{k+1} &= \Big(\nu_1 a\cos\xi + (1 - \nu_1 a + \frac{1}{2}\Delta t c) + i(\frac{1}{2}\nu_2 b)\sin\xi\Big)\widehat{V}_k, \\
\Big(\nu_1 a(1 - \cos\xi) + 1 - \frac{1}{2}\Delta t c + i(-\frac{1}{2}\nu_2 b)\sin\xi\Big)\widehat{V}_{k+1} &= \Big(-\nu_1 a(1 - \cos\xi) + 1 + \frac{1}{2}\Delta t c + i(\frac{1}{2}\nu_2 b)\sin\xi\Big)\widehat{V}_k \\
\Big(1 + X - Y - X\cos\xi - iZ\sin\xi\Big)\widehat{V}_{k+1} &= \Big(1 - X + Y + X\cos\xi + iZ\sin\xi\Big)\widehat{V}_k,
\end{aligned}
$$

where $X = 2\nu_1 a$, $Y = \frac{1}{2}\Delta t c$ and $Z = \frac{1}{2}\nu_2 b$. Since

$$
\begin{aligned}
& (1 - X + Y)^2 + 2X\cos\xi(1 - X + Y) + Z^2\sin^2\xi - \Big((1 + X - Y)^2 - 2X\cos\xi(1 + X - Y) + Z^2\sin^2\xi\Big) \\
=\ & 4(Y - X + X\cos\xi) \\
=\ & 4\left(Y - 2X\cos^2\frac{\xi}{2}\right) \le 0,
\end{aligned}
$$

we complete the proof[3]. ■

The Crank-Nicolson method can be written in the matrix form

$$
\begin{bmatrix}
-A_{k+1,1} & 1 - B_{k+1,1} & -C_{k+1,1} & \cdots & 0 & 0 & 0 \\
0 & -A_{k+1,2} & 1 - B_{k+1,2} & \cdots & 0 & 0 & 0 \\
0 & 0 & -A_{k+1,3} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & -A_{k+1,I-1} & 1 - B_{k+1,I-1} & -C_{k+1,I-1}
\end{bmatrix}
\begin{bmatrix}
V_{k+1,0} \\ V_{k+1,1} \\ \vdots \\ \vdots \\ V_{k+1,I-1} \\ V_{k+1,I}
\end{bmatrix}
$$

$$
= \begin{bmatrix}
A_{k,1} & 1 + B_{k,1} & C_{k,1} & \cdots & 0 & 0 & 0 \\
0 & A_{k,2} & 1 + B_{k,2} & \cdots & 0 & 0 & 0 \\
0 & 0 & A_{k,3} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & A_{k,I-1} & 1 + B_{k,I-1} & C_{k,I-1}
\end{bmatrix}
\begin{bmatrix}
V_{k,0} \\ V_{k,1} \\ \vdots \\ \vdots \\ V_{k,I-1} \\ V_{k,I}
\end{bmatrix}.
$$

[3] This proof is wrong. Why?

The two matrices have $I-1$ rows and $I+1$ columns. These are not a square matrices. So we need two boundary conditions. This matrix form can be written as

$$\begin{bmatrix} 1-B_{k+1,1} & -C_{k+1,1} & \cdots & 0 & 0 \\ -A_{k+1,2} & 1-B_{k+1,2} & \cdots & 0 & 0 \\ 0 & -A_{k+1,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -A_{k+1,I-1} & 1-B_{k+1,I-1} \end{bmatrix} \begin{bmatrix} V_{k+1,1} \\ \vdots \\ \vdots \\ V_{k+1,I-1} \end{bmatrix} + \begin{bmatrix} -A_{k+1,1}V_{k+1,0} \\ 0 \\ \vdots \\ \vdots \\ 0 \\ -C_{k+1,I-1}V_{k+1,I} \end{bmatrix}$$

$$= \begin{bmatrix} A_{k,1} & 1+B_{k,1} & C_{k,1} & \cdots & 0 & 0 & 0 \\ 0 & A_{k,2} & 1+B_{k,2} & \cdots & 0 & 0 & 0 \\ 0 & 0 & A_{k,3} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_{k,I-1} & 1+B_{k,I-1} & C_{k,I-1} \end{bmatrix} \begin{bmatrix} V_{k,0} \\ V_{k,1} \\ \vdots \\ \vdots \\ V_{k,I-1} \\ V_{k,I} \end{bmatrix}. \tag{15.5.5}$$

Now the first matrix is $(I-1)\times(I-1)$. If we know $V_{k+1,0}$ and $V_{k+1,I}$, we can solve this system of equations.

There are many forms of the boundary condition that is to be implemented.

1. $V_{k+1,0}$ and $V_{k+1,I}$ are given.
2. The relationships between $V_{k+1,0}$ and $V_{k+1,1}$ and between $V_{k+1,I}$ and $V_{k+1,I-1}$ are given. For example, if the value f at some S between $i=0$ and $i=1$ is given, we can extrapolate $V_{k+1,0}$ from $V_{k+1,0}$ and f (see Figure 15.3). i.e.

$$V_{k+1,0} = \frac{1}{\alpha}\left(f-(1-\alpha)V_{k+1,1}\right)$$

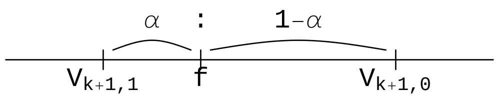


Figure 15.3: Extrapolation from $V_{k+1,1}$ and f

More generally, suppose we have

$$\begin{aligned} V_{k+1,0} &= a+bV_{k+1,1}, \\ V_{k+1,I} &= c+dV_{k+1,I-1}. \end{aligned}$$

The left-hand side of Eq(15.5.5) can be written

$$\begin{bmatrix} 1-B_{k+1,1} & -C_{k+1,1} & \cdots & 0 & 0 \\ -A_{k+1,2} & 1-B_{k+1,2} & \cdots & 0 & 0 \\ 0 & -A_{k+1,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -A_{k+1,I-1} & 1-B_{k+1,I-1} \end{bmatrix} \begin{bmatrix} V_{k+1,1} \\ \vdots \\ \vdots \\ V_{k+1,I-1} \end{bmatrix} + \begin{bmatrix} -A_{k+1,1}(a+bV_{k+1,1}) \\ 0 \\ \vdots \\ \vdots \\ 0 \\ -C_{k+1,I-1}(c+dV_{k+1,I-1}) \end{bmatrix}.$$

By absorbing the $V_{k+1,1}$ and $V_{k+1,I-1}$ terms into the matrix, we can write this as

$$\begin{bmatrix} 1-B_{k+1,1}-bA_{k+1,1} & -C_{k+1,1} & \cdots & 0 & 0 \\ -A_{k+1,2} & 1-B_{k+1,2} & \cdots & 0 & 0 \\ 0 & -A_{k+1,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -A_{k+1,I-1} & 1-B_{k+1,I-1}-dC_{k+1,I-1} \end{bmatrix} \begin{bmatrix} V_{k+1,1} \\ \vdots \\ \vdots \\ V_{k+1,I-1} \end{bmatrix} + \begin{bmatrix} -aA_{k+1,1} \\ 0 \\ \vdots \\ \vdots \\ 0 \\ -cC_{k+1,I-1} \end{bmatrix}.$$

3. $V_{k+1,0} = 2V_{k+1,1} - V_{k+1,2}$ and $V_{k+1,I} = 2V_{k+1,I-1} - V_{k+1,I-2}$.
   In this case, the left-hand side of Eq(15.5.5) can be written

$$\begin{bmatrix} 1-B_{k+1,1}-2A_{k+1,1} & -C_{k+1,1}+A_{k+1,1} & \cdots & 0 & 0 \\ -A_{k+1,2} & 1-B_{k+1,2} & \cdots & 0 & 0 \\ 0 & -A_{k+1,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -A_{k+1,I-1}+C_{k+1,I-1} & 1-B_{k+1,I-1}-2C_{k+1,I-1} \end{bmatrix} \begin{bmatrix} V_{k+1,1} \\ \vdots \\ \vdots \\ V_{k+1,I-1} \end{bmatrix}.$$

15.6 θ-Method

The θ-method gives the explicit, implicit and Crank-Nicolson when $\theta = 0$, $\theta = 1$ and $\theta = \frac{1}{2}$, respectively. For the PDE(15.5.3), the θ-method is given by

$$\begin{aligned} &\frac{V_i^k - V_i^{k+1}}{\Delta t} + (1-\theta)a_i^k \frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{\Delta x^2} + \theta a_i^{k+1} \frac{V_{i+1}^{k+1} - 2V_i^{k+1} + V_{i-1}^{k+1}}{\Delta x^2} \\ &+(1-\theta)b_i^k \frac{V_{i+1}^k - V_{i-1}^k}{2\Delta x} + \theta b_i^{k+1} \frac{V_{i+1}^{k+1} - V_{i-1}^{k+1}}{2\Delta x} + (1-\theta)c_i^k V_i^k + \theta c_i^{k+1} V_i^{k+1} \quad = \quad 0. \end{aligned}$$

This scheme can be rewritten as

$$\begin{aligned} & -\theta \left(\frac{\Delta t}{\Delta x^2} a_i^{k+1} - \frac{\Delta t}{2\Delta x} b_i^{k+1} \right) V_{i-1}^{k+1} + \left[1 - \theta \left(-\frac{2\Delta t}{\Delta x^2} a_i^{k+1} + \Delta t c_i^{k+1} \right) \right] V_i^{k+1} - \theta \left(\frac{\Delta t}{\Delta x^2} a_i^{k+1} + \frac{\Delta t}{2\Delta x} b_i^{k+1} \right) V_{i+1}^{k+1} \\ = \quad & (1-\theta) \left(\frac{\Delta t}{\Delta x^2} a_i^k - \frac{\Delta t}{2\Delta x} b_i^k \right) V_{i-1}^k + \left[1 + (1-\theta) \left(-\frac{2\Delta t}{\Delta x^2} a_i^k + \Delta t c_i^k \right) \right] V_i^k + (1-\theta) \left(\frac{\Delta t}{\Delta x^2} a_i^{k+1} + \frac{\Delta t}{2\Delta x} b_i^{k+1} \right) V_{i+1}^k \end{aligned}$$

and

$$-\theta A_i^{k+1} V_{i-1}^{k+1} + \left[1 - \theta B_i^{k+1}\right] V_i^{k+1} - \theta C_i^{k+1} V_{i+1}^{k+1} \quad = \quad (1-\theta) A_i^k V_{i-1}^k + \left[1 + (1-\theta) B_i^k\right] V_i^k + (1-\theta) C_i^k V_{i+1}^k,$$

where

$$\begin{aligned} A_i^k &= \frac{\Delta t}{\Delta x^2} a_i^k - \frac{\Delta t}{2\Delta x} b_i^k, \\ B_i^k &= -\frac{2\Delta t}{\Delta x^2} a_i^k + \Delta t c_i^k, \\ C_i^k &= \frac{\Delta t}{\Delta x^2} a_i^k + \frac{\Delta t}{2\Delta x} b_i^k, \quad i = 1, 2, \cdots, I-1. \end{aligned}$$

The matrix form of this scheme can be written as

$$\begin{bmatrix} 1-\theta B_1^{k+1} & -\theta C_1^{k+1} & \cdots & 0 & 0 \\ -\theta A_2^{k+1} & 1-\theta B_2^{k+1} & \cdots & 0 & 0 \\ 0 & -\theta A_3^{k+1} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -\theta A_{I-1}^{k+1} & 1-\theta B_{I-1}^{k+1} \end{bmatrix} \begin{bmatrix} V_1^{k+1} \\ \vdots \\ \vdots \\ V_{I-1}^{k+1} \end{bmatrix} + \begin{bmatrix} -\theta A_1^{k+1} V_0^{k+1} \\ 0 \\ \vdots \\ \vdots \\ 0 \\ -\theta C_{I-1}^{k+1} V_I^{k+1} \end{bmatrix} \tag{15.6.1}$$

$$= \begin{bmatrix} (1-\theta)A_1^k & 1+(1-\theta)B_1^k & (1-\theta)C_1^k & \cdots & 0 & 0 & 0 \\ 0 & (1-\theta)A_2^k & 1+(1-\theta)B_2^k & \cdots & 0 & 0 & 0 \\ 0 & 0 & (1-\theta)A_3^k & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & (1-\theta)A_{I-1}^k & 1+(1-\theta)B_{I-1}^k & (1-\theta)C_{I-1}^k \end{bmatrix} \begin{bmatrix} V_0^k \\ V_1^k \\ \vdots \\ \vdots \\ V_{I-1}^k \\ V_I^k \end{bmatrix}.$$

✎ **Boundary Conditions**

1. Simply Linear Boundary Conditions: When the linear boundary condition is given, i.e.

$$V_0^{k+1} = 2V_1^{k+1} - V_2^{k+1}, \quad V_I^{k+1} = 2V_{I-1}^{k+1} - V_{I-2}^{k+1}$$

then left hand side is

$$\begin{bmatrix} 1-\theta B_1^{k+1}-2\theta A_1^{k+1} & -\theta C_1^{k+1}+\theta A_1^{k+1} & \cdots & 0 & 0 \\ -\theta A_2^{k+1} & 1-\theta B_2^{k+1} & \cdots & 0 & 0 \\ 0 & -\theta A_3^{k+1} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -\theta A_{I-1}^{k+1}+\theta C_{I-1}^{k+1} & 1-\theta B_{I-1}^{k+1}-2\theta C_{I-1}^{k+1} \end{bmatrix} \begin{bmatrix} V_1^{k+1} \\ \vdots \\ \vdots \\ V_{I-1}^{k+1}. \end{bmatrix}.$$

The linear boundary conditions are imperfect. We place boundaries far enough from the region of interest, so that even if the imposed boundary conditions are imperfect, it does not materially affect the solution.

2. Modified Linear Boundary Conditions: Assume we have transformed $x = \log S$. The boundary condition

$$\frac{\partial^2 f}{\partial S^2} = 0$$

is equivalent to

$$\frac{\partial^2 f}{\partial x^2}\frac{1}{S^2} - \frac{\partial f}{\partial x}\frac{1}{S^2} = 0. \tag{15.6.2}$$

(15.6.2) can be approximated by

$$\frac{f(x+\Delta x)-2f(x)+f(x-\Delta x)}{\Delta x^2} = \frac{f(x+\Delta x)-f(x-\Delta x)}{2\Delta x}.$$

At $i = 0$ we have

$$\frac{V_2^{k+1}-2V_1^{k+1}+V_0^{k+1}}{\Delta x^2} = \frac{V_2^{k+1}-V_0^{k+1}}{2\Delta x}$$

and

$$V_0^{k+1} = \frac{4}{\Delta x+2}V_1^{k+1} + \frac{\Delta x-2}{\Delta x+2}V_2^{k+1}.$$

Similary, at $i = I$ we have

$$V_I^{k+1} = \frac{\Delta x+2}{\Delta x-2}V_{I-2}^{k+1} + \frac{-4}{\Delta x-2}V_{I-1}^{k+1}.$$

Thus (15.6.1) can be written as

$$\begin{bmatrix} 1-\theta B_1^{k+1}-\frac{4}{\Delta x+2}\theta A_1^{k+1} & -\theta C_1^{k+1}-\frac{\Delta x-2}{\Delta x+2}\theta A_1^{k+1} & \cdots & 0 & 0 \\ -\theta A_2^{k+1} & 1-\theta B_2^{k+1} & \cdots & 0 & 0 \\ 0 & -\theta A_3^{k+1} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -\theta A_{I-1}^{k+1}-\frac{\Delta x+2}{\Delta x-2}\theta C_{I-1}^{k+1} & 1-\theta B_{I-1}^{k+1}-\frac{-4}{\Delta x-2}\theta C_{I-1}^{k+1} \end{bmatrix} \begin{bmatrix} V_1^{k+1} \\ \vdots \\ \vdots \\ V_{I-1}^{k+1}. \end{bmatrix}.$$

3. Solving Alternative PDEs at Boundaries: We can use a simple PDE version by eliminating the diffusion part,

$$\frac{\partial f}{\partial t} + a(x,t)\frac{\partial^2 f}{\partial x^2} + b(x,t)\frac{\partial f}{\partial x} + c(x,t)f = 0,$$

$$\frac{\partial f}{\partial t} + \underbrace{\widetilde{b}(x,t)\frac{\partial f}{\partial x}}_{\text{convection term}} + c(x,t)f = 0.$$

For example, $b(x,t) = r - \delta - \frac{1}{2}\sigma^2, \widetilde{b}(x,t) = r - \delta$.

At $i = 0$ we use forward differences for convection term:

$$\frac{V_0^k - V_0^{k+1}}{\Delta t} + (1-\theta)\widetilde{b}_0^k \frac{V_1^k - V_0^k}{\Delta x} + \theta \widetilde{b}_0^{k+1}\frac{V_1^{k+1} - V_0^{k+1}}{\Delta x} + (1-\theta)c_0^k V_0^k + \theta c_0^{k+1}V_0^{k+1} = 0.$$

This equation can be written as

$$\left[1 - \theta\left(-\frac{\Delta t}{\Delta x}\widetilde{b}_0^{k+1} + \Delta t c_0^{k+1}\right)\right] V_0^{k+1} - \theta\frac{\Delta t}{\Delta x}\widetilde{b}_0^{k+1}V_1^{k+1} = \left[1 + (1-\theta)\left(-\frac{\Delta t}{\Delta x}\widetilde{b}_0^k + \Delta t c_0^k\right)\right] V_0^k + (1-\theta)\frac{\Delta t}{\Delta x}\widetilde{b}_0^k V_1^k.$$

At $i = I$ we use backward differences for convection term:

$$\frac{V_I^k - V_I^{k+1}}{\Delta t} + (1-\theta)\widetilde{b}_I^k \frac{V_I^k - V_{I-1}^k}{\Delta x} + \theta \widetilde{b}_I^{k+1}\frac{V_I^{k+1} - V_{I-1}^{k+1}}{\Delta x} + (1-\theta)c_I^k V_I^k + \theta c_I^{k+1}V_I^{k+1} = 0.$$

This equation can be written as

$$\theta\frac{\Delta t}{\Delta x}\widetilde{b}_I^{k+1}V_{I-1}^{k+1} + \left[1 - \theta\left(\frac{\Delta t}{\Delta x}\widetilde{b}_I^{k+1} + \Delta t c_I^{k+1}\right)\right] V_I^{k+1} = -(1-\theta)\frac{\Delta t}{\Delta x}\widetilde{b}_I^k V_{I-1}^k + \left[1 + (1-\theta)\left(\frac{\Delta t}{\Delta x}\widetilde{b}_I^k + \Delta t c_I^k\right)\right] V_I^k.$$

These are of the matrix form

$$\begin{bmatrix} 1-\theta B_0^{k+1} & -\theta C_0^{k+1} & 0 & \cdots & 0 & 0 & 0 \\ -\theta A_1^{k+1} & 1-\theta B_1^{k+1} & -\theta C_1^{k+1} & \cdots & 0 & 0 & 0 \\ 0 & -\theta A_2^{k+1} & 1-\theta B_2^{k+1} & \cdots & 0 & 0 & 0 \\ 0 & 0 & -\theta A_3^{k+1} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\theta A_{I-1}^{k+1} & 1-\theta B_{I-1}^{k+1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\theta A_I^{k+1} & 1-\theta B_I^{k+1} \end{bmatrix} \begin{bmatrix} V_0^{k+1} \\ V_1^{k+1} \\ \vdots \\ \vdots \\ V_{I-1}^{k+1} \\ V_I^{k+1} \end{bmatrix}$$

$$= \begin{bmatrix} 1+(1-\theta)B_0^k & (1-\theta)C_0^k & 0 & \cdots & 0 & 0 & 0 \\ (1-\theta)A_1^k & 1+(1-\theta)B_1^k & (1-\theta)C_1^k & \cdots & 0 & 0 & 0 \\ 0 & (1-\theta)A_2^k & 1+(1-\theta)B_2^k & \cdots & 0 & 0 & 0 \\ 0 & 0 & (1-\theta)A_3^k & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & (1-\theta)A_{I-1}^k & 1+(1-\theta)B_{I-1}^k & (1-\theta)C_{I-1}^k \\ 0 & 0 & 0 & \cdots & 0 & (1-\theta)A_I^k & 1+(1-\theta)B_I^k \end{bmatrix} \begin{bmatrix} V_0^k \\ V_1^k \\ \vdots \\ \vdots \\ V_{I-1}^k \\ V_I^k \end{bmatrix},$$

where for $i = 1, 2, \cdots, I-1$

$$A_i^k = \frac{\Delta t}{\Delta x^2}a_i^k - \frac{\Delta t}{2\Delta x}b_i^k, \quad B_i^k = -\frac{2\Delta t}{\Delta x^2}a_i^k + \Delta t c_i^k, \quad C_i^k = \frac{\Delta t}{\Delta x^2}a_i^k + \frac{\Delta t}{2\Delta x}b_i^k,$$

and

$$\begin{aligned} B_0^k &= -\frac{\Delta t}{\Delta x}\widetilde{b}_0^k + \Delta t c_0^k, \quad C_0^k = \frac{\Delta t}{\Delta x}\widetilde{b}_0^k, \\ A_I^k &= -\frac{\Delta t}{\Delta x}\widetilde{b}_I^k, \quad B_I^k = \frac{\Delta t}{\Delta x}\widetilde{b}_I^k + \Delta t c_I^k. \end{aligned}$$

In this case, the use of convection terms at the boundaries did not alter the tridiagonal structure of the matrix. Notice that the one-side discretization we used is first order accurate. If we had used second-order one-sided differencing for the convection terms, we would have altered the tridiagonal nature of the dicretization matrix. Likewise, if we had used the full equation at the boundaries, one-sided discretization would have altered the tridiagonal structure of the discretization matrix.

However, the PDE boundary conditions and modified linear boundary conditions are, in general, more accurate than simple linear boundary conditions.

✎ **Stability**

Let $X = a\frac{\Delta t}{\Delta x^2}$, $Y = c\frac{\Delta t}{2}$ and $Z = b\frac{\Delta t}{2\Delta x}$. After taking the discrete Fourier transform, we have

$$\begin{aligned}
\left[-\theta A e^{-i\xi} + 1 - \theta B - \theta C e^{i\xi}\right]\hat{V}^{k+1} &= \left[(1-\theta)Ae^{-i\xi} + 1 + (1-\theta)B + (1-\theta)Ce^{i\xi}\right]\hat{V}^{k}, \\
\left[2\theta X(1-\cos\xi) + 1 - 2\theta Y - 2i\theta Z\sin\xi\right]\hat{V}^{k+1} &= \left[-2(1-\theta)X(1-\cos\xi) + 1 + 2(1-\theta)Y + 2i(1-\theta)Z\sin\xi\right]\hat{V}^{k}, \\
\left[1 + 2\theta X - 2\theta Y - 2\theta X\cos\xi - 2i\theta Z\sin\xi\right]\hat{V}^{k+1} &= \left[1 - 2(1-\theta)X + 2(1-\theta)Y + 2(1-\theta)X\cos\xi + 2i(1-\theta)Z\sin\xi\right]\hat{V}^{k},
\end{aligned}$$

Let

$$\begin{aligned}
P &:= 1 + 2\theta X - 2\theta Y - 2\theta X\cos\xi - 2i\theta Z\sin\xi, \\
Q &:= 1 - 2(1-\theta)X + 2(1-\theta)Y + 2(1-\theta)X\cos\xi + 2i(1-\theta)Z\sin\xi.
\end{aligned}$$

We want to find a sufficient condition for $\frac{|Q|}{|P|} \le 1$. Straightforward calculations yield

$$\begin{aligned}
&|Q|^2 - |P|^2 \\
= \;& 4(X - Y - X\cos\xi)\Big(-1 + X(1-2\theta)(1-\cos\xi) - Y(1-2\theta)\Big) + 4Z^2(1-2\theta)\sin^2\xi.
\end{aligned}$$

We want this equation to be less than 0.

1. $\frac{1}{2} \le \theta \le 1$: In this case,
$$1 - 2\theta \le 0.$$
Hence we have $|Q|^2 - |P|^2 \le 0$ and $\frac{|Q|}{|P|} \le 1$.

2. $0 \le \theta < \frac{1}{2}$: If
$$\begin{cases} |Z| \le X, \\ 1 - 2(1-2\theta)(X-Y) \ge 0, \\ -\dfrac{1}{1-2\theta} \le Y \le 0, \end{cases}$$
then
$$|Q|^2 - |P|^2 \le 0.$$
The reason is as follows.
$$\begin{aligned}
&|Q|^2 - |P|^2 \\
= \;& -4(X-Y)\Big(1 - (1-2\theta)(X-Y)\Big) + 4X\Big(1 - (1-2\theta)(X-Y)\Big)\cos\xi \\
& 4X^2(1-2\theta)\cos^2\xi + 4Z^2(1-2\theta)\sin^2\xi \\
\le \;& -4(X-Y)\Big(1 - (1-2\theta)(X-Y)\Big) + 4X\Big(1 - (1-2\theta)(X-Y)\Big)\cos\xi + 4X^2(1-2\theta) \qquad (15.6.3)
\end{aligned}$$

We want to Eq(15.6.3) is less than 0. It is equivalent to

$$4X\Big(1-(1-2\theta)(X-Y)\Big)\cos\xi \le 4(X-Y)\Big(1-(1-2\theta)(X-Y)\Big) - 4X^2(1-2\theta).$$

Also, this is equivalent to

$$\cos\xi \quad \le \quad \frac{4(X-Y)\Big(1-(1-2\theta)(X-Y)\Big) - 4X^2(1-2\theta)}{4X\Big(1-(1-2\theta)(X-Y)\Big)}. \tag{15.6.4}$$

The following equation is a sufficient condition for Eq(15.6.4).

$$\begin{aligned} 1 &\le \frac{4(X-Y)\Big(1-(1-2\theta)(X-Y)\Big) - 4X^2(1-2\theta)}{4X\Big(1-(1-2\theta)(X-Y)\Big)} \\ \Longleftarrow \quad 0 &\le 4(X-Y)\Big(1-(1-2\theta)(X-Y)\Big) - 4X^2(1-2\theta) - 4X\Big(1-(1-2\theta)(X-Y)\Big) \end{aligned} \tag{15.6.5}$$

The right hand side of Eq(15.6.5) can be expanded as

$$-4Y - 4(1-2\theta)Y^2 \quad = \quad -4Y\Big((1-2\theta)Y+1\Big).$$

The condition for this to be greater than 0 is

$$-\frac{1}{1-2\theta} \le Y \le 0.$$

This completes the proof.

3. $\theta = 0$: Then

$$\begin{aligned} & |Q|^2 - |P|^2 \\ = \quad & -1 + (1-2X+2Y)^2 - 4X(1-2X+2Y)\cos\xi + X^2\cos^2\xi + Z^2\sin^2\xi. \end{aligned}$$

Example 15.6.1. Assume that the parameters are given as follows:

$$\begin{aligned} a &= \frac{1}{2}\sigma^2, \\ b &= r - \delta - \frac{1}{2}\sigma^2, \\ c &= -r. \end{aligned}$$

Then we have

$$\begin{aligned} A &= \frac{1}{2}\sigma^2\,\frac{\Delta t}{\Delta x^2} - \frac{\Delta t}{2\Delta x}\left(r - \delta - \frac{1}{2}\sigma^2\right), \\ B &= -\sigma^2\,\frac{\Delta t}{\Delta x^2} - r\,\Delta t, \\ C &= \frac{1}{2}\sigma^2\,\frac{\Delta t}{\Delta x^2} + \frac{\Delta t}{2\Delta x}\left(r - \delta - \frac{1}{2}\sigma^2\right), \\ X &= \frac{1}{2}\sigma^2\,\frac{\Delta t}{\Delta x^2}, \end{aligned}$$

$$\begin{aligned} Y &= -\frac{r}{2}\Delta t, \\ Z &= \left(r-\delta-\frac{1}{2}\sigma^2\right)\frac{\Delta t}{2\Delta x}. \end{aligned}$$

If $0 \le \theta < \frac{1}{2}$, then the convergence conditions are as follows.

1. $|Z| \le X$:

$$\begin{aligned} & \left|r-\delta-\frac{1}{2}\sigma^2\right|\frac{\Delta t}{2\Delta x} \le \frac{1}{2}\sigma^2\frac{\Delta t}{\Delta x^2} \\ \iff \quad & \Delta x \le \frac{\sigma^2}{\left|r-\delta-\frac{1}{2}\sigma^2\right|}. \end{aligned}$$

2. $0 \le 1-2(1-2\theta)(X-Y)$: This is equivalent to

$$\begin{aligned} & 0 \le 1-2(1-2\theta)(X-Y) \\ \iff \quad & X-Y \le \frac{1}{2(1-2\theta)} \\ \iff \quad & \Delta t\left(\frac{\sigma^2}{\Delta x^2}+r\right) \le \frac{1}{1-2\theta} \\ \iff \quad & \Delta t \le \frac{1}{1-2\theta}\times\frac{1}{\frac{\sigma^2}{\Delta x^2}+r} \approx \frac{\Delta x^2}{(1-2\theta)\sigma^2} \end{aligned}$$

3. $-\frac{1}{1-2\theta} \le Y \le 0$: We have

$$\Delta t \quad \le \quad \frac{2}{(1-2\theta)r}$$

$\square$

15.7 Greeks with Transformation

1. Delta: From $x = \log S$ and chain rule we have

$$\begin{aligned}\Delta &= \frac{\partial f}{\partial S} \\ &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial S} \\ &= \frac{\partial f}{\partial x}\frac{1}{S}.\end{aligned}$$

2. Gamma: Similarly we get the gamma as follows.

$$\begin{aligned}\Gamma &= \frac{\partial^2 f}{\partial S^2} \\ &= \frac{\partial}{\partial S}\left(\frac{\partial f}{\partial S}\right) \\ &= \frac{\partial}{\partial S}\left(\frac{\partial f}{\partial x}\frac{1}{S}\right) \\ &= \frac{\partial^2 f}{\partial x^2}\frac{1}{S^2} - \frac{\partial f}{\partial x}\frac{1}{S^2} \\ &= \frac{\partial^2 f}{\partial x^2}\frac{1}{S^2} - \Delta\frac{1}{S}.\end{aligned}$$

15.8 Alternating Direction Implicit(ADI)

We do not generally want to try to solve an equation such as

$$Q_1\mathbf{u}^{n+1} = \mathbf{u}^n \tag{15.8.1}$$

by Gaussian elimination because, in general, Gaussian elimination is too expensive to use for solving the matrix equations associated with difference schemes in two or more dimensions.

Another approach is to use any of the iterative schemes to solve equation (15.8.1) (Gauss Seidel, Jacobi, SOR, etc.) Iterative solvers can be and are sometimes used. Generally, iterative solvers are expensive to use to solve equation (15.8.1) at each time step.

A way around solver an equation such as equation (15.8.1) is to use an alternating direction implicit(ADI) scheme.

15.8.1 Peaceman-Rachford Scheme

$$\boxed{\begin{aligned}(1-\frac{r_x}{2}\delta_x^2)u_{jk}^{n+\frac{1}{2}} &= (1+\frac{r_y}{2}\delta_y^2)u_{jk}^n,\\ (1-\frac{r_y}{2}\delta_y^2)u_{jk}^{n+1} &= (1+\frac{r_x}{2}\delta_x^2)u_{jk}^{n+\frac{1}{2}}.\end{aligned}}$$

We evaluate the derivative with respect to x implicitly, the derivative with respect to y explicitly and use a time step of $\Delta t/2$, we get the scheme

$$\frac{u_{jk}^{n+\frac{1}{2}} - u_{jk}^n}{\Delta t/2} = \frac{\nu}{\Delta x^2}\delta_x^2 u_{jk}^{n+\frac{1}{2}} + \frac{\nu}{\Delta y^2}\delta_y^2 u_{jk}^n. \tag{15.8.2}$$

Recall that

$$\begin{aligned}\delta_x^2 u_{jk} &:= u_{j+1k} - 2u_{jk} + u_{j-1k},\\ \delta_y^2 u_{jk} &:= u_{jk+1} - 2u_{jk} + u_{jk-1}.\end{aligned}$$

We rewrite difference scheme (15.8.2) as

$$(1-\frac{r_x}{2}\delta_x^2)u_{jk}^{n+\frac{1}{2}} = (1+\frac{r_y}{2}\delta_y^2)u_{jk}^n, \tag{15.8.3}$$

where $r_x = \frac{\Delta t}{\Delta x^2}$ and $r_y = \frac{\Delta t}{\Delta y^2}$. If we take the discrete Fourier transform of Eq.(15.8.3), we get

$$\begin{aligned}\left(1+2r_x\sin^2\frac{\xi}{2}\right)\hat{u}^{n+\frac{1}{2}} &= \left(1-2r_y\sin^2\frac{\eta}{2}\right)\hat{u}^n,\\ \rho(\xi,\eta) &= \frac{1-2r_y\sin^2\frac{\eta}{2}}{1+2r_x\sin^2\frac{\xi}{2}}.\end{aligned}$$

We see that scheme (15.8.2) is conditionally stable with

$$r_y \leq 1.$$

Now consider the another half step:

$$\frac{u_{jk}^{n+1}-u_{jk}^{n+\frac{1}{2}}}{\Delta t/2} = \frac{\nu}{\Delta x^2}\delta_x^2 u_{jk}^{n+\frac{1}{2}} + \frac{\nu}{\Delta y^2}\delta_y^2 u_{jk}^{n+1}, \tag{15.8.4}$$

(this time explicit in x and implicit in y), or

$$(1-\frac{r_y}{2}\delta_y^2)u_{jk}^{n+1} = (1+\frac{r_x}{2}\delta_x^2)u_{jk}^{n+\frac{1}{2}}, \tag{15.8.5}$$

and take the discrete Fourier transform of the result, we get

$$\begin{aligned}
\Big(1+2r_y\sin^2\frac{\eta}{2}\Big)\hat{u}^{n+1} &= \Big(1-2r_x\sin^2\frac{\xi}{2}\Big)\hat{u}^{n+\frac{1}{2}} \\
&= \Big(1-2r_x\sin^2\frac{\xi}{2}\Big)\frac{1-2r_y\sin^2\frac{\eta}{2}}{1+2r_x\sin^2\frac{\xi}{2}}\hat{u}^n, \\
\rho(\xi,\eta) &= \frac{\big(1-2r_x\sin^2\frac{\xi}{2}\big)\big(1-2r_y\sin^2\frac{\eta}{2}\big)}{\big(1+2r_x\sin^2\frac{\xi}{2}\big)\big(1+2r_y\sin^2\frac{\eta}{2}\big)}.
\end{aligned}$$

It is not hard to see that ρ satisfies $|\rho|\le 1$. Hence, the two step scheme (15.8.2) and (15.8.4), called the Peaceman-Rachford, is unconditionally stable.

15.8.2 D'Yakonov scheme

$$\boxed{\begin{aligned}
(1-\frac{r_x}{2}\delta_x^2)u_{jk}^{*} &= (1+\frac{r_x}{2}\delta_x^2)(1+\frac{r_y}{2}\delta_y^2)u_{jk}^{n}, \\
(1-\frac{r_y}{2}\delta_y^2)u_{jk}^{n+1} &= u_{jk}^{*}.
\end{aligned}}$$

If we operate on both sides of (15.8.3) and (15.8.4) by $(1+\frac{r_x}{2}\delta_x^2)$ and $(1-\frac{r_x}{2}\delta_x^2)$ respectively, we get

$$\begin{aligned}
(1+\frac{r_x}{2}\delta_x^2)(1-\frac{r_x}{2}\delta_x^2)u_{jk}^{n+\frac{1}{2}} &= (1+\frac{r_x}{2}\delta_x^2)(1+\frac{r_y}{2}\delta_y^2)u_{jk}^{n}, \\
(1-\frac{r_x}{2}\delta_x^2)(1+\frac{r_x}{2}\delta_x^2)u_{jk}^{n+\frac{1}{2}} &= (1+\frac{r_x}{2}\delta_x^2)(1+\frac{r_y}{2}\delta_y^2)u_{jk}^{n}, \\
(1-\frac{r_x}{2}\delta_x^2)(1-\frac{r_y}{2}\delta_y^2)u_{jk}^{n+1} &= (1-\frac{r_x}{2}\delta_x^2)(1+\frac{r_x}{2}\delta_x^2)u_{jk}^{n+\frac{1}{2}}, \\
(1-\frac{r_x}{2}\delta_x^2)(1-\frac{r_y}{2}\delta_y^2)u_{jk}^{n+1} &= (1+\frac{r_x}{2}\delta_x^2)(1+\frac{r_y}{2}\delta_y^2)u_{jk}^{n},
\end{aligned}$$

If we split this equation into

$$\begin{aligned}
(1-\frac{r_x}{2}\delta_x^2)u_{jk}^{*} &= (1+\frac{r_x}{2}\delta_x^2)(1+\frac{r_y}{2}\delta_y^2)u_{jk}^{n}, \\
(1-\frac{r_y}{2}\delta_y^2)u_{jk}^{n+1} &= u_{jk}^{*}.
\end{aligned}$$

This scheme is referred to as the D'Yakonov scheme.

15.8.3 Douglas-Rachford Scheme

$$
\begin{aligned}
(1 - r_x\delta_x^2)u^*_{jk} &= (1 + r_y\delta_y^2)u^n_{jk}, \\
(1 - r_y\delta_y^2)u^{n+1}_{jk} &= u^*_{jk} - r_y\delta_y^2 u^n_{jk}.
\end{aligned}
$$

If we write the implicit scheme,

$$u^{n+1}_{jk} - r_x\delta_x^2 u^{n+1}_{jk} - r_y\delta_y^2 u^{n+1}_{jk} = u^n_{jk},$$

as

$$(1 - r_x\delta_x^2 - r_y\delta_y^2)u^{n+1}_{jk} = u^n_{jk}, \tag{15.8.6}$$

we see that the left hand side naturally factors into

$$(1 - r_x\delta_x^2)(1 - r_y\delta_y^2)u^{n+1}_{jk}.$$

Thus to factor equation(15.8.6), we must add a term of the form

$$r_x r_y \delta_x^2\delta_y^2 u^{n+1}_{jk}.$$

If we add

$$r_x r_y \delta_x^2\delta_y^2 u^n_{jk}$$

to the left hand side, the Douglas-Rachford scheme can be written as

$$(1 - r_x\delta_x^2)(1 - r_y\delta_y^2)u^{n+1}_{jk} = (1 + r_x r_y\delta_x^2\delta_y^2)u^n_{jk}.$$

The form of the Douglas-Rachford scheme used for computations is

$$
\begin{aligned}
(1 - r_x\delta_x^2)u^*_{jk} &= (1 + r_y\delta_y^2)u^n_{jk} \\
(1 - r_y\delta_y^2)u^{n+1}_{jk} &= u^*_{jk} - r_y\delta_y^2 u^n_{jk}.
\end{aligned} \tag{15.8.7}
$$

We must select the boundary conditions for u^* carefully. Solving u^* from equation(15.8.7) yields

$$u^*_{jk} = (1 - r_y\delta_y^2)u^{n+1}_{jk} + r_y\delta_y^2 u^n_{jk}.$$

Then the boundary conditions for u^* at $j = 0$ and $j = M_x$ can be given in terms of the given Dirichlet boundary conditions of u by

$$
\begin{aligned}
u^*_{0k} &= (1 - r_y\delta_y^2)u^{n+1}_{0k} + r_y\delta_y^2 u^n_{0k}, \\
u^*_{M_x k} &= (1 - r_y\delta_y^2)u^{n+1}_{M_x k} + r_y\delta_y^2 u^n_{M_x k}.
\end{aligned}
$$

For a detail, refer to Thomas(1997).

15.8.4 ADI with Mixed Derivatives

It is well-known that "cross terms" can be difficult to handle implicitly using ADI technique. In this subsection, we study an ADI scheme with mixed derivatives.

We consider initial value problems of the form

$$\frac{\partial u}{\partial t} = \mathcal{L}u,$$

where $\mathcal{L}$ is a partial differential operator

$$\mathcal{L} = \sum_{i=1}^{N}\sum_{j=1}^{N} q_{ij}\partial_i\partial_j, \quad \partial_i = \frac{\partial}{\partial x_i}, q_{ij} = q_{ji}. \tag{15.8.8}$$

The parabolicity of problem(15.8.8) implies that the symmetric matrix $Q = (q_{ij})$ is positive definite so that problem(15.8.8) is well-posed.

In the follows we shall use the symbol $u^n_{j_1,\cdots,j_N}$ to denote the finite difference solution at the node point $(j_1\Delta x_1, \cdots, j_N\Delta x_N, n\Delta t)$ under the assumption that $\Delta x_i = \Delta$ defines a uniform space mesh. Operators such as(for example, in two space dimensions)

$$\begin{aligned}
\delta_x u_{ij} &= u_{i+1,j} - 2u_{ij} + u_{i-1,j}, \\
\delta_{xj} u_{ij} &= u_{i+1,j+1} - u_{i+1,j-1} - u_{i-1,j+1} + u_{i-1,j-1}, \\
\nabla_x u_{ij} &= u_{i+1,j} - u_{i-1,j}
\end{aligned}$$

define conventional central difference operators.

The basic scheme we consider can be split into ADI form and resolved as a sequence of N tridiagonal matrix operators, namely

$$\begin{aligned}
(1 - r\theta q_{11}\delta^2_{x_1})u^{n+1(1)} &= \left[1 + r(1-\theta)q_{11}\delta^2_{x_1} + r\sum_{i=1}^{N} q_{ii}\delta^2_{x_i} + \frac{1}{2}r\sum_{i=2}^{N}\sum_{j=1}^{i-1} q_{ij}\delta_{x_i x_j}\right]u^n, \\
(1 - r\theta q_{22}\delta^2_{x_2})u^{n+1(2)} &= u^{n+1(1)} - r\theta q_{22}\delta^2_{x_2}u^n, \\
&\vdots \\
(1 - r\theta q_{NN}\delta^2_{x_N})u^{n+1} &= u^{n+1(N-1)} - r\theta q_{NN}\delta^2_{x_N}u^n,
\end{aligned}$$

where $r = \Delta t/\Delta^2$, $u^{n+1(i)}$ denotes the approximation to u^{n+1} at split level (i) and θ is a real parameter that determines the implicitness of the method.

We note, first of all that this scheme can be regarded as a natural extension of previous ADI methods for parabolic equations. For example when $N = 2$ and the cross terms are absent (i.e. $q_{ij} = 0, i \neq j$), the choice $\theta = 1$ defines an unconditionally stable Douglas-Rachford scheme of $O(\Delta t) + O(\Delta^2)$, whereas the choice $\theta = \frac{1}{2}$ yields a higher-order scheme $O(\Delta t^2) + O(\Delta^2)$ which again unconditionally stable. When mixed derivatives are present however, the accuracy remains $O(\Delta t) + O(\Delta^2)$ independent of θ, owing to the one-sided time differencing of the cross terms. For a detail, see Craig and Sneyd(1988).

In the case of two space dimensions with mixed derivatives, this scheme has the same structure as the stable($\theta \geq \frac{1}{2}$) ADI.

Now, we consider a more general initial value problems of the form

$$\frac{\partial u}{\partial t} = \mathcal{L}u, \tag{15.8.9}$$

where $\mathcal{L}$ is a partial differential operator

$$\mathcal{L} = \sum_{i=1}^{N}\sum_{j=1}^{N} q_{ij}\partial_i\partial_j + \sum_{i=1}^{n} p_i\partial_i + g, \quad \partial_i = \frac{\partial}{\partial x_i}, q_{ij} = q_{ji}.$$

The proposed ADI scheme for problem(15.8.9) is given by

$$\begin{aligned}
\left\{1-\theta\left(rq_{11}\delta_{x_1}^2+\lambda p_1\nabla_{x_1}+\eta g\right)\right\}u^{n+1(1)} &= \left[1+(1-\theta)\left(rq_{11}\delta_{x_1}^2+\lambda p_1\nabla_{x_1}+\eta g\right)+r\sum_{i=1}^{N}q_{ii}\delta_{x_i}^2+\frac{1}{2}r\sum_{i=2}^{N}\sum_{j=1}^{i-1}q_{ij}\delta_{x_ix_j}\right]u^n,\\
\left\{1-\theta\left(rq_{22}\delta_{x_2}^2+\lambda p_2\nabla_{x_2}+\eta g\right)\right\}u^{n+1(2)} &= u^{n+1(1)}-\theta\left(rq_{22}\delta_{x_2}^2+\lambda p_2\nabla_{x_2}+\eta g\right)u^n,\\
&\vdots\\
\left\{1-\theta\left(rq_{NN}\delta_{x_N}^2+\lambda p_N\nabla_{x_N}+\eta g\right)\right\}u^{n+1} &= u^{n+1(N-1)}-\theta\left(rq_{NN}\delta_{x_N}^2+\lambda p_N\nabla_{x_N}+\eta g\right)u^n,
\end{aligned}$$

where $\lambda = \frac{\Delta t}{2\Delta}$ and $\eta = \frac{\Delta t}{N}$.

15.9 Operator Splitting Method

In this section, we follow Duffy(2004) and *Numerical Recipes in C++*(2002).

The basic idea of operator splitting, which is called *time splitting* or *the method of fractional step*, is this: Suppose you have an initial value equation of the form

$$\frac{\partial u}{\partial t} = \mathcal{L}u$$

where $\mathcal{L}$ is some operator. Suppose that it can be written as a linear sum of m pieces, which act additively on u,

$$\mathcal{L}u = \mathcal{L}_1 u + \mathcal{L}_2 u + \cdots \mathcal{L}_m u.$$

Finally, suppose that for each of the pieces, you already know a scheme for updating the variable u from time step n to time step $n+1$, i.e

$$\begin{aligned} u^{n+1} &= \mathcal{U}_1(u^n, \Delta t), \\ u^{n+1} &= \mathcal{U}_2(u^n, \Delta t), \\ &\cdots \\ u^{n+1} &= \mathcal{U}_m(u^n, \Delta t). \end{aligned}$$

Now, one form of operator splitting would be to get from n to $n+1$ by the following sequence of updating:

$$\begin{aligned} u^{n+\left(\frac{1}{m}\right)} &= \mathcal{U}_1(u^n, \Delta t), \\ u^{n+\left(\frac{2}{m}\right)} &= \mathcal{U}_2(u^{n+\left(\frac{1}{m}\right)}, \Delta t), \\ &\cdots \\ u^{n+1} &= \mathcal{U}_m(u^{n+\left(\frac{m-1}{m}\right)}, \Delta t). \end{aligned}$$

Example 15.9.1. Consider the two-dimensional heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}.$$

In the case of the explicit splitting scheme we get

$$\begin{aligned} \frac{\widetilde{u}_{ij} - u^n_{ij}}{\Delta t} &= \frac{u^n_{i+1,j} - 2u^n_{i,j} + u^n_{i-1,j}}{\Delta x^2}, \\ \frac{u^{n+1}_{ij} - \widetilde{u}_{ij}}{\Delta t} &= \frac{\widetilde{u}_{i,j+1} - 2\widetilde{u}_{i,j} + \widetilde{u}_{i,j-1}}{\Delta y^2}. \end{aligned}$$

If the mesh size in both x and y direction is a constant $\Delta x = \Delta y = h$, the explicit scheme is stable under the condition

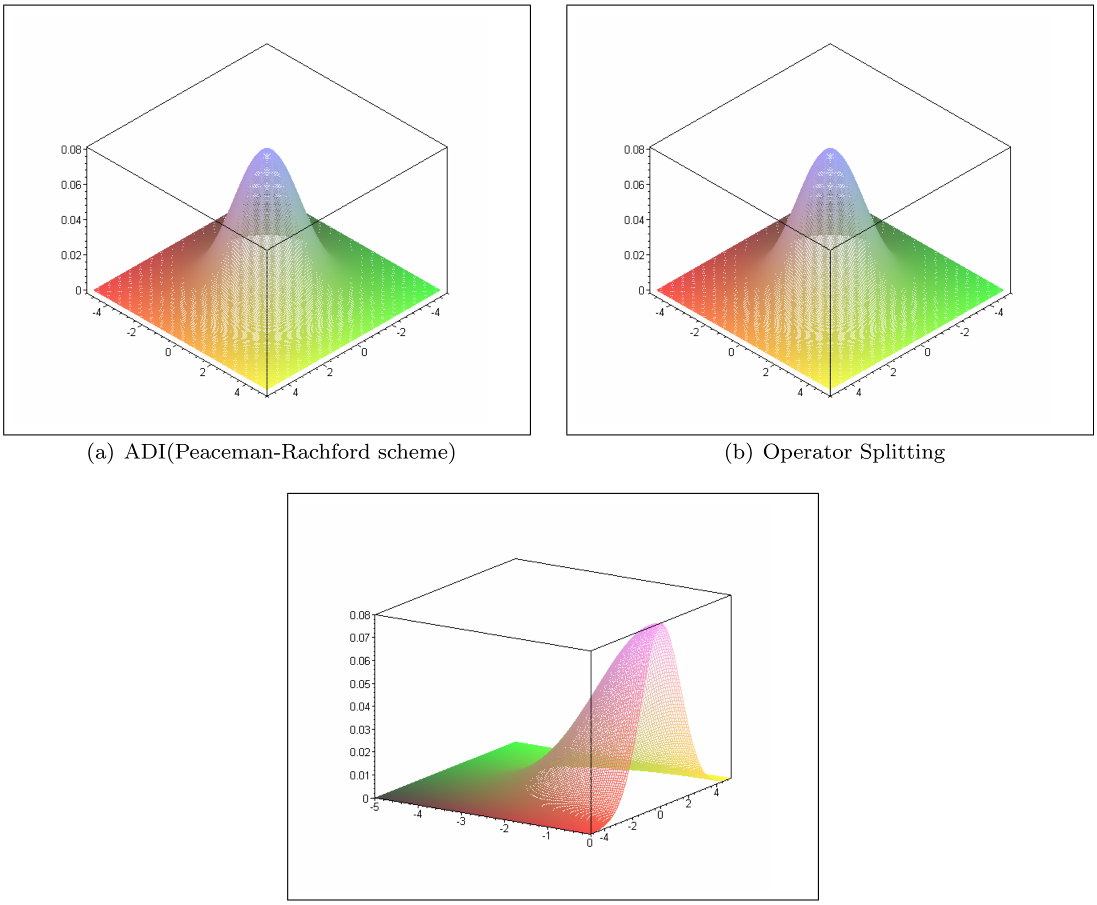
$$\frac{\Delta t}{h^2} \leq \frac{1}{2}.$$

The implicit splitting scheme for the two-dimensional heat equation is given by

$$\frac{\widetilde{u}_{ij} - u^n_{ij}}{\Delta t} = \frac{\widetilde{u}_{i+1,j} - 2\widetilde{u}_{i,j} + \widetilde{u}_{i-1,j}}{\Delta x^2},$$

$$\frac{u_{ij}^{n+1}-\widetilde{u}_{ij}}{\Delta t} = \frac{u_{i,j+1}^{n+1}-2u_{i,j}^{n+1}+u_{i,j-1}^{n+1}}{\Delta y^2}.$$

Figure(15.4) shows the comparison of ADI and Splitting method for the fundamental solution of heat equation. □



(a) ADI(Peaceman-Rachford scheme) (b) Operator Splitting

(c) Overlapped graph(indistinguishable!!!)

Figure 15.4: Numerical solutions of 2-dimension heat equation.

Example 15.9.2. Consider the following equations:

$$\frac{\partial u}{\partial t} = \mathcal{L}u, \quad \mathcal{L}u=\sum_{i,j=1}^{2} a_{ij}\frac{\partial^2 u}{\partial x_i \partial x_j}, \quad a_{11}a_{22}-a_{12}^2>0, \quad a_{12}=a_{21}, a_{11}>0, a_{22}>0.$$

Yanenko(1971) proposed the following stable and convergent splitting scheme

$$\begin{aligned}\frac{\widetilde{u}_{ij}-u_{ij}^n}{\Delta t} &= a_{11}\frac{\widetilde{u}_{i+1,j}-2\widetilde{u}_{i,j}+\widetilde{u}_{i-1,j}}{\Delta x_1^2}+a_{12}\frac{u_{i+1,j+1}^n-u_{i+1,j-1}^n-u_{i-1,j+1}^n+u_{i-1,j-1}^n}{4\Delta x_1\Delta x_2},\\ \frac{u_{ij}^{n+1}-\widetilde{u}_{ij}}{\Delta t} &= a_{22}\frac{u_{i,j+1}^{n+1}-2u_{i,j}^{n+1}+u_{i,j-1}^{n+1}}{\Delta x_2^2}+a_{12}\frac{\widetilde{u}_{i+1,j+1}-\widetilde{u}_{i+1,j-1}-\widetilde{u}_{i-1,j+1}+\widetilde{u}_{i-1,j-1}}{4\Delta x_1\Delta x_2}.\end{aligned}$$

□

✎ **Two-Dimensional Black-Scholes PDE**
After log transformation we have

$$
\begin{aligned}
-\frac{\partial f}{\partial t} &= \mathcal{L}f, \quad \text{where} \\
\mathcal{L} &:= \frac{1}{2}\sigma_1^2\frac{\partial^2}{\partial x^2}+\frac{1}{2}\sigma_2^2\frac{\partial^2}{\partial y^2}+\rho\sigma_1\sigma_2\frac{\partial^2}{\partial x\partial y}+\nu_1\frac{\partial}{\partial x}+\nu_2\frac{\partial}{\partial y}-r,
\end{aligned}
$$

where $\nu_1 = r-\delta_1-\frac{1}{2}\sigma_1^2$, and $\nu_2 = r-\delta_2-\frac{1}{2}\sigma_2^2$.
We can split $\mathcal{L}$ into two operators $\mathcal{L}_1$ and $\mathcal{L}_2$ such that

$$
\begin{aligned}
\mathcal{L}_1 f &:= \frac{1}{2}\sigma_1^2\frac{\partial^2 f}{\partial x^2}+\nu_1\frac{\partial f}{\partial x}+\lambda_1\rho\sigma_1\sigma_2\frac{\partial^2 f}{\partial x\partial y}-\lambda_2 r f, \\
\mathcal{L}_2 f &:= \frac{1}{2}\sigma_2^2\frac{\partial^2 f}{\partial y^2}+\nu_2\frac{\partial f}{\partial y}+(1-\lambda_1)\rho\sigma_1\sigma_2\frac{\partial^2 f}{\partial x\partial y}-(1-\lambda_2) r f,
\end{aligned}
$$

where $\lambda_1, \lambda_2 \in [0,1]$. Let

$$
u_{ij}^n := f(x_{min}+i\Delta x, y_{min}+j\Delta y, T-n\Delta t),
$$

where T is the option maturity. Thus we have the following splitting scheme.

$$
\begin{aligned}
&\frac{u_{ij}^n-\widetilde{u}_{ij}}{\Delta t}+(1-\theta)\frac{\sigma_1^2}{2}\frac{u_{i+1,j}^n-2u_{i,j}^n+u_{i-1,j}^n}{\Delta x^2}+\theta\frac{\sigma_1^2}{2}\frac{\widetilde{u}_{i+1,j}-2\widetilde{u}_{i,j}+\widetilde{u}_{i-1,j}}{\Delta x^2} \\
&+(1-\theta)\nu_1\frac{u_{i+1,j}^n-u_{i-1,j}^n}{2\Delta x}+\theta\nu_1\frac{\widetilde{u}_{i+1,j}-\widetilde{u}_{i-1,j}}{2\Delta x}-(1-\theta)\lambda_2 r u_{i,j}^n-\theta\lambda_2 r\widetilde{u}_{i,j} \\
&+\lambda_1\rho\sigma_1\sigma_2\frac{u_{i+1,j+1}^n-u_{i+1,j-1}^n-u_{i-1,j+1}^n+u_{i-1,j-1}^n}{4\Delta x\Delta y} = 0 \\
&\frac{\widetilde{u}_{ij}-u_{ij}^{n+1}}{\Delta t}+(1-\theta)\frac{\sigma_2^2}{2}\frac{\widetilde{u}_{i,j+1}-2\widetilde{u}_{i,j}+\widetilde{u}_{i,j-1}}{\Delta y^2}+\theta\frac{\sigma_2^2}{2}\frac{u_{i,j+1}^{n+1}-2u_{i,j}^{n+1}+u_{i,j-1}^{n+1}}{\Delta y^2} \\
&+(1-\theta)\nu_2\frac{\widetilde{u}_{i,j+1}-\widetilde{u}_{i,j-1}}{2\Delta y}+\theta\nu_2\frac{u_{i,j+1}^{n+1}-u_{i,j-1}^{n+1}}{2\Delta y}-(1-\theta)(1-\lambda_2) r\widetilde{u}_{i,j}-\theta(1-\lambda_2) r u_{i,j}^{n+1} \\
&+(1-\lambda_1)\rho\sigma_1\sigma_2\frac{\widetilde{u}_{i+1,j+1}-\widetilde{u}_{i+1,j-1}-\widetilde{u}_{i-1,j+1}+\widetilde{u}_{i-1,j-1}}{4\Delta x\Delta y} = 0.
\end{aligned}
$$

Note that the updating order is as follows:

$$
u_*^n \implies \widetilde{u}_* \implies u_*^{n+1}, \qquad n=0,1,\cdots.
$$

Now we can simplify the equations:

- for fixed j

$$
\begin{aligned}
&\theta\left(-\frac{1}{2}\sigma_1^2\frac{\Delta t}{\Delta x^2}+\nu_1\frac{\Delta t}{2\Delta x}\right)\widetilde{u}_{i-1,j}+\left(1+\theta\sigma_1^2\frac{\Delta t}{\Delta x^2}+\theta\lambda_2 r\Delta t\right)\widetilde{u}_{i,j}+\theta\left(-\frac{1}{2}\sigma_1^2\frac{\Delta t}{\Delta x^2}-\nu_1\frac{\Delta t}{2\Delta x}\right)\widetilde{u}_{i+1,j} \\
= &(1-\theta)\left(\frac{1}{2}\sigma_1^2\frac{\Delta t}{\Delta x^2}-\nu_1\frac{\Delta t}{2\Delta x}\right)u_{i-1,j}^n+\left(1-(1-\theta)\sigma_1^2\frac{\Delta t}{\Delta x^2}-(1-\theta)\lambda_2 r\Delta t\right)u_{ij}^n+(1-\theta)\left(\frac{1}{2}\sigma_1^2\frac{\Delta t}{\Delta x^2}+\nu_1\frac{\Delta t}{2\Delta x}\right)u_{i+1,j}^n \\
&+\lambda_1\rho\sigma_1\sigma_2\Delta t\frac{u_{i+1,j+1}^n-u_{i+1,j-1}^n-u_{i-1,j+1}^n+u_{i-1,j-1}^n}{4\Delta x\Delta y},
\end{aligned}
$$

- for fixed i

$$
\begin{aligned}
&\theta\left(-\frac{1}{2}\sigma_2^2\frac{\Delta t}{\Delta y^2}+\nu_2\frac{\Delta t}{2\Delta y}\right)u_{i,j-1}^{n+1}+\left(1+\theta\sigma_2^2\frac{\Delta t}{\Delta y^2}+\theta(1-\lambda_2) r\Delta t\right)u_{i,j}^{n+1}+\theta\left(-\frac{1}{2}\sigma_2^2\frac{\Delta t}{\Delta y^2}-\nu_2\frac{\Delta t}{2\Delta y}\right)u_{i,j+1}^{n+1} \\
= &(1-\theta)\left(\frac{1}{2}\sigma_2^2\frac{\Delta t}{\Delta y^2}-\nu_2\frac{\Delta t}{2\Delta y}\right)\widetilde{u}_{i,j-1}+\left(1-(1-\theta)\sigma_2^2\frac{\Delta t}{\Delta y^2}-(1-\theta)(1-\lambda_2) r\Delta t\right)\widetilde{u}_{i,j}+(1-\theta)\left(\frac{1}{2}\sigma_2^2\frac{\Delta t}{\Delta y^2}+\nu_2\frac{\Delta t}{2\Delta y}\right)\widetilde{u}_{i,j+1} \\
&+(1-\lambda_1)\rho\sigma_1\sigma_2\Delta t\frac{\widetilde{u}_{i+1,j+1}-\widetilde{u}_{i+1,j-1}-\widetilde{u}_{i-1,j+1}+\widetilde{u}_{i-1,j-1}}{4\Delta x\Delta y}.
\end{aligned}
$$

Remark 15.9.3. The cross-derivative $\frac{\partial^2 f}{\partial x \partial y}$ can be approximated by

$$\frac{\partial^2 f}{\partial x \partial y} \approx \frac{f_{i+1,j+1} - f_{i+1,j-1} - f_{i-1,j+1} + f_{i-1,j-1}}{4\Delta x \Delta y}.$$

Alternatively, we have

$$\begin{aligned} f_{i+1,j+1} &\approx f_{ij} + \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\Delta x^2 + 2\frac{\partial^2 f}{\partial x \partial y}\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}\Delta y^2\right) \\ f_{i-1,j-1} &\approx f_{ij} - \frac{\partial f}{\partial x}\Delta x - \frac{\partial f}{\partial y}\Delta y + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\Delta x^2 + 2\frac{\partial^2 f}{\partial x \partial y}\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}\Delta y^2\right), \end{aligned}$$

then the summation gives us

$$\begin{aligned} \frac{\partial^2 f}{\partial x \partial y} &\approx \frac{f_{i+1,j+1} - 2f_{i,j} + f_{i-1,j-1}}{2\Delta x \Delta y} - \frac{\Delta x}{2\Delta y}\frac{\partial^2 f}{\partial x^2} - -\frac{\Delta y}{2\Delta x}\frac{\partial^2 f}{\partial y^2} \\ &\approx \frac{1}{2}\frac{f_{i+1,j+1} - f_{i,j+1} - f_{i+1,j} + f_{i,j}}{\Delta x \Delta y} + \frac{1}{2}\frac{f_{i,j} - f_{i-1,j} - f_{i,j-1} + f_{i-1,j-1}}{\Delta x \Delta y}. \end{aligned}$$

Example 15.9.4 (2-dimensional ELS).
기초자산이 2개인 조기상환 double barrier ELS에 대하여, ADI와 operator splitting method로 각각 해를 구해 보면 Figure(15.5)와 같은 결과를 얻을 수 있다.

1. ADI로 부터 얻어진 결과를 그린 Figure(15.5(a))와 operator splitting method로 구한 결과를 그린 Figure(15.5(b))를 서로 비교해 보면 거의 같은 값임을 알 수 있다. ADI Method에서는 cross term을 없애기 위해 heat equation으로 변환을 하기 때문에, 최종적인 결과의 grid line이 비스듬하다는 것을 확인 할 수 있다. 반면, operator splitting의 결과에서는 grid line이 수직으로 교차하고 있음을 볼 수 있다.

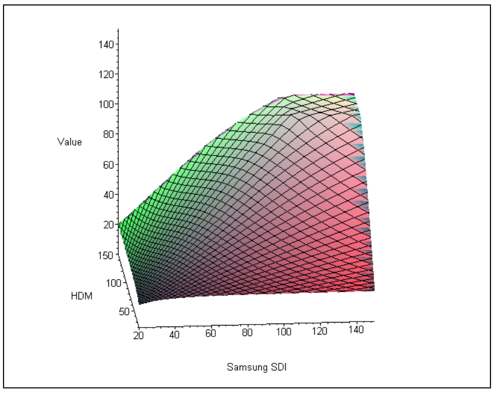
2. Figure (15.5(c))에서는 두 가지 방법을 통해 구한 결과를 contour plot 형태로 보여주고 있다. 얼핏 보아서는 두 결과가 눈으로 구별할 수 없을 정도로 일치하지만, 자세히 보면, 가장 위쪽의 등고선은 약간 불일치 하다는 것을 볼 수 있다. 또한 그 중 하나는 등고선이 매끈하지 못하다는 것도 볼 수 있다. 매끈하지 못한 결과가 ADI로 부터 얻은 결과 인데, 그 원인은 heat equation변환 과정에서 발생하는 interpolation 때문으로 여겨진다.

3. operator splitting method에서는 단순한 log변환만을 거치기 때문에, 그 결과가 부드럽운 등고선으로 나타난다.

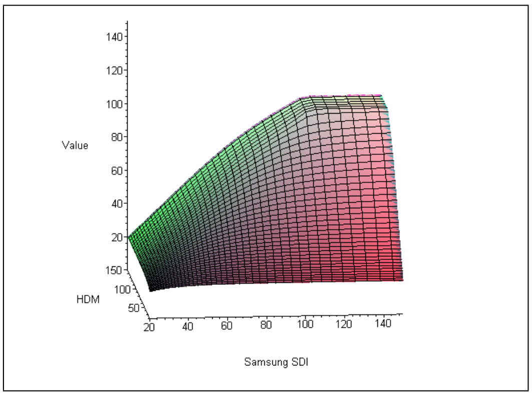
4. grid를 200×200으로 하고, 하루를 3번 분할(만기 3년일 경우 약 $3 \times 3 \times 365 = 3285$번 분할) 했을 경우, ADI는 약 82초 정도가 소요되고, operator splitting method는 약 52초가 소요된다. 실행 시간은 컴퓨터의 성능에 따라 달라질 수 있지만, operator splitting method가 ADI보다 좀 더 빠르다는 것을 알 수 있다 [4].

□

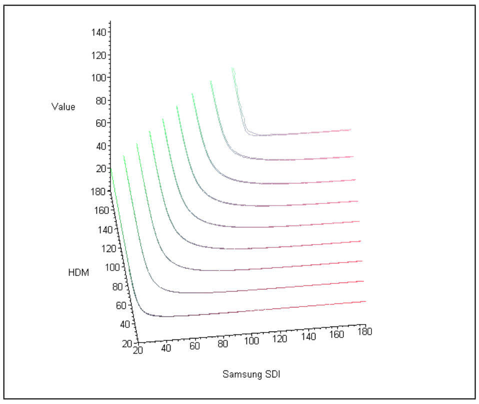
[4] Intel centrino 1.6GHz에서 실행한 결과임. grid를 300×300으로 했을 경우, 각각 196초, 146초가 소요된다.



(a) ADI(Peaceman-Rachford scheme)



(b) Operator Splitting



(c) Overlapped contour plot

Figure 15.5: Numerical solutions for a 2-dimension ELS.

15.10 How To Handle Jumps in FDM

15.10.1 Discrete Cashflow(a jump in the value)

Assume that we divide one day by N.

$$t_{d-1,N} = t_{d,0}, \quad t_{d,1}, \quad \cdots, \quad t_{d,N} = t_{d+1,0}.$$

The FDM is a relation between the value at time

$$t^+_{d,i} \quad \sim \quad t^-_{d,i+1}$$

1. If day t_d is a coupon date, the contract value must jump by the amount(C) of cashflow. In continuous time we have

$$V(S, t^-_{d-1,N}) \quad = \quad V(S, t^+_{d,0}) + C$$

2. If day t_d is a callable or puttable date, the condition is applicable to the relation between

$$t^+_{d,i} \quad \sim \quad t^-_{d,i+1}, \quad i = 0, \cdots, N-1.$$

15.10.2 Discrete Dividend(a jump in the underlying asset)

Suppose that a dividend of D is paid on date t_d. Across this dividend date the asset value drops by an amount D but the option value does not change. Mathematically we write

$$V(S, t_d^-) = V(S - D, t_d^+).$$

15.11 Summing Up

15.11.1 Heat Equation Version

	Scheme	Matrices
explicit	$u_k^{n+1} = (1 + r_x\delta_x^2)u_k^n$	$\lvert r_x \quad 1-2r_x \quad r_x \rvert$
implicit	$(1 - r_x\delta_x^2)u_k^{n+1} = u_k^n$	$\lvert -r_x \quad 1+2r_x \quad -r_x \rvert$
C-R	$(1 - \frac{r_x}{2}\delta_x^2)u_k^{n+1} = (1 + \frac{r_x}{2}\delta_x^2)u_k^n$	$\lvert -\frac{r_x}{2} \quad 1+r_x \quad -\frac{r_x}{2} \rvert \frac{r_x}{2} \quad 1-r_x \quad +\frac{r_x}{2} \rvert$

Table 15.2: 1-dimensional FDM schemes

	Scheme	Matrices
step 1	$(1 - \frac{r_x}{2}\delta_x^2)u_{jk}^{n+\frac{1}{2}} = (1 + \frac{r_y}{2}\delta_y^2)u_{jk}^n$	$\lvert -\frac{r_x}{2} \quad 1+r_x \quad -\frac{r_x}{2} \rvert \frac{r_y}{2} \quad 1-r_y \quad +\frac{r_y}{2} \rvert$
step 2	$(1 - \frac{r_y}{2}\delta_y^2)u_{jk}^{n+1} = (1 + \frac{r_x}{2}\delta_x^2)u_{jk}^{n+\frac{1}{2}}$	$\lvert -\frac{r_y}{2} \quad 1+r_y \quad -\frac{r_y}{2} \rvert \frac{r_x}{2} \quad 1-r_x \quad +\frac{r_x}{2} \rvert$

Table 15.3: 2-dimensional ADI(Peaceman-Rachford) schemes

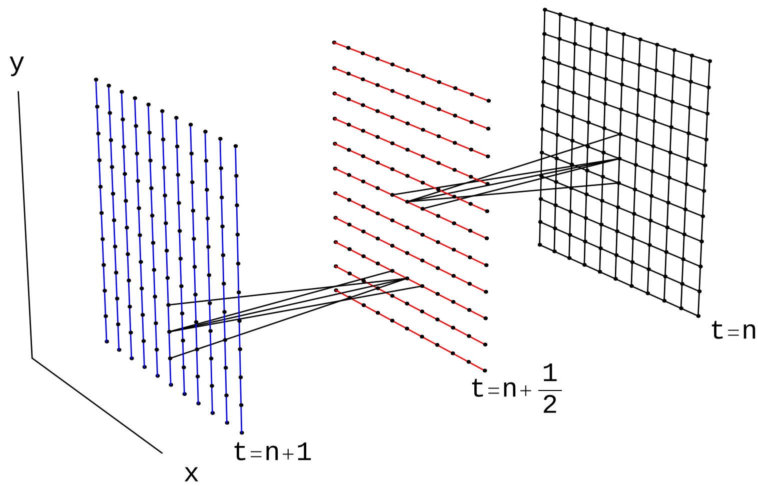


Figure 15.6: Alternating direction implicit method

15.12 Boundary Conditions

In financial problems, we frequently deal with infinite or semi-infinite domains. For example, if we transform the pricing equation for an equity option whose underlying process is log normal to $x = \log S$ coordinates, then the solution domain is $-\infty < x < \infty$.

Obviously, an infinite grid cannot be represented in the computer, so we must artificially truncate the solution domain and replace the deleted portions with boundary conditions that minimize the deleterious effects of the truncation. It is a particularly difficult problem for hyperbolic PDEs where a poor boundary condition can reflect an out-going disturbance back into the solution domain with little loss of amplitude, destroying the desired solution.

We are a bit more fortunate in financial modeling, because there is usually a dominant diffusion process, volatility, that attenuates disturbances from an imperfect boundary condition. We place boundaries far enough from the region of interest, so that even if the imposed boundary conditions are imperfect, it dose not materially affect the solution. For a detail, see Tavella and Randall(2000).

15.12.1 Pricing Barrier Options In The PDE Framework

In the PDE setting, barrier options are even simpler than plain vanilla options because we have more information about the boundaries.

Let $V(S,t)$ denote the value of the barrier contract before the barrier has been triggered. $V(S,t)$ still satisfies the Black-Scholes PDE

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r-\delta) S \frac{\partial V}{\partial S} - rV \quad = \quad 0.$$

✎ **Down-and-In Barrier**
Let T be the expiry of barrier contract and S_d the barrier. If the underlying asset does not reach

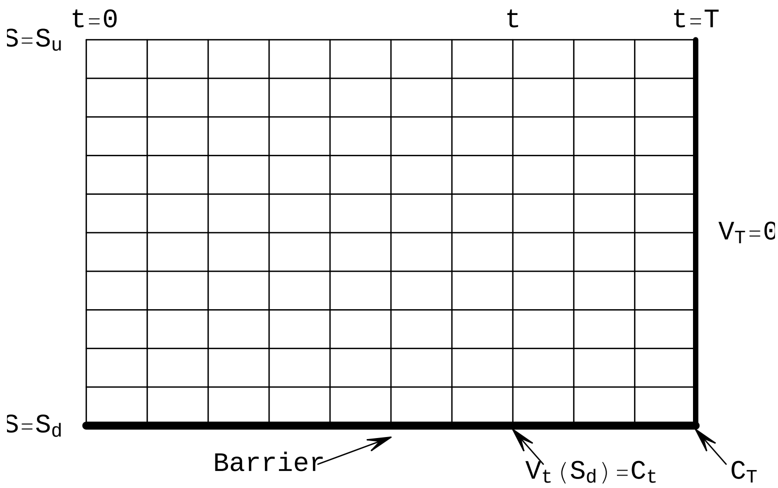


Figure 15.7: Boundary condition of a down-and-in barrier option

the barrier before expiry, the contract becomes worthless. This leads to the final condition

$$V(S,T) \quad = \quad 0 \quad \text{for } S > S_d.$$

From the continuity of V we have the boundary condition

$$V(S_d,t) \quad = \quad \lim_{S\to S_d^+} V(S,t) \quad = \quad \text{value of vanilla call option} \quad = \quad C(S,t),$$

where $C(S,t)$ is given by

$$C(S,t) \quad = \quad e^{-r(T-t)}E[C(S,T)] \quad = \quad e^{-r(T-t)}E[(S-K)^+].$$

We now can extend $V(S,t)$ to $\widetilde{V}(S,t)$ defined where $(S,t) \in (0,\infty) \times [0,T]$ as follows:

$$\widetilde{V}(S,t) \quad := \quad \begin{cases} V(S,t), & S \geq S_d, \\ C(S,t), & S \leq S_d. \end{cases}$$

Note that $V(S_d,t) = C(S_d,t)$ for all t.

Since each of $V(S,t)$ and $C(S,t)$ satisfies the Black-Scholes equation, $\widetilde{V}$ also satisfies the Black-Scholes equation.

Chapter 16

Monte Carlo Simulations

16.1 Introduction to Monte Carlo Method

16.1.1 What is Monte Carlo Simulation?

사람들 중에는 randomness를 좋아하지 않는 사람들이 있다. 이는 randomness가 위험(risk)을 뜻하기 때문일 것이다. 그러나 이러한 randomness를 이용해서 우리가 원하는 계산을 할 수도 있다. random한 시행들의 기대값을 취함으로써 그로 부터 적분값을 구하거나 편미분 방정식을 풀 수도 있다. 해석적인(analytic) 방법으로 해결 할 수 없는 수 많은 문제들이 random sample 생성을 통하여 해결될 수도 있다. 그 중 random sample을 생성하여 얻어지는 결과값들의 평균을 통해 해를 찾는 방법을 Monte Carlo Simulation 이라 한다.

다음의 간단한 예를 통해 Monte Carlo Simulation이 어떤 것이지 살펴보자.

Example 16.1.1. Monte Carlo simulation을 이용하여, 원주율 π 를 추정해 보자. 반지름 1인 사분원의 면적은 $\frac{\pi}{4}$이다.

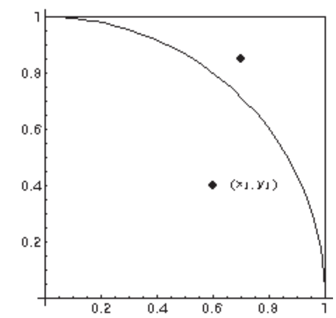


Figure 16.1: Monte Carlo Simulation을 이용한 π 계산

0과 1사이의 random number의 쌍 (x_j, y_j)를 생성해보자. 그리고, 이로 부터 확률변수 X_j를 다음과 같이 정의하자.

$$X_j = \begin{cases} 1, & \sqrt{x_j^2 + y_j^2} \le 1 \\ 0, & \sqrt{x_j^2 + y_j^2} > 1. \end{cases}$$

그러면, IID[1] X_j의 기대값은

$$
\begin{aligned}
E[X_j] &= \text{점}(x_j, y_j)\text{가 사분원의 내부에 있을 확률} \\
&= \text{전체 면적에 대한 사분원의 면적} \\
&= \frac{\pi}{4}
\end{aligned}
$$

이다. 이때, 추정량

$$
\begin{aligned}
\hat{\theta} &= \frac{1}{N}(X_1 + X_2 + \cdots + X_N) \\
&= \frac{\text{사분원의 내부에 있는 점의 갯수}}{N}
\end{aligned}
$$

은 참값 $\frac{\pi}{4}$의 불편 추정량(unbiased estimator)이 된다. 즉,

$$
E[\hat{\theta}] = \frac{\pi}{4}
$$

따라서, 추정량 $4\hat{\theta}$는 우리가 구하고자 하는 원주율 π의 불편추정량이 된다. □

우리는 Monte Carlo simulation 방법을 금융분야에 적용하고자 한다. 경제나 금융에 등장하는 상품들(주식,채권,등)과 이들로 부터 파생되는 상품들(option, futures, swap, cap, floor)은 모두 그 가격의 결정이 확률적인 구조에 의존한다. 이들의 가격을 Monte Carlo Simulation을 이용하여 구하는 기본적인 방법은 만기에서의 손익(payoff)에 대한 할인된 기대값을 구하는 것이다. 기대값을 계산하기 위해서는 기초자산(underlying asset)의 확률적인 구조를 알아야한다.

그러나 각각의 기초자산의 확률분포를 알아도 그 자산들 간의 결합된 확률구조나 provision 등이 복잡한 경우가 많다. 또한 이러한 구조를 가진 파생금융상품들은 가격결정을 위한 analytic formula를 가지고 있지 않는 경우가 많다. 이러한 경우 Monte Carlo Simulation 방법은 단순하면서도 유연한 하나의 대안이 될 수 있다.

또한 Monte Carlo 방법은 기초자산의 가격움직임이 나타내는 현실적인 문제인 jump를 반영할 수도 있고 path-dependent한 상품(Asian option)의 가격결정이 쉽다. 그리고 최근의 금융 계산은 어마어마한 계산량을 요한다. 이러한 경우에도 Monte Carlo Simulation 방법은 강력한 도구가 될 수 있다.

그리고, Monte Carlo Simulation으로 추정한 값은 Strong Law of Large Number에 의해 시행횟수가 증가함에 따라 참값에 수렴함을 보장한다. 그리고 여기서 문제가 될 수 있는 것은 수렴 속도인데 비교적 느린 Monte Carlo simulation의 수렴 속도를 향상시키는 여러가지 방법이 개발되어지고 있다.

16.1.2 Pros And Cons

- 장점
    1. 수학적 기반이 간단하다.
    2. 자산들간의 correlation에 대한 고려가 쉽다.
    3. 정확도를 높이기 위해서는 좀 더 많은(아주 많이?) 시행해 보면 된다.
    4. 모델을 바꾸기가 쉽다. 일반적으로 많이 사용하는 정규분포 뿐만 아니라, 비정규분포 모델에 대한 구현도 쉽게 할 수 있다.

[1] Independent Identically Distributed

5. 복잡한 sample path에 대한 의존관계가 비교적 쉽게 구현된다. 따라서, payoff가 자산의 path에 의존할 때도 가격을 어렵지 않게 구할 수 있다.

- 단점

1. analytic한 방법에 비해 계산 시간이 많이 소요된다.
2. American option과 같이 조기행사(early exercise)가 가능한 상품에 대한 계산이 쉽지 않다.
3. 추정량(estimator)에 대한 표준 오차(standard error)는 M번 시행한 경우 $\frac{1}{\sqrt{M}}$에 비례한다. 이에 관해서는 다음에 좀 더 자세히 다룰 것이다.

16.2 Implementing Monte Carlo Simulation

risk neutral 가정하에서 random sample을 생성하여 손익(payoff)에 대한 기대값으로 금융상품이나 포트폴리오의 가치를 계산한다.

16.2.1 Example: Black-Scholes Call Option Model

Monte Carlo Simulation의 예로, 현재가치 $S(0)$인 주식에 대한, 행사가격 K, 만기 T인 European call option을 생각해 보자. 물론, European call option의 가격을 Black-Scholes Model로 가격을 구할 경우에는 Black-Scholes Formula가 있기 때문에, 굳이 Monte Carlo simulation을 이용하지 않아도 된다. 하지만 이 예는 Monte Carlo simulation의 구현과정을 설명하기에 적합하기 때문에 선택되었다.

Step 1. **모델 결정:** 기초자산의 가치의 움직임에 대한 모델을 결정한다. 지금의 경우는 $S(t)$에 대한 모델로 Geometric Brownian Motion을 택하자. 즉,

$$dS_t = S_t(rdt + \sigma dW_t). \tag{16.2.1}$$

(16.2.1)은 시간의 흐름이 연속적임을 가정한 모델이다[2]. 이것을 이산적인 모델로 만들기 위해서 만기까지의 시간을 N 개로 등분하여 구간간격을 Δt로 만들면, (16.2.1)은 아래와 같다.

$$S(t + \Delta t) - S(t) = S(t)(r\Delta t + \sigma\sqrt{\Delta t}\ \epsilon), \tag{16.2.2}$$

여기서 ϵ은 표준정규분포를 따르는 확률변수이다. 이 경우에는Δt가 충분히 작아야(즉, N이 충분히 커야), (16.2.2)가 (16.2.1)의 근사식이 될 수 있다.

Remark 16.2.1. 대부분의 확률미분방정식(stochastic differential equation)은 명시적인 모양(explicit form)의 식으로 표현되지 않는다. 그러나, 다행히도 (16.2.1)은 명시적인(explicit) 식을 가진다.

$$S(t + \Delta t) = S(t) \exp\left(\left(r - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t}\ \epsilon\right). \tag{16.2.3}$$

따라서, 이러한 경우는 Δt가 작지 않아도 되며, (16.2.3)은 path에 의존하지 않는 European-type의 상품을 계산하는데에 사용가능하다.

[2](16.2.1)에서 drift항이 일반적인 항이 아니라 이자률 r인 것은 앞에서도 언급했듯이, risk-neutral 확률공간에서 가치를 구해야 하기 때문이다.

Step 2. **Sample Path 생성:**

$$\begin{aligned} S(0) &= \text{현재 가격},\ S(n) := S(0+n\Delta t),\ N\Delta t = \text{만기}(=T) \\ S(1) &= S(0)+S(0)(r\Delta t+\sigma\sqrt{\Delta t}\ \epsilon_1) \\ S(2) &= S(1)+S(1)(r\Delta t+\sigma\sqrt{\Delta t}\ \epsilon_2) \\ &\vdots \\ S(n) &= S(n-1)+S(n-1)(r\Delta t+\sigma\sqrt{\Delta t}\ \epsilon_n), \end{aligned}$$

여기서, ϵ_i들은 서로독립인 표준 정규분포이다.

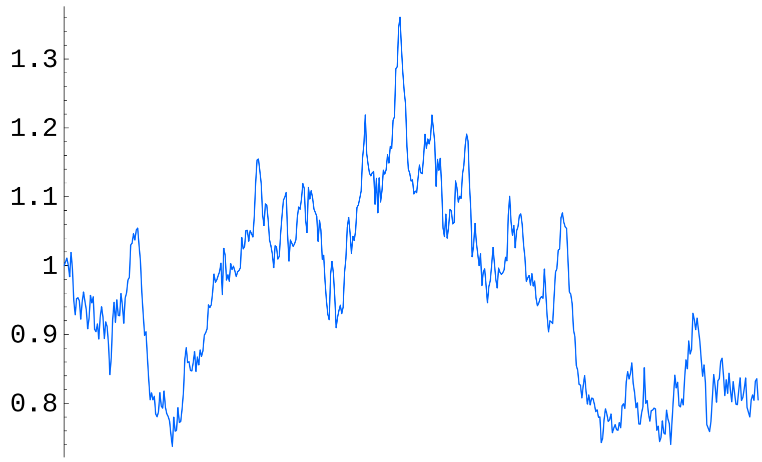


Figure 16.2: A sample path with geometric Brownian motion: S(0)=1,drift=10%, $\sigma = 40\%$

Step 3. **만기에서의 가치 계산:** Step 2에서 생성된 random sample에 대하여, 만기에서의 가치(sample payoff)를 결정한다. 행사가 K인 call option의 경우

$$\max\{S(T)-K,0\}.$$

Step 4. **반복,평균,할인:** Step 2, Step 3을 반복하여, 평균을 구한다. 이렇게 구한 평균을 이자율로 할인한다.

$$e^{-rT}\left[\frac{1}{N}\sum_{i=1}^{N}\max\{S_T^i-K,0\}\right]. \tag{16.2.4}$$

Remark 16.2.2 (Error Estimation). 한번의 sample path로부터 얻어낸, 만기에서의 가치를 할인한 것을 C_j라 할 때, 시뮬레이션 회수가 M이면,

$$\widehat{C} = \frac{1}{M}\sum_{j=1}^{M} C_j,$$

$\widehat{C}$는 참값 C의 추정치가 된다. strong law of large number에 의해 $\widehat{C}$는, 시뮬레이션 회수 M 이 증가함에 따라, 참값 C에 수렴한다. 추정치가 갖는 오차를 재는 측도 중에서, $\widehat{C}$의 표준편차(SD)를

표준오차 (standard error,SE)라 한다.

$$\begin{aligned} SD(C) &= \sqrt{\frac{1}{M-1}\sum_{j=1}^{M}(C_j - \widehat{C})^2}, \\ SE(\widehat{C}) &= \frac{SD(C)}{\sqrt{M}}. \end{aligned}$$

따라서, 2배 정확한 값을 얻기 위해서는 4배 많은 양의 simulation이 필요하다. ■

16.3 Examples

16.3.1 Delta Hedging Simulation

어떤 상품(instrument)을 Hedging 하는 방법에 대해 살펴보자.

$$\begin{aligned} r &= \text{이자율} \\ \Pi &= \text{Hedging 대상 상품} \\ S_i &= \text{기초자산} \\ & t_0, \quad t_i = t_0 + \Delta t \times i, \quad t_N = T = t_0 + \Delta t \times N, \\ & t_0(\text{현재}), \quad t_1, \quad t_2, \quad \cdots, \quad t_{N-1}, \quad t_N(\text{만기}) \end{aligned}$$

에 대하여, 각각의 t_i 시점에 $\Delta_i S_i$만큼을 돈을 빌려, Δ_i만큼의 기초자산을 매입한다($\Delta_i, -\Delta_i S_i$). 이렇게 구성한 포트폴리오를 t_{i+1} 시점에서 현금화하여, 만기까지 가져간다.

$$\begin{array}{cccccc} (\Delta_0, -\Delta_0 S_0) & \xrightarrow{t_1:\text{현금화}} & \Delta_0(S_1 - S_0 e^{r\Delta t}) & \xrightarrow{T} & \Delta_0(S_1 - S_0 e^{r\Delta t})e^{r(T-\Delta t)} \\ (\Delta_1, -\Delta_1 S_1) & \xrightarrow{t_2:\text{현금화}} & \Delta_1(S_2 - S_1 e^{r\Delta t}) & \xrightarrow{T} & \Delta_1(S_2 - S_1 e^{r\Delta t})e^{r(T-2\Delta t)} \\ & \vdots & & & \\ (\Delta_{N-1}, -\Delta_{N-1} S_{N-1}) & \xrightarrow{t_N = T:\text{현금화}} & \Delta_{N-1}(S_N - S_{N-1} e^{r\Delta t}) & & \end{array}$$

이것을 모두 합하면,

$$\sum_{i=0}^{N-1} \Delta_i(S_{i+1} - S_i e^{r\Delta t})e^{r(T-(i+1)\Delta t)}. \tag{16.3.1}$$

t_0에 Π_0을 가지고 시작했다면, 우리는 다음을 얻을 수 있다.

$$\begin{aligned} &\Pi_0 e^{rT} + \sum_{i=0}^{N-1} \Delta_i(S_{i+1} - S_i e^{r\Delta t})e^{r(T-(i+1)\Delta t)} = \Pi_T + \text{ hedging error}, \\ &\Pi_0 e^{rT} = \Pi_T - \sum_{i=0}^{N-1} \Delta_i(S_{i+1} - S_i e^{r\Delta t})e^{r(T-(i+1)\Delta t)} + \text{ hedging error}, \\ \therefore\ &\Pi_0 = e^{-rT}\Big[\Pi_T - \sum_{i=0}^{N-1} \Delta_i(S_{i+1} - S_i e^{r\Delta t})e^{r(T-(i+1)\Delta t)}\Big] + \text{ hedging error}. \end{aligned} \tag{16.3.2}$$

하나의 random sample로 부터, (16.3.2)를 얻었다. hedging error의 평균이 0이므로,M개의 random sample을 생성하여, (16.3.2)들의 평균을 구하면, 그것이 Π의 현재 가치가 된다.

16.3.2 Path-Dependent Asian Option

Asian option은 만기에서의 payoff가 정해진 기간 동안의 자산의 평균(S_{ave})–산술평균,또는 기하평균–에 의해 결정되는 option이다. Asian call, put option의 만기에서의 가치는 각각

$$\max(S_{\text{ave}} - K, 0), \quad \max(K - S_{\text{ave}}, 0) \tag{16.3.3}$$

이다. 하나의 sample path로 생성된 주가

$$S(0), S(1), S(2), \cdots, S(N)$$

에 대하여,

$$\begin{aligned} S_{\text{ave}} &= \frac{S(1) + S(2) + \cdots + S(N)}{N}, \quad \text{또는} \\ S_{\text{ave}} &= \{S(1)S(2) + \cdots + S(N)\}^{\frac{1}{N}} \end{aligned}$$

이다. 따라서 M개의 sample path를 생성했다면, asian option의 가격은

$$e^{-rT}\frac{1}{M}\sum_{j=1}^{M}\max(S^j_{\text{ave}} - K, 0), \; e^{-rT}\frac{1}{M}\sum_{j=1}^{M}\max(K - S^j_{\text{ave}}, 0)$$

이 된다.

16.3.3 Interest Futures

(단기)금리 선물의 경우, 기초 자산은 정기예금으로 이해 할 수 있고,가격은 미래의 예금기간에 적용되는 고정금리라 할 수 있다. 즉 금리 선물을 사는 것은 예금을 하는것과 동일하며, 반면에 금리 선물을 파는 것은 예금을 받는 것과 동일하다. **잔존 만기 91일 양도성정기예금증서(CD)**에 대한 선물이 그 예라 할 수 있다.
금리 선물의 가치를 결정하기 위해서는 이자율의 움직임에 대한 모델이 필요하다. 이자율 모델은 여러가지가 있지만, 비교적 간단한 Vasicek 모델을 예로 들자.

$$\begin{aligned} dr_t &= a(b - r_t)dt + \sigma dW_t \\ r_{t+1} - r_t &= a(b - r_t)\Delta t + \sigma\sqrt{\Delta t}\;\epsilon. \end{aligned}$$

r_t를 91일 CD 금리라 가정하며,

$$\begin{aligned} r(0) &= \text{현재 CD 금리}, \\ r(1) &= r(0) + a(b - r(0))\Delta t + \sigma\sqrt{\Delta t}\;\epsilon_1 \\ &\vdots \\ r(n) &= r(n-1) + a(b - r(n-1))\Delta t + \sigma\sqrt{\Delta t}\;\epsilon_n. \end{aligned}$$

현재 행사가 K로 만기 T인, CD금리 선물에 대한 1개의 매입 포지션을 가지고 있다면, 만기에서의 손익은

$$(K - r(n)) \times \frac{91}{364} \times 500,000,000 \tag{16.3.4}$$

이다.
만약 M개 random sample을 생성했다면, 금리선물의 가치는 (16.3.4)들의 평균으로, 다음과 같다.

$$\frac{\exp(-\overline{r}T)}{M}\sum_{j=1}^{M}(K-r(n)_j)\times\frac{91}{364}\times 500,000,000 \quad ,$$
$$\overline{r}=\frac{1}{M}\sum_{j=1}^{M}r(j).$$

16.3.4 Hedging Sensitivities

hedging sensitivity를 나타내는 delta, gamma, vega, theta, rho는 다음과 같이 finite difference ratio를 이용하여 계산한다.

$$\begin{aligned}
\text{delta} &= \frac{\partial V}{\partial S} \approx \frac{V(S+\Delta S)-V(S-\Delta S)}{2\Delta S}\\
\text{gamma} &= \frac{\partial^2 V}{\partial S^2} \approx \frac{V(S+\Delta S)-2V(S)+V(S-\Delta S)}{(\Delta S)^2}\\
\text{vega} &= \frac{\partial V}{\partial \sigma} \approx \frac{V(\sigma+\Delta\sigma)-V(\sigma-\Delta\sigma)}{2\Delta\sigma}\\
\text{theta} &= \frac{\partial V}{\partial t} \approx \frac{V(t+\Delta t)-V(t)}{\Delta t}\\
\text{rho} &= \frac{\partial V}{\partial r} \approx \frac{V(r+\Delta r)-V(r-\Delta r)}{2\Delta r}
\end{aligned}$$

예를 들어, delta가 위와 같은 모양의 근사식으로 표현되는 이유는 delta가 기초자산(S)에 대한 미분이므로, 다음과 같이 근사될 수 있기 때문이다.

$$\begin{aligned}
\text{delta} &= \lim_{\Delta S\to 0}\frac{V(S+\Delta S)-V(S-\Delta S)}{2\Delta S}\\
&\approx \frac{V(S+\Delta S)-V(S-\Delta S)}{2\Delta S}
\end{aligned}$$

따라서, Monte Carlo Simulation을 통해서 delta를 계산하고자 한다면, 작은 $\Delta S (e.g.\ \Delta S = 0.001S)$에 대하여, $V(S+\Delta S)$와 $V(S-\Delta S)$를 같은 random number에 대해서 계산하여야 한다. 그렇지 않으며, 오차를 상쇄시키는 효과도 없고, 또한 작은 값 ΔS로 나누므로 오차는 더욱 커진다.

Remark 16.3.1 (Common Random Number). For the estimator of delta

$$\frac{V(S+\epsilon)-V(S)}{\epsilon},$$

the decisive element for the variance of estimator is the variance of the numerator, which turns out to be equal to

$$\mathrm{Var}(V(S))+\mathrm{Var}(V(S+\epsilon))-2\mathrm{Cov}(V(S),V(S+\epsilon)).$$

Therefore, the estimator of the Greeks is all the more efficient that $V(S+\epsilon)$ and $V(S)$ are correlated since the covariance of the two terms is maximum for a perfect correlation of the two terms. This is why using common random numbers for the simulation of the two option prices: $V(S+\epsilon)$ and $V(S)$, is very efficient.

여기서, Monte Carlo simulation을 이용하여, delta를 구하는 효율적인 방법을 한 가지 더 소개하고자 한다. 예를 들어, European call option의 경우 delta는 다음과 같이 표현된다.

$$\begin{aligned}\text{delta} &= \frac{\partial C}{\partial S} = \frac{\partial}{\partial S}(e^{-rT}E[(S_T - K)\ 1_{S_T>K}]), \quad \text{여기서}\\ S_T &= S\exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\epsilon_T\right).\end{aligned}$$

따라서,

$$\text{delta} = e^{-rT}E[\exp((r - \frac{1}{2}\sigma^2)T + \sigma\epsilon_T)1_{\{S_T>K\}}] \tag{16.3.5}$$

이 된다[3]. (16.3.5)은 Monte Carlo Simulation을 통하여, 쉽게 구현될 수 있다. 즉, 만기에서의 손익이 다음과 같이 주어진 상품의 현재가치(할인가치)를 구하면 된다.

$$\begin{cases}\exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\epsilon_T\right), & S_T > K\\ 0, \quad S_T \le K.\end{cases}$$

그러나, 이러한 방법을 통해서, gamma를 구할 수는 없다[4]. 그렇지만, gamma는 (16.3.5)을 통해서 구한 delta를 이용하여, 다음의 식을 통해 구할 수 있다.

$$\text{gamma} \;=\; \frac{\partial^2 V}{\partial S^2} \approx \frac{delta(S + \Delta S) - delta(S - \Delta S)}{2\Delta S}.$$

16.4 Efficiency, Speeding Up & Variance Reduction Methods

구하고자 하는 값 θ가 있다고 하자. θ를 구하기 위해, random sample로 부터 $\{\hat{\theta}_i, i = 1, 2, ...\}$를 생성할 수 있다. 여기서, $\hat{\theta}_i$는 평균이 θ이고, 분산이 σ^2이라 하자. 그러면, n개의 sample로 부터 얻어지는, θ의 추정치는 표본평균($\tilde{\theta}$):

$$\tilde{\theta} = \frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_i$$

이다. 그러면, n이 큰 값이면, 중심극한정리에 의해, 표분평균은 근사적으로, 평균이 θ, 분산이 $\frac{\sigma^2}{n}$인, 정규분포를 따르다. 따라서,95% 신뢰도의 신뢰구간은

$$\tilde{\theta} - 1.96\frac{\sigma}{\sqrt{n}} < \theta < \tilde{\theta} + 1.96\frac{\sigma}{\sqrt{n}}.$$

따라서, 신뢰구간의 길이는 $\frac{1}{\sqrt{n}}$에 비례한다. 즉, 10배 정확한 추정을 하려면, 100배의 시행을 해야 한다.

이제, 두 가지의 추정방법이 있을 때, 효율성 측면에 대해 살펴보자. 다음의 θ에 대한, 두가지 추정량이 있다고 하자.

$$\hat{\theta}^{(1)} = \{\hat{\theta}_i^{(1)}, i = 1, 2, \cdots\}, \quad \hat{\theta}^{(2)} = \{\hat{\theta}_i^{(2)}, i = 1, 2, \cdots\}.$$

두 가지의 추정량 모두 불편추정량(unbiased)이라 하자. 즉,

$$E[\hat{\theta}_i^{(1)}] = E[\hat{\theta}_i^{(2)}] = \theta,\ Var[\hat{\theta}_i^{(1)}] = \sigma_1^2,\ Var[\hat{\theta}_i^{(2)}] = \sigma_2^2$$

[3] $1_{\{S_T>K\}}$는 $S_T > K$이면 1이고, 아니면 0인 함수이다.

[4] (16.3.5)을 미분하면, Dirac delta 함수가 나오는데, 이는 Monte Carlo Simulation을 통해서 구현하기가 어렵다.

만약, $\sigma_1 < \sigma_2$ 라 하면, 같은 n개의 sample로 부터 얻어지는 $\hat{\theta}^{(1)}, \hat{\theta}^{(2)}$에 대하여, $\hat{\theta}^{(1)}$이 더 나은 추정량라 할 수 있을까?

추정량 계산을 위한 sample을 생성하는데, 소요되는 시간에 대한 고려가 없는 상태에서, 분산이 작다고, 더 나은 추정량라고 하기는 적절치 못 하다.

이제, 하나의 sample을 생성하는데, 소요되는 시간을 각각, b_1, b_2라 하자. 주어진 t 시간에 생성할 수 있는 sample의 갯수는 각각, $\frac{t}{b_1}, \frac{t}{b_2}$ 이다. 따라서, 시간 t에 얻을 수 있는 θ의 추정량는 각각,

$$\frac{b_1}{t}\sum_{i=1}^{t/b_1}\hat{\theta}_i^{(1)}, \; \frac{b_2}{t}\sum_{i=1}^{t/b_2}\hat{\theta}_i^{(2)}$$

이다. 충분히 큰 시간 t에 대해서, 추정량은 근사적으로, 평균이 θ, 표준편차가 각각

$$\sigma_1\sqrt{\frac{b_1}{t}}, \; \sigma_2\sqrt{\frac{b_2}{t}}$$

인 정규분포를 따른다. 따라서,

$$\sigma_1^2 b_1 < \sigma_2^2 b_2 \tag{16.4.1}$$

이 만족된다면, 추정량 $\hat{\theta}^{(1)}$이 $\hat{\theta}^{(2)}$보다 더 나은 추정량이라 할 수 있다.

이제, Monte Carlo Simulation의 효율성을 높이기 위한 방법에 대해서 살펴보자. 효율성이 높다는 것은 simulation의 횟수가 적어도, 표준오차(standard error)가 작다는 것이다. 우리는 여기서, 다음과 같은 variance reduction technique에 대해서 알아보자.

- Antithetic Variates
- Control Variates
- Moments[5] Matching Methods

16.4.1 Antithetic Variates

Antithetic Variates는 Variance Reduction 방법 중 가장 간단하고, 많이 사용되는 방법 중에 하나다. European call option을 통해서, Antithetic Variates 방법을 설명하고자 한다. Black-Scholes option model에서는, 주식의 가치가 log-normal process를 따른다는 가정하에, risk-neutral measure하에서, 다음과 같이 생성된다.

$$S_T^{(i)} = S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z_i\right), \quad i = 1, \cdots, n.$$

n개의 random sample로 부터, 얻어지는 행사가 K인, call option 가치의 불편추정량은

$$\hat{C} = \frac{1}{n}\sum_{i=1}^{n} C_i = \frac{1}{n}\sum_{i=1}^{n} e^{-rT}\max(0, S_T^{(i)} - K)$$

이다.
Antithetic Variates 방법은 Z_i가 표준정규분포를 따르고, $-Z_i$도 표준정규분포를 따른다는 사실에 주목한다. 즉,

$$Z_i \sim N(0,1), \quad -Z_i \sim N(0,1).$$

[5] 확률변수 X에 대하여, $E[X], E[X^2], E[X^3], \cdots$ 등을 Moment라 한다.

이로부터, 주식가치와 option 가치의 불편 추정량을 새롭게 얻을 수 있다.

$$\tilde{S}_T^{(i)} = S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}(-Z_i)\right), \quad i = 1, \cdots, n.$$

$$\hat{C} = \frac{1}{n}\sum_{i=1}^{n} \tilde{C}_i = \frac{1}{n}\sum_{i=1}^{n} e^{-rT}\max(0, \tilde{S}_T^{(i)} - K)$$

따라서, Antithetic Variates 방법을 통한 option 가치 $\hat{C}_{AV}$를 다음과 같이 얻는다.

$$\hat{C}_{AV} = \frac{1}{n}\sum_{i=1}^{n} \frac{C_i + \tilde{C}_i}{2}.$$

직관적인 측면에서 봐도, n개의 $(Z_i, -Z_i)$는 $2n$개의 random sample보다 더 고른 분포를 가진다는 것을 알 수 있다. 또한, $2n$개의 표준정규분포로 부터의 random sample은 평균이 0이 아니지만, n개의 $(Z_i, -Z_i)$는 평균이 0임이 보장된다. $\{Z_i\}$가 고르게 잘 분포되어 있으면, $(Z_i, -Z_i)$는 더욱 더 고르게 분포되어 있는 것을 기대할 수 있다.

이제, 좀 더 엄밀하게 Antithetic Variate 방법의 효율성(Variance Reduction)에 대해서 살펴보자.
$C_i, \tilde{C}_i$가 같은 분산을 가지므로,

$$\mathrm{Var}\left[\frac{C_i + \tilde{C}_i}{2}\right] = \frac{1}{2}\left(\mathrm{Var}[C_i] + \mathrm{Cov}[C_i, \tilde{C}_i]\right).$$

그런데, $C_i, \tilde{C}_i$를 계산하기 위해서는 C_i만 계산 할 때 보다 2배의 계산 시간이 소요된다. 따라서,(16.4.1)에 의하여, Antithetic Variates가 보다 나은 추정량이 되기 위해서는

$$\mathrm{Var}[\hat{C}_{AV}] \leq \mathrm{Var}[\hat{C}]$$

이 만족되어야 한다. 이는

$$\mathrm{Cov}[C_i, \tilde{C}_i] \leq 0 \tag{16.4.2}$$

이 만족되면 된다. 그런데, European option의 경우, (16.4.2)이 만족된다 [6].
이와 같이, 음의 상관관계를 갖는 가상적인 자산을 생성하는 방법을 *antithetic variance reduction*이라하고, 생성된 음의 상관관계를 갖는 자산을 *antithetic variate*라 한다.

16.4.2 Control Variates Technique

Control Variates 방법은 Antithetic Variates 방법과 더불어, 가장 많이 사용되는 variance reduction technique 이다. P_A를 우리가 구하고자 하는 값이라 하자. P_G는 우리가 그 값을 아는 것이라 하자 [7]. $\hat{P}_A, \hat{P}_G$를 각각, P_A, P_G의 sample path를 통한 불편추정량이라 하자. 즉,

$$P_A = E[\hat{P}_A], \quad P_G = E[\hat{P}_G].$$

그러면,

$$P_A = P_G + E[\hat{P}_A - \hat{P}_G]$$

[6] 단조함수 ϕ에 대하여, $C_i = \phi(Z_i)$로 표현된다. 이 경우 $Cov[Z_i, -Z_i] = -1$ 이므로, $Cov[\phi(Z_i), \phi(-Z_i)] \leq 0$이 된다. barrier option의 경우, 단조함수 ϕ에 관한 식으로 표현 되지 않으므로, Antithetic Variate 방법이 효율적이지 못하다.

[7] 예를 들면, P_A는 Asian option에서 그 평균을 산술평균으로 하는 것이라 할 수 있고, P_G는 평균을 기하평균으로 하는 것이라 할 수 있다. P_G에 대해서는 option의 가치를 구하는 공식이 알려져있지만, P_A는 그렇지 못하다.

가 성립된다. 따라서, P_A의 새로운 불편추정량 $\hat{P}_A^{cv}$를 얻을 수 있다.

$$\hat{P}_A^{cv} = \hat{P}_A + (P_G - \hat{P}_G). \tag{16.4.3}$$

이 때, $P_G - \hat{P}_G$를 $control$이라 한다.

이제, $\hat{P}_A^{cv}$가 $\hat{P}_A$보다 더 나은 추정량이 되기 위한 조건에 대하여 살펴보자.

$$\mathrm{Var}[\hat{P}_A^{cv}] = \mathrm{Var}[\hat{P}_A] + \mathrm{Var}[\hat{P}_G] - 2\mathrm{Cov}[\hat{P}_A, \hat{P}_G]$$

이므로, $\hat{P}_A$와 $\hat{P}_G$의 공분산이 크면 클수록, $\hat{P}_A^{cv}$는 더 효율적인 추정량이 된다.
여기서 (16.4.3)보다 좀 더 나은 새로운 추정량 $\hat{P}_A^{\beta}$을 생각해 보자.

$$\hat{P}_A^{\beta} = \hat{P}_A + \beta(P_G - \hat{P}_G). \tag{16.4.4}$$

그러면,

$$\mathrm{Var}[\hat{P}_A^{\beta}] = \mathrm{Var}[\hat{P}_A] + \beta^2\mathrm{Var}[\hat{P}_G] - 2\beta\mathrm{Cov}[\hat{P}_A, \hat{P}_G]$$

임을 알 수 있고, $\mathrm{Var}[\hat{P}_A^{\beta}]$가 최소가 되게 하는 β는 다음과 같다.

$$\beta^* = \frac{\mathrm{Cov}[\hat{P}_A, \hat{P}_G]}{\mathrm{Var}[\hat{P}_G]}.$$

이 때,

$$Var[\hat{P}_A^{\beta^*}] \le \mathrm{Var}[\hat{P}_A]$$

이 된다. 따라서, (16.4.3)은 분산(표준오차)를 줄일 수도 있고 증가시킬 수도 있지만, (16.4.4)는, $\mathrm{Cov}[\hat{P}_A, \hat{P}_G]$가 0만 아니면,

$$\mathrm{Var}[\hat{P}_A^{\beta^*}] < \mathrm{Var}[\hat{P}_A]$$

이 되어, $\hat{P}_A$보다 효과적인 추정치가 된다.
이제, β^*의 추정에 대해서 살펴보자. 만약, n개의 random sample $\{(P_{Ai}, P_{Gi})\}$로 부터, β^*의 추정량 $\hat{\beta}$를 구하고, 이로 부터 구한

$$\frac{1}{n}\sum_{i=1}^{n} P_{Ai} + \hat{\beta}\big(P_G - \frac{1}{n}\sum_{i=1}^{n} P_{Gi}\big) \tag{16.4.5}$$

는 $\hat{\beta}$와 P_{Gi}의 의존관계 때문에, 불편추정량이 되지 못 한다. 따라서, n개의 random sample중 일부인 n_1개를 β^*의 추정에 사용하고, 나머지 $n - n_1$개를 (16.4.5)의 추정에 사용한다면, (16.4.5)는 불편추정량이 된다 ($n_1 \ll n$).
그러나, Control Variates 방법에서, β^*의 추정은 아주 정확하지 않아도 분산을 줄이는데 큰 문제가 되지 않는다.

16.4.3 Moments Matching Methods

$\{Z_i, i = 1, 2, \cdots, n\}$을 표준정규분포를 따르는 random sample이라고 하면, 이것들의 표본평균 $\tilde{Z} = \sum_{i=1}^{n} Z_i$은 0이 아닐 것이 확실하다.

$$\tilde{Z}_i = Z_i - \tilde{Z}, \quad i = 1, 2, \cdots, n \tag{16.4.6}$$

로,$\tilde{Z}_i$를 정의하면, $\{\tilde{Z}_i, i = 1, 2, \cdots, n\}$의 표본평균은 0이 된다. 여기서 주의해서 볼 점은,

$$\{\tilde{Z}_i, i = 1, 2, \cdots, n\}$$

이 표준정규분포를 따르지만, 서로 독립이 아니라는 점이다. 이러한 이유로, Moments Matching 방법에서, 표준오차를 구하고, 이로부터 신뢰구간을 구하는 것이 쉽지 않다.
(16.4.6)는 첫번째 Moment(평균)을 일치시키는 방법을 보여준다. 이제 좀 더 일반적으로, 평균, 표준편차가 각각, μ, σ인 경우 다음과 같이 변환시킬 수 있다.

$$\begin{aligned}\tilde{Z}_i &= (Z_i - \tilde{Z})\frac{\sigma}{\tilde{s}} + \mu, \quad \text{여기서}, \\ \tilde{s} &= \frac{1}{n-1}\sum_{i=1}^{n}(Z_i - \tilde{Z})^2.\end{aligned} \tag{16.4.7}$$

그런데, (16.4.8)을 통한 변환 후, $\tilde{Z}_i$는 더 이상, 정규분포가 아니다. 이로부터 구한 option의 가치는 불편추정량이 아니다. 이것은 대부분의 경우 크게 문제 될 것이 없지만, 어떤 경우는 심각한 문제를 야기할 수도 있다.

Moment matching의 또 다른 응용으로, Geometric Brownian motion에 적용된 예를 보자. 현재의 주가가 S_0일 때 T시점의 주가 S_T의 평균,분산은 각각,

$$\begin{aligned}E[S_T] &= \mu_{S_T} = S_0 e^{rT} \\ Var[S_T] &= \sigma^2_{S_T} = S_0^2\sqrt{e^{2rT}(e^{\sigma^2 T} - 1)}\end{aligned}$$

이다. 생성된 $\{S_T(i), i = 1, 2, \cdots, n\}$을 다음과 같이 변환할 수 있다.

$$\tilde{S}_T(i) = (S_T(i) - \tilde{S}_T)\frac{\sigma_{S_T}}{s_{S_T}} + \mu_{S_T}.$$

여기서,

$$\begin{aligned}\tilde{S}_T &= \frac{1}{n}\sum_{i=1}^{n} S_T(i), \\ s_{S_T} &= \frac{1}{n-1}\sum_{i=1}^{n}(S_T(i) - \tilde{S}_T)^2.\end{aligned}$$

16.5 Multivariate Monte Carlo

Monte Carlo Simulation은 European-style의 상품이 여러개의 기초자산에 의존하여, 그 가치가 결정될 때, 사용할 수 있는 자연스러운 방법이다. 만약 option의 만기에서의 손익이 $S_1, S_2, \cdots, S_n$의 함수로 표시된다고 하자. 그러면, 이론적으로는 $n+1$차원의 편미분 방정식을 풀면 된다. 이것을 푸는 것은 많은 시간을 요하게 되지만, Monte Carlo Simulation을 이용하면, 쉽게 해결될 수 있다. 또는 여러개의 상품으로 구성된 포트폴리오의 대한 것도 Multivariate Monte Carlo를 이용해서 해결할 수 있다. 포트폴리오의 Value at Risk를 구할 때도 Monte Carlo simulation은 사용된다.

$$\begin{aligned}S_i(t + \Delta t) &= S_i(t)\exp\left(\left(r - \frac{\sigma_i^2}{2}\right)\Delta t + \sigma_i\sqrt{\Delta t}\ \epsilon_i\right), \\ E[\epsilon_i\epsilon_j] &= \rho_{ij}.\end{aligned} \tag{16.5.1}$$

그러면, 어떻게 (16.5.1)를 만족하는 정규분포를 따르는 random sample $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$를 생성할 것인가? 이 문제를 해결해 주는 방법이 Cholesky Decomposition이다.

16.5.1 Cholesky Decomposition

다음과 같이 covariance(correlation) matrix Σ가 주어져 있다고 하자.

$$\Sigma = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ & & \ddots & \\ s_{n1} & s_{n2} & \cdots & s_{nn} \end{pmatrix}$$

Σ는 covariance(correlation) matrix 이므로 대칭행렬이다($s_{ij} = s_{ji}$). 우리의 목적은 covariance(correlation) matrix가 Σ인 정규분포를 따르는 random sample을 생성하는 것이다. 일단, 서로 독립인, 즉 covariance(correlation) matrix가 단위행렬인 정규분포를 따르는 random sample X를 생성하는 것은 쉽다. 만약 Σ가

$$\Sigma = AA^T \tag{16.5.2}$$

으로 표현된다면, $\underline{AX}$ 가 우리가 원하는 random sample이 된다.

$$\begin{aligned} \text{Variance}(AX) &= A\text{Variance}(X)A^T \\ &= AIA^T \\ &= AA^T \\ &= \Sigma. \end{aligned}$$

(16.5.2)이 되도록 Σ를 만들수 있는 방법 중에 하나가 Cholesky Decomposition 이다. 주목할 점은 Σ가 어떤 조건을 만족한다면, Cholesky Decomposition은 행렬 A를 하삼각행렬로 만들어 준다.

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ & & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

로 두고 (16.5.2)을 풀면 다음을 쉽게 얻을 수 있다.

$$\begin{aligned} a_{ii} &= \sqrt{s_{ii} - \sum_{k=1}^{i-1} a_{ik}^2}, \quad i = 1, 2, \cdots, n, \\ a_{ij} &= \frac{1}{a_{jj}}\Big(s_{ij} - \sum_{k=1}^{j-1} a_{ik}a_{jk}\Big), \quad \text{for each } i, j = 1, 2, \cdots, i-1. \\ a_{11} &\to a_{21} \to a_{22} \to a_{31} \to \cdots \to a_{nn-1} \to a_{nn}. \end{aligned}$$

Example 16.5.1. For a 2×2 symmetric matrix, we have

$$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} 1 & \rho \\ 0 & \sqrt{1-\rho^2} \end{pmatrix}.$$

Also for a 3×3 matrix, we have

$$\begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ \rho_{12} & \sqrt{1-\rho_{12}^2} & 0 \\ \rho_{13} & \frac{\rho_{23}-\rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}} & \sqrt{1-\rho_{23}^2-\frac{(\rho_{23}-\rho_{12}\rho_{13})^2}{1-\rho_{12}^2}} \end{pmatrix} \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ 0 & \sqrt{1-\rho_{12}^2} & \frac{\rho_{23}-\rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}} \\ 0 & 0 & \sqrt{1-\rho_{23}^2-\frac{(\rho_{23}-\rho_{12}\rho_{13})^2}{1-\rho_{12}^2}} \end{pmatrix}.$$

□

16.5.2 Example: Value at Risk

다음과 같이 구성된 포트폴리에에 대하여, 5-days Value at Risk를 구하고자 한다.

P_B: 만기 1년, \$1,000,000 US-Zero-Bond
P_O: 만기 1달, \$1,000,000에 대한 put option, 행사가격=1,300원.

Value at Risk 계산에 필요한 parameter가 다음과 같이 주어져 있다고 하자.

P_{FX}: 환율(Won/USD): 현재-1,300원, 수익율-1-day vol=0.15%
채권(USD): 현재 이자율(r_f)-5%, 수익율-1-day vol=0.07%
ρ=-0.07 (환율-채권 수익율간의 correlation), 이자율(won,r_D)-8%.

표준 정규 분포를 따르고, correlation이 ρ인 확률 변수 Z_B, Z_{FX}에 대하여,

$$\begin{aligned} P_{FX} &= 1300e^{0.0015\times\sqrt{5}\times Z_{FX}} \\ P_B &= 951,229e^{0.0007\times\sqrt{5}\times Z_B}. \end{aligned}$$

이것으로 부터, Black-Scholes option 공식을 사용하여 계산하면, 5일 후 포트폴리오의 가치는 다음과 같다.

$$P_B \times P_{FX} + 1,000,000 \times BS(P_{FX}, 1300, \frac{1}{12} - \frac{5}{365}, r_D, r_f, 0.0015 \times \sqrt{365}). \quad (16.5.3)$$

이젠, random sample (Z_{FX}, Z_B)을 충분히 생성하여,(16.5.3)을 계산하면, 우리가 원하는 Value at Risk를 계산할 수 있다.

16.6 Generating Distributions

16.6.1 Inverse Transform Method

✎ **Continuous Distribution**
Suppose we are given the distribution function $F(x) = P(X \leq x)$ and that we want to generate random variable according to F. If we are able to invert F easily, we may apply the following inverse transform method.

1. Draw a uniform random number between 0 and 1. Denote uniform random variable by $U \sim U(0, 1)$.
2. Return $X = F^{-1}(U)$.

It is easy to see that the random variable X generated by this method is actually characterized by the distribution function F.

$$P(X \leq x) = P\Big(F^{-1}(U) \leq x\Big) = P\Big(U \leq F(x)\Big) = F(x).$$

Example 16.6.1 (Conditional Distributions). Suppose X has distribution F, and consider the problem of sampling X conditional on $a < X \leq b$, with $F(a) < F(b)$. Using the inverse transform method, there is no more difficult than generating X unconditionally. If $U \sim U(0, 1)$. then the random variable V defined by

$$V = F(a) + \big(F(b) - F(a)\big)U$$

is uniformly distributed. To see this, observe that

$$\begin{aligned} P(F^{-1}(V) \leq x) &= P\Big(F(a) + (F(b) - F(a))U \leq F(a)\Big) \\ &= P\left(U \leq \frac{F(x) - F(a)}{F(b) - F(a)}\right) \\ &= \frac{F(x) - F(a)}{F(b) - F(a)} \end{aligned}$$

and this is precisely the distribution of X given $a < X \leq b$.

✎ **Discrete Distribution**
Consider a discrete distribution with a finite support:

$$P(X = x_i) = p_i, \quad i = 1, 2, \cdots, n.$$

The following method gives us the discrete random variable X.

1. First, generate a uniform random variable U.

2. Return X as

$$X = \begin{cases} x_1 & \text{if } U < p_1, \\ x_2 & \text{if } p_1 \leq U < p_1 + p_2, \\ \vdots & \\ x_i & \text{if } \sum_{k=1}^{i-1} p_k \leq U < \sum_{k=1}^{i} p_k \\ \vdots & \end{cases}$$

16.6.2 Acceptance-Rejection Method

Suppose we must generate random variables according to a probability density function $f(x)$, and that the difficulty in inverting the corresponding distribution function makes the inverse transform method unattractive. Thus acceptance-rejection method gives us a alternative. The rejection method does not require that the cumulative distribution function be readily computable, much less the inverse of that function.

1. Assume that we know a function $t(x)$ such that

$$t(x) \geq f(x), \quad \forall x \in I,$$

   where I is the support of f. The function $t(x)$ is necessary to be a probability density function. If the support I is bounded, a typical choice for $r(x)$ is simply the uniform distribution on I, and we may choose

$$t(x) = \max_{x \in I} f(x).$$

2. Let

$$c := \int_I t(x) dx.$$

   Then

$$r(x) := \frac{t(x)}{c}$$

   is a probability density function. If the distribution $r(x)$ is easy to generate, the following acceptance-rejection method generates a random variable X distributed according to the density f:

   (a) Generate $R \sim r$.

   (b) Generate $U \sim U(0, 1)$, independent of R.

   (c) If

$$U \leq \frac{f(R)}{t(R)}$$

   return $X = Y$; otherwise, repeat the procedure.

Note that it can be shown that the average number of iterations to generate one random sample is c.

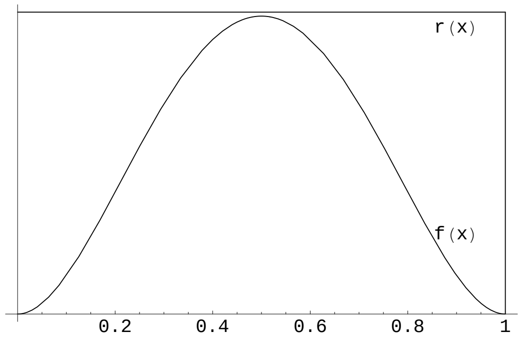


Figure 16.3: Graphical example of the acceptance-rejection method

16.6.3 Generating Normal Distribution

✎ **Using The Central Limit Theorem**
Let $X_i, (i = 1, \cdots, n)$ be IID random variables which are uniformly distributed between 0 and 1. For each i, We have

$$\begin{aligned} E[X_i] &= \frac{1}{2}, \\ \mathrm{Var}[X_i] &= \frac{1}{12}, \\ \mathrm{Kutosis}[X_i] &= \frac{1}{80}. \end{aligned}$$

By the central limit theorem,

$$\frac{\sum_{i=1}^{n} X_i - n \cdot \mu}{\sqrt{n}\sigma} = \frac{\sum_{i=1}^{n} X_i - n \cdot \frac{1}{2}}{\sqrt{\frac{n}{12}}}$$

approximately is normally distributed. i.e. $N(0, 1)$. If $n = 12$ we get

$$Y := \sum_{i=1}^{12} X_i - 6$$

which approximately normally distributed. Note that the standard normally distributed random variable has kutosis 3, but Y has kutosis $\frac{1}{80} \times 12 = 0.15$, which means that too many values close to the mean will be generated.

✎ **Box-Muller Method**
Let Y_1, Y_2 be a random sample from uniform distribution over $0 < y < 1$. Define X_1 and X_2 by

$$\begin{aligned} X_1 &= \sqrt{-2 \log Y_1} \cos(2\pi Y_2), \\ X_2 &= \sqrt{-2 \log Y_1} \sin(2\pi Y_2). \end{aligned}$$

The corresponding transformation is one-to-one and maps

$$\{(y_1, y_2) : 0 < y_1 < 1, 0 < y_2 < 1\}$$


```
while(X > 1)
      generate U_1, U_2 ~ U(0,1)
      U_1 <- 2U_1 - 1, U_2 <- 2U_2 - 1
      X = U_1^2 + U_2^2
end
R <- sqrt(-2 log X)
X_1 <- (U_1/sqrt(X)) R = U_1 sqrt(-2 log(X)/X),
X_2 <- (U_2/sqrt(X)) R = U_2 sqrt(-2 log(X)/X)
return X_1, X_2
```

Figure 16.4: A pseudo code for polar rejection method

onto

$$\{(x_1, x_2) : -\infty < x_1 < \infty, -\infty < x_2 < \infty\} - \{(0,0)\}.$$

The inverse transformation is given by

$$\begin{aligned} y_1 &= \exp\left(-\frac{x_1^2 + x_2^2}{2}\right), \\ y_2 &= \frac{1}{2\pi} \arctan\left(\frac{x_2}{x_1}\right). \end{aligned}$$

This has the Jacobian determinant

$$\begin{aligned} J &= \begin{vmatrix} -x_1 \exp\left(-\frac{x_1^2+x_2^2}{2}\right) & -x_2 \exp\left(-\frac{x_1^2+x_2^2}{2}\right) \\ -\frac{1}{2\pi}\frac{x_2}{x_1^2+x_2^2} & \frac{1}{2\pi}\frac{x_1}{x_1^2+x_2^2} \end{vmatrix} \\ &= \frac{1}{2\pi} \exp\left(-\frac{x_1^2 + x_2^2}{2}\right). \end{aligned}$$

Hence the joint distribution of X_1 and X_2 is bivariate normal distribution with correlation 0(i.e. independent).

Remark 16.6.2. $R := X_1^2 + X_2^2 = -2 \log Y_1$ is exponentially distributed with mean 2, i.e.

$$P(R \leq t) = 1 - e^{-\frac{t}{2}}.$$

$V := 2\pi Y_2$ is a random angle uniformly between 0 and 2π.

✎ Polar Rejection Method

Marsaglia and Bray developed a modification of the Box-Muller method that reduces computing time by avoiding evaluation of the sine and cosine functions. Figure (16.4) demonstrates a pseudo code for polar rejection method.

1. Note that conditional on acceptance, X is uniformly distributed between 0 and 1. Let $D_r = \{(u_1, u_2) | u_1^2 + u_2^2 \leq r\},\ 0 \leq r \leq 1$.

$$P(X \leq r) = \iint_{D_r} \frac{1}{\pi}\, du_1 du_2$$

$$
\begin{aligned}
&= \int_0^{2\pi} \int_0^{\sqrt{r}} \frac{r}{\pi}\, dr d\theta \\
&= r.
\end{aligned}
$$

This shows that X is uniformly distributed. Thus R has the same effect as the Box-Muller method.

2. $\frac{U_1}{\sqrt{X}}$ and $\frac{U_2}{\sqrt{X}}$ are independent of X conditional on $X \leq 1$.

Remark 16.6.3. Polar rejection method is inapplicable with quasi-Monte Carlo simulation.

16.7 Quasi Monte Carlo Simulation

Sequences of n-tuples that fill n-space more uniformly than uncorrelated random points are called quasi-random sequences.

Assume that we want to generate a sequence of N random vectors.

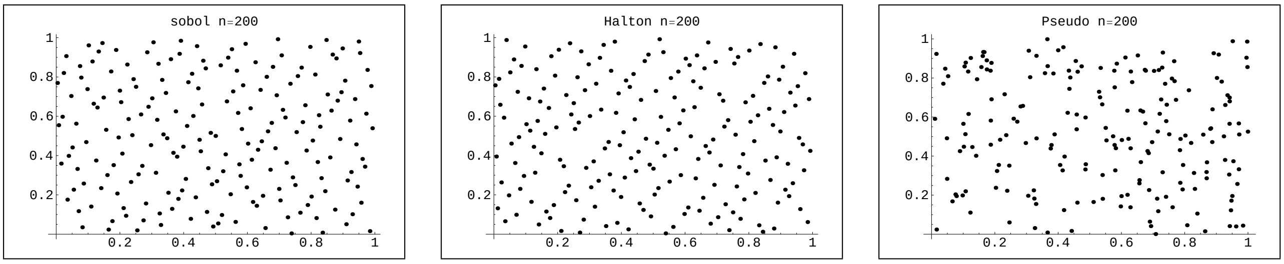
Definition 16.7.1 (Low Discrepancy). Given vector $\mathbf{x} = (x_1, x_2, \cdots, x_m)$, the rectangular subset $G_{\mathbf{x}}$ defined as

$$G_{\mathbf{x}} = [0, x_1) \times [0, x_2) \times \cdots \times [0, x_m),$$

which has a volume $x_1 x_2 \cdots x_m$. If we denote by $S(G)$ the function counting the number of points in the sequence, which are contained in a subset $G \subset \mathbf{I}^m$, the discrepancy is defined as

$$\sup_{X \in \mathbf{I}^m} |S(G_{\mathbf{x}}) - N x_1 x_2 \cdots x_m|.$$

□



(a) Sobol Sequence (b) Halton Sequence (c) Pseudo Random Sequence

Figure 16.5: 200 pairs 2D-random numbers

Figure (16.5) demonstrate 200 pairs 2-dimensional Sobol, Halton and pseudo random numbers.

16.7.1 Halton Sequence

✎ **Construction of Halton sequence**

Let us introduce notations. For integer n and base b, n has a unique b-ary expansion. We can write this as

$$n = \sum_{j=0}^{\infty} a_j(n) b^j,$$

with all but finitely many of the coefficients $a_j(n)$ equal to zero. Define radical inversion function $\psi_b : \mathbb{Z}^+ \to [0, 1)$ as follows.

$$\psi_b(n) = \sum_{j=0}^{\infty} \frac{a_j(n)}{b^{j+1}}$$

Halton's low-discrepancy sequences are given by the following algorithm:

1. Representing a positive integer number n in a base b, where b is a prime number:

$$n = (\cdots d_4 d_3 d_2 d_1 d_0)_b = \sum_{j=0}^{m_n} d_j b^j$$

   where $d_j = a_j(n)$.

2. Reflecting the digits and adding a radix point to obtain a number within the unit interval:

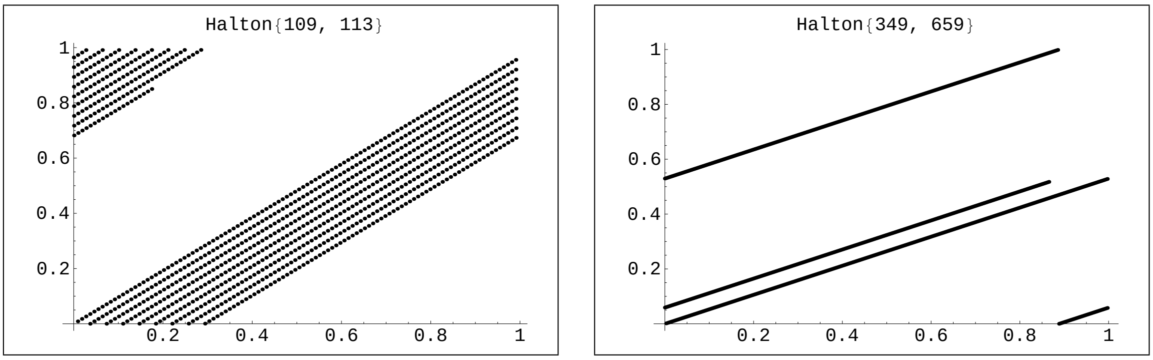
$$h(n, b) = (0.d_1 d_2 d_3 d_4 \cdots)_b = \sum_{j=0}^{m_n} \frac{d_j}{b^{j+1}} = \psi_b(n).$$

3. To get a sequence of d-tuples in d-space, you make each component a Halton sequence with a different prime base b. Typically, the first d primes are used. For example, nth Halton sequence can be given by

$$\mathbf{x}_n = \Big(\psi_{b_1}(n), \psi_{b_2}(n), \cdots, \psi_{b_d}(n)\Big),$$

   where $b_1, b_2, \cdots, b_d$ are primes.

✎ **Drawback of Halton sequence** If the base is large, the sequence produces long monotone segments, and projections of Halton sequence onto coordinates using large bases will have long diagonal segments in the projection hypercube. Figure (16.6) demonstrates this phenomenon.



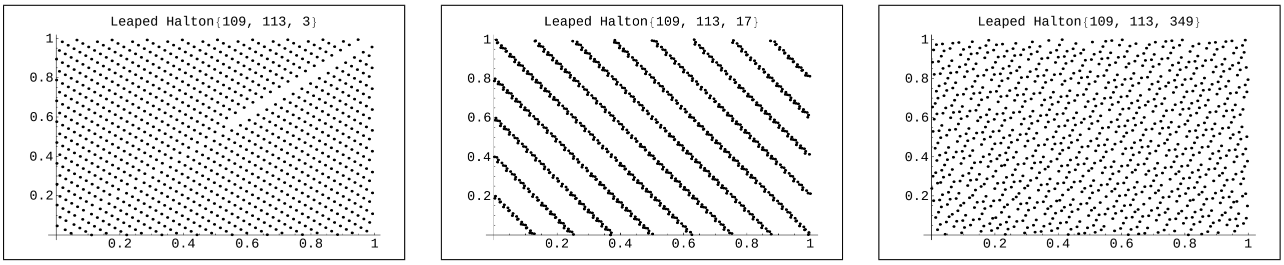
(a) bases 109 and 113 (b) bases 349 and 659

Figure 16.6: 1000 points of the Halton sequence

✎ **Leaped Halton sequence**
A possible remedy for the problem illustrated in Figure (16.6) is a leaped Halton sequence:

$$\mathbf{x}_n = \Big(\psi_{b_1}(ln), \psi_{b_2}(ln), \cdots, \psi_{b_d}(ln)\Big), \quad n = 0, 1, \cdots,$$

for some integer $l \geq 2$. A recommended choice for l is relative prime to the base $b_1, b_2, \cdots, b_d$. Figure (16.7) suggest that leaping can indeed improve uniformity, but its effect is very sensitive to the choice of leap parameter l.



(a) $l = 3$ (b) $l = 17$ (c) $l = 349$

Figure 16.7: First 10000 points of leaped Halton sequence with bases 109 and 113

16.7.2 Faure Sequence

We noted in the previous subsection that the uniformity of Halton sequences degrades in higher dimensions. In particular, the dth coordinate use a base at least as large as the dth prime, and this grows superexponentially with d.

Faure developed a different method in which all coordinates use a common base. This base must be at least as large as the dimension itself, but can be smaller than the largest base used for a Halton sequence of equal dimension.

In a d-dimensional Faure sequence, for the base b we choose the smallest prime number greater than or equal to d. To generate nth quasi-random vector, we can write n as b-ary expansion. i.e.

$$n = \sum_{j=0}^{\infty} a_j(n) b^j = \sum_{j=0}^{r_n - 1} a_j(n) b^j,$$

where r_n is the smallest integer such that

$$a_{r_n - 1}(n) \neq 0, \quad a_j(n) = 0 \text{ for all } j \geq r_n.$$

The ith coordinate, $i = 1, \cdots, d$, of the nth point in the Faure sequence is given by

$$\sum_{k=1}^{\infty} \frac{y_k^{(i)}(n)}{b^k}, \tag{16.7.1}$$

where

$$\begin{pmatrix} y_1^{(i)}(n) \\ y_2^{(i)}(n) \\ \vdots \\ y_{r_n}^{(i)}(n) \end{pmatrix} = \mathbf{C}^{(i-1)} \begin{pmatrix} a_0(n) \\ a_1(n) \\ \vdots \\ a_{r_n - 1}(n) \end{pmatrix} \mod b,$$

where $\mathbf{C}^{(i)}$ is the $r_n \times r_n$ matrix with entries

$$\mathbf{C}_{st}^{(i)} = \binom{t-1}{s-1} i^{t-s}.$$

Note that $\mathbf{C}^{(0)}$ is identity matrix and these generator matrices have the following cyclic relation:

$$\mathbf{C}^{(i)} = \mathbf{C}^{(1)} \mathbf{C}^{(i-1)}, \quad i = 1, 2, \cdots.$$

Figure (16.8) displays pseudo code for generating matrix $\mathbf{C}^{(i)}$.


```
FaureMat(r,i)
{
    C[1][1] = 0;
    for(s=2;s<=r;s++){
        C[s][s] = 1;  // diagonal entries
        C[1][s] = i;  // the first row
    }
    for(t=3;t<=r;t++){
        for(s=2;s<=t-1;s++){
            C[s,t] = C[s-1][t-1] + i*C[s][t-1];
        }
    }
}
```

Figure 16.8: Pseudo code for generating matrix $\mathbf{C}^{(i)}$ with size $r \times r$

Example 16.7.2. Suppose we want to generate 4-dimensional Faure points and have chosen $b = 5$ as base. To decide the generation matrix size, we have to determine how many generate Faure points. For example, let us generate 12 Faure points. Since $11 = (2,1)_5$ is 2-digit, matrix size r is 2. Thus the generating matrices are

$$\begin{aligned} \mathbf{C}^{(0)} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{C}^{(1)} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \\ \mathbf{C}^{(2)} &= \mathbf{C}^{(1)}\mathbf{C}^{(1)} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{C}^{(3)} = \mathbf{C}^{(1)}\mathbf{C}^{(2)} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

For $n = 0, 1, 2, \cdots, 11$, the vector $\mathbf{a}(n) = \mathbf{C}^{(0)}\mathbf{a}(n)$ are

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

The vectors $\mathbf{C}^{(1)}\mathbf{a}(n)$ are

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

The vectors $\mathbf{C}^{(2)}\mathbf{a}(n)$ are

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

Finally, the vectors $\mathbf{C}^{(2)}\mathbf{a}(n)$ are

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

To convert each of these sets of vectors into fractions, we can apply Eq(16.7.1). By multiplying each vector by $(\frac{1}{5}, \frac{1}{25})$, we have

$$\begin{array}{cccccccccccc} 0 & \frac{1}{5} & \frac{2}{5} & \frac{3}{5} & \frac{4}{5} & \frac{1}{25} & \frac{6}{25} & \frac{11}{25} & \frac{16}{25} & \frac{21}{25} & \frac{2}{25} & \frac{7}{25} \\ 0 & \frac{1}{5} & \frac{2}{5} & \frac{3}{5} & \frac{4}{5} & \frac{6}{25} & \frac{11}{25} & \frac{16}{25} & \frac{21}{25} & \frac{1}{25} & \frac{12}{25} & \frac{17}{25} \\ 0 & \frac{1}{5} & \frac{2}{5} & \frac{3}{5} & \frac{4}{5} & \frac{11}{25} & \frac{16}{25} & \frac{21}{25} & \frac{1}{25} & \frac{6}{25} & \frac{22}{25} & \frac{2}{25} \\ 0 & \frac{1}{5} & \frac{2}{5} & \frac{3}{5} & \frac{4}{5} & \frac{16}{25} & \frac{21}{25} & \frac{1}{25} & \frac{6}{25} & \frac{11}{25} & \frac{7}{25} & \frac{12}{25} \end{array}.$$

□

16.7.3 Sobol's Sequence

A Sobol's sequence is generated on the basis of a set of "direction numbers" $v_1, v_2, \cdots, v_r$, and a generating matrix $\mathbf{V}$. The elements of $\mathbf{V}$ are equal to 0 or 1. Its columns are the binary expansion of a set of direction numbers $v_1, v_2, \cdots, v_r$. Here r could be arbitrarily large, in constructing the kth point in the sequence, think of r as the number of terms in the binary expansion of k. The matrix $\mathbf{V}$ will be upper triangular, so regardless of the number of rows in the full matrix, it suffices to consider the $r \times r$ submatrix of $\mathbf{V}$.

Definition 16.7.3 (Primitive Polynomial). A polynomial

$$x^q + c_1 x^{q-1} + \cdots + c_{q-1} x + 1 \tag{16.7.2}$$

over the field $\mathbb{Z}_2$ is called primitive polynomial if it satisfies the following two properties:

1. it is irreducible.
2. the smallest power p for which the polynomial divides $x^p + 1$ is $p = 2^q - 1$.

□

The polynomial (16.7.2) defines a recurrence relation

$$m_j = 2c_1 m_{j-1} \oplus 2^2 c_2 m_{j-2} \oplus \cdots \oplus 2^{q-1} c_{q-1} m_{j-q+1} \oplus 2^q m_{j-q} \oplus m_{j-q}, \quad (j > q) \tag{16.7.3}$$

where $\oplus$ is bit-wise XOR operation. Some numbers $m_1, \cdots, m_q$ are needed to initialize the recursion. They may be chosen arbitrarily, provided that each m_i is odd and $m_i < 2^i$. From m_j, the direction numbers are defined by setting

$$v_j = \frac{m_j}{2^j}.$$

Note that $m_1, \cdots, m_q$ are referred to as the initial numbers.

To get nth number in the Sobol sequence, denoted by x_n, consider the binary representation of the integer n:

$$n = (\cdots b_3 b_2 b_1)_2. \tag{16.7.4}$$

The result is obtained by computing the bitwise XOR [8] of the direction numbers v_i, for which $b_i \neq 0$:

$$\begin{aligned} x_n &= \mathbf{V} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_r \end{pmatrix} \\ &= b_1 v_1 \oplus b_2 v_2 \oplus \cdots b_r v_r, \end{aligned}$$

where generating matrix V is defined by

$$\mathbf{V} = \begin{pmatrix} \\ v_1 & v_2 & \cdots & v_{r-1} & v_r \\ \\ \end{pmatrix}.$$

[8] $\oplus$ denotes binary addition as follow.

$$0 \oplus 0 = 0, 1 \oplus 0 = 1, 0 \oplus 1 = 1, 1 \oplus 1 = 0.$$

Example 16.7.4. To get Sobol sequence with length $31 = (11111)_2$(5-digits), we need at least five m_i's. Consider the primitive polynomial

$$x^3 + x^2 + 1$$

with degree $q = 3$. The recurrence (16.7.3) becomes

$$m_j = 2m_{j-1} \oplus 8m_{j-3} \oplus m_{j-3}$$

and suppose we initialize it with $m_1 = 1, m_2 = 3, m_3 = 3$. The next two elements in the sequence are as follows:

$$\begin{aligned}
m_4 &= (2 \cdot 3) \oplus (8 \cdot 1) \oplus 1 \\
&= 0110 \oplus 1000 \oplus 0001 \\
&= 1111_{(2)} \\
&= 15, \\
m_5 &= (2 \cdot 15) \oplus (8 \cdot 3) \oplus 3 \\
&= 11110 \oplus 11000 \oplus 00011 \\
&= 00101_{(2)} \\
&= 5.
\end{aligned}$$

From these five values of m_j, we can calculate the corresponding values of v_j by dividing by 2^j. But dividing by 2^j is equivalent to shifting the binary point to the left j places in the representation of m_j. Hence the first five direction numbers are

$$v_1 = \frac{m_1}{2} = 0.1, \quad v_2 = \frac{m_2}{2^2} = 0.11, \quad v_3 = 0.011, \quad v_4 = 0.1111, \quad v_5 = 0.00101$$

and the corresponding generator matrix is

$$\mathbf{V} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \\ v_1 & v_2 & v_3 & v_4 & v_5 \\ \\ \end{pmatrix}.$$

For example, 29th sequence can be obtained from the following procedure:

$$\begin{aligned}
29 &= (11101)_2, \\
x_n &= \mathbf{V} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \\
&= v_1 \oplus v_3 \oplus v_4 \oplus v_5 \\
&= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \oplus \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \oplus \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \oplus \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
&= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \\
&= \frac{1}{2^3} + \frac{1}{2^4} + \frac{1}{2^5} \\
&= \frac{7}{32} \\
&= 0.21875.
\end{aligned}$$

(to be continued)

□

✎ **Gray Code Construction**

Antanov and Saleev point out that Sobol's method simplifies if the usual binary representation of n in (16.7.4) is replaced with a Gray code representation. They also proved the discrepancy is not changed by using the Gray code representation of n. All we need to know is the following:

1. A Gray code is a function mapping an integer n to a corresponding binary representation $G(n)$; the function, for a given integer N, is one-to-one for $0 \leq n \leq 2^N - 1$.
2. A Gray code representation for the integer n is obtained from its binary representation by computing
$$\cdots g_3 g_2 g_1 \quad = \quad (\cdots b_3 b_2 b_1)_2 \oplus (\cdots b_4 b_3 b_2)_2.$$
For example, the Gray code for $n = 13$ is given by
$$G(13) \quad = \quad (1101)_2 \oplus (0110)_2 = (1011).$$
3. The main feature of such a code is that the codes for consecutive number n and $n+1$ differ only one position. That position corresponds to the rightmost zero bit in the binary representation of n.
4. Using the feature of Gray codes, we may streamline generation of a Sobol sequence. Given x_n, we have
$$x_{n+1} \quad = \quad x_n \oplus v_c.$$
where c is the index of the rightmost zero bit b_c in the binary representation of n. Gray code $G(n)$ and $G(n+1)$ differ in the cth bit. For example, when $n = 11 = (1\underline{0}11)_2$, we have $G(11) = (1110)$, $G(12) = (1010)$ and $c = 3$.

Example 16.7.5. (Continued) Since the last bit of the binary representation of 4 is zero, the Gray code for 4 and 5 differ in the last bit.

Let us get x_5 from $x_4 = (0.10100)_2$. The index of the rightmost zero bit of $n = 4 = (010\underline{0})_2$ is $c = 1$. Thus

$$x_5 \quad = \quad x_4 \oplus v_1 \quad = (10100) \oplus (10000) \quad = \quad (00100) \quad = \quad (0.00100)_2 \quad = \quad 0.1250.$$

□

n	$(n)_2$	$G(n)$	Sobol(binary)	Sobol(decimal)	c	x_n
0	0000	0000	0.00000	0.0000		
1	0001	0001	0.10000	0.5000	1	$x_0 \oplus v_1$
2	0010	0011	0.01000	0.2500	2	$x_1 \oplus v_2$
3	0011	0010	0.11000	0.7500	1	$x_2 \oplus v_1$
4	0100	0110	0.10100	0.6250	3	$x_3 \oplus v_3$
5	0101	0111	0.00100	0.1250	1	$x_4 \oplus v_1$
6	0110	0101	0.11100	0.8750	2	$x_5 \oplus v_2$
7	0111	0100	0.01100	0.3750	1	$x_6 \oplus v_1$
8	1000	1100	0.10010	0.5625	4	$x_7 \oplus v_4$
9	1001	1101	0.00010	0.0625	1	$x_8 \oplus v_1$
10	1010	1111	0.11010	0.8125	2	$x_9 \oplus v_2$
11	1011	1110	0.01010	0.3125	1	$x_{10} \oplus v_1$
12	1100	1010	0.00110	0.1875	3	$x_{11} \oplus v_3$
13	1101	1011	0.10110	0.6875	1	$x_{12} \oplus v_1$
14	1110	1001	0.01110	0.4375	2	$x_{13} \oplus v_2$
15	1111	1000	0.11110	0.9375	1	$x_{14} \oplus v_1$

Table 16.1: Gray codes

✎ **Choosing Initial Direction Numbers**

Sobol' provides some guidance for choosing initial direction values.

1. Sobol's Property A: A d-dimensional sequences $x_0, x_1, \cdots$ satisfies Sobol's Property A if for every $j = 0, 1, \cdots$ exactly one of the points $x_k, j2^d \le k < (j+1)2^d$ falls in each of the 2^d cubes of the form
$$\prod_{i=1}^{d}\left[\frac{a_i}{2}, \frac{a_i+1}{2}\right), \quad a_i \in \{0, 1\}.$$
2. Sobol's Property A′: The sequence satisfies Sobol's Property A′ if for every $j = 0, 1, \cdots$ exactly one of the points $x_k, j2^{2d} \le k < (j+1)2^{2d}$ falls in each of the 2^{2d} cubes of the form
$$\prod_{i=1}^{d}\left[\frac{a_i}{4}, \frac{a_i+1}{4}\right), \quad a_i \in \{0, 1, 2, 3\}.$$

Example 16.7.6. To illustrate, suppose $d = 3$ and suppose the first three m_j values for the first three coordinates are as follows:

$$(m_1, m_2, m_3) = (1, 3, 5), \quad (m_1, m_2, m_3) = (1, 1, 7), \quad (m_1, m_2, m_3) = (1, 3, 7).$$

From the first three m_j values in each coordinates, we determine the matrices

$$V^{(1)} = (1, 3, 5) = (1, 11, 101) = \begin{pmatrix} \mathbf{1} & \mathbf{1} & \mathbf{1} \\ 0 & \mathbf{1} & \mathbf{0} \\ 0 & 0 & \mathbf{1} \end{pmatrix}$$

$$V^{(2)} = (1, 1, 7) = (1, 01, 111) = \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{1} \\ 0 & \mathbf{1} & \mathbf{1} \\ 0 & 0 & \mathbf{1} \end{pmatrix}$$

$$V^{(3)} = (1,3,7) = (1,11,111) = \begin{pmatrix} \mathbf{1} & \mathbf{1} & \mathbf{1} \\ 0 & \mathbf{1} & \mathbf{1} \\ 0 & 0 & \mathbf{1} \end{pmatrix}$$

From the first row of each $V^{(1)}, V^{(2)}, V^{(3)}$ we assemble the matrix

$$D = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Because the matrix has a determinate of 1, Sobol's Property A fails.

Remark 16.7.7.

1. There have been reports in various publication that Sobol' numbers also suffer the problem of rapid breakdown of homogeneity in higher dimensions. There is the notion that no low-discrepancy number generator is suitable for high dimensions(see Jäckel(2002)).
2. Low-discrepancy numbers, or more specifically, Sobol' numbers, unlikely pseudo-random number, have the antithetic feature built into them, but only approximately. We use a recommended number of draws such as $2^n - 1$ for some n. Even though when we are not using $2^n - 1$ draws with low-discrepancy numbers, adding the antithetic method to the use of low-discrepancy numbers is unlikely to improve the accuracy, and instead can lead to erroneous result[9].
3. Box-Muller and polar rejection methods are highly dangerous(and should not be used, really) in conjunction with low-discrepancy numbers[10].
4. QMC & Path Dependent Option: To evaluate path dependent option by Monte Carlo method we need to generate daily path. But simple path generation using quasi random number may give rise to serious trouble[11].
   (a) In a straightforward application of Sobol' points to the generation of Brownian paths, the ith coordinate of each point would be transformed to a sample from the standard normal distribution using Φ^{-1}, and these would be scaled and summed using the random walk construction. To the extent that the initial coordinates of a Sobol' sequence have uniformity superior to that of higher-indexed coordinates. This construction does a particularly good job of sampling the first increments of the Brownian path. However, many financial contracts would be primarily sensitive to the terminal value of the Brownian path.
   (b) A simple effective way to generate daily path is the combination of Sobol' sequences with the Brownian bridge construction. Through the Brownian bridge construction, the first coordinate of Sobol' sequence determines the terminal value of the Brownian path, so this value should be particularly well distributed. Moreover, the first several coordinates of the Sobol' sequence determine the general shape of the Brownian path; the last few coordinates influence only the fine detail of the path, which os often less important.

[9] See Jäckel(2002) pp 112.
[10] See Jäckel(2002) pp 106.
[11] See Glasserman(2004).

16.8 Brownian Bridge

Suppose that we are given a Brownian motion $\{W_t\}$ and $0 < t_1 < t_2 < t_3$. Consider the problem of generating W_{t_2} conditional on $W_{t_1} = w_1$ and $W_{t_3} = w_3$. We use the conditional distribution of W_{t_2} given in Theorem (A.3). We know that the conditional distribution is given by

$$\begin{pmatrix} W_{t_1} \\ W_{t_2} \\ W_{t_3} \end{pmatrix} \sim N\left(\mathbf{0}, \begin{pmatrix} t_1 & t_1 & t_1 \\ t_1 & t_2 & t_2 \\ t_1 & t_2 & t_3 \end{pmatrix}\right)$$

After permutation we have

$$\begin{pmatrix} W_{t_2} \\ W_{t_1} \\ W_{t_3} \end{pmatrix} \sim N\left(\mathbf{0}, \begin{pmatrix} t_2 & t_1 & t_2 \\ t_1 & t_1 & t_1 \\ t_2 & t_1 & t_3 \end{pmatrix}\right)$$

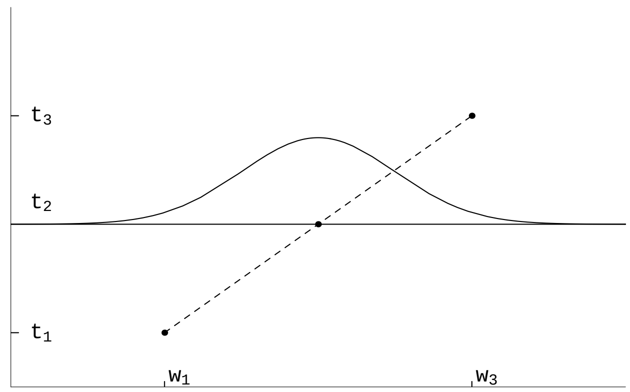


Figure 16.9: Brownian bridge construction

By the Theorem (A.3), we can see that, given $W_{t_1} = w_1, W_{t_3} = t_3$, W_{t_2} is normally distributed with mean

$$\begin{aligned} E[W_{t_2}|W_{t_1} = w_1, W_{t_3} = w_3] &= \mathbf{0} - \begin{pmatrix} t_1 & t_2 \end{pmatrix} \begin{pmatrix} t_1 & t_1 \\ t_1 & t_3 \end{pmatrix}^{-1} \begin{pmatrix} w_1 \\ w_3 \end{pmatrix} \\ &= \frac{1}{t_3 - t_1} \begin{pmatrix} t_1 & t_2 \end{pmatrix} \begin{pmatrix} \frac{t_3}{t_1} & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_3 \end{pmatrix} \\ &= \frac{w_1(t_3 - t_2) + w_3(t_2 - t_1)}{t_3 - t_1} \end{aligned}$$

and variance

$$\begin{aligned} t_2 - \begin{pmatrix} t_1 & t_2 \end{pmatrix} \begin{pmatrix} t_1 & t_1 \\ t_1 & t_3 \end{pmatrix}^{-1} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} &= t_2 - \frac{t_1(t_3 - t_2) + t_2(t_2 - t_1)}{t_3 - t_1} \\ &= \frac{(t_3 - t_2)(t_2 - t_1)}{t_3 - t_1}. \end{aligned}$$

Note that the conditional mean may be obtained by linear interpolation between w_1 and w_3.

Why we use a Brownian bridge construction? The Brownian bridge construction has no computational advantage over simple Brownian path construction. The potential advantage of the Brownian bridge construction arises when it is used with certain variance reduction techniques and low-discrepancy methods.

Example 16.8.1. Suppose we are given S_0 and S_T, i.e.

$$S_T = S_0 \exp\left(\mu T + \sigma W_T\right).$$

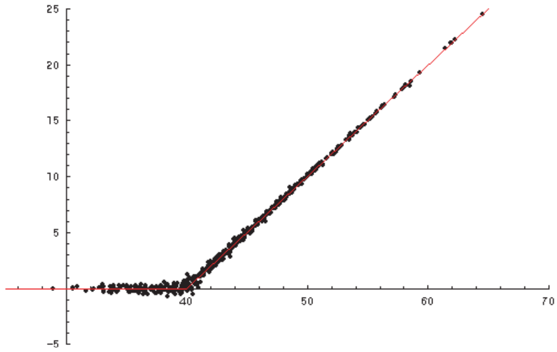
Then for $0 \le t \le T$ we have

$$\begin{aligned}
S_t &= S_0 \exp\left(\mu t + \sigma\left(\frac{t}{T} W_T + \sqrt{\frac{t(T-t)}{T}} Z\right)\right), \quad Z \sim N(0,1) \\
&= S_0 \exp\left(\mu t + \sigma\left(\frac{t}{T}\,\frac{1}{\sigma}\left(\log \frac{S_T}{S_0} - \mu T\right) + \sqrt{\frac{t(T-t)}{T}} Z\right)\right) \\
&= S_0 \exp\left(\frac{t}{T} \log \frac{S_T}{S_0} + \sigma\sqrt{\frac{t(T-t)}{T}} Z\right) \\
&= S_0 \left(\frac{S_T}{S_0}\right)^{\frac{t}{T}} \exp\left(\sigma\sqrt{\frac{t(T-t)}{T}} Z\right).
\end{aligned}$$

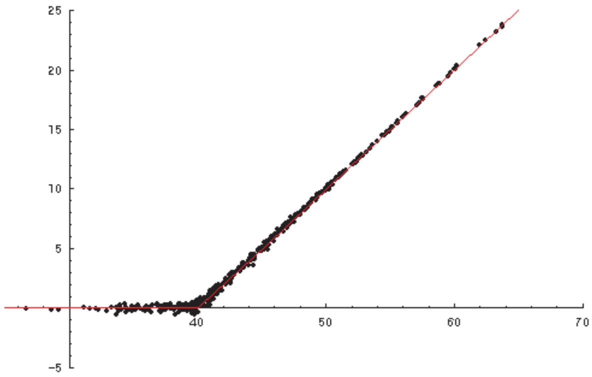
□

16.9 Example: Discrete Delta Hedging

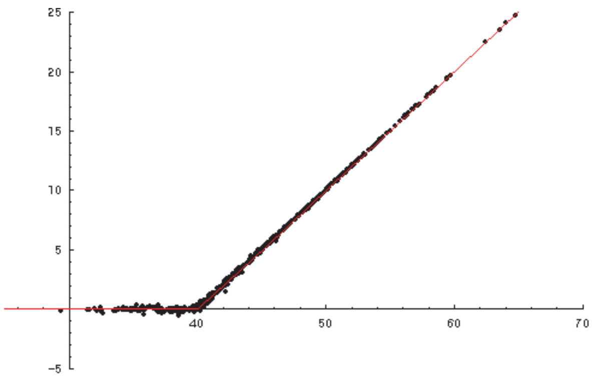
Figure (16.10) demonstrates discrete hedging errors.



(a) Delta Hedging: $M = 500$, $N = 50$, $S(0) = 42$, $K = 40$, $r = 0.1$, $\sigma = 0.2$



(b) Delta Hedging: $M = 500$, $N = 100$, $S(0) = 42$, $K = 40$, $r = 0.1$, $\sigma = 0.2$



(c) Delta Hedging: $M = 500$, $N = 180$, $S(0) = 42$, $K = 40$, $r = 0.1$, $\sigma = 0.2$

Figure 16.10: Discrete hedging errors

Chapter 17

Finite Element Methods

The finite-difference approach with equidistant grids is easy to understand and straightforward to implement. The resulting uniform rectangular grids are comfortable, but in many applications not flexible enough. Steep gradients of the solution require a finer grid such that the difference quotients provide good approximations of the differentials. On the other hand, a flat gradient may be well modeled on a coarse grid. Such a flexibility of the grid is hard to obtain with finite-difference methods. An alternative type of methods for solving PDEs that does provide the desired flexibility is the class of finite-element methods.

The basic idea of *finite element method* is to approximate the solution of a given differential equation with a set of algebraically simple functions. The spacial domain of the differential equation is divided into subdomains called **elements**. A "finite element" designates a mathematical topic such as a subinterval and defined thereupon a piece of function. In a narrow sense, a finite element is a pair consisting of one piece of subdomain and the corresponding function defined thereupon, mostly a polynomial.

There are alternative names as *variational methods*, or *weighted residuals*, or *Galerkin methods*. As these different names suggest, there are several different approaches leading to finite elements.

In this chapter, we follow Seydel(2002) and Topper(2005).

17.1 The Principle of Weighted Residuals

To explain the principle of weighted residuals we discuss the formally simple case of the differential equation

$$\mathcal{L}u \quad = \quad f \qquad \text{on } \mathcal{D} \subset \mathbb{R}^n.$$

Here $\mathcal{L}$ symbolizes a linear differential operator. The piecewise approach starts with a partition of the domain into a finite number of subdomains $\mathcal{D}_k$,

$$\mathcal{D} \quad = \quad \bigcup_k \mathcal{D}_k, \tag{17.1.1}$$

where the partition is assumed disjoint up to the boundaries of $\mathcal{D}_k$. In the one-dimensional case($n = 1$), for example, the $\mathcal{D}_k$ are subintervals of a whole interval $\mathcal{D}$. In the two-dimensional case (17.1.1) may describe a partition into triangles.

An approximation to the solution u can be represented by

$$w(x) \quad := \quad \sum_{i=1}^{N} c_i \varphi_i(x), \quad c_i \in \mathbb{R}, \tag{17.1.2}$$

where φ are called *basis functions*, or *trivial functions* or *shape functions*. The free parameters $c_1, \cdots, c_N$ are to be determined such that $w \approx u$. Define the residual as

$$R \quad := \quad \mathcal{L}w - f.$$

We look for a w such that R becomes "small". In order to determine $c_1, \cdots, c_N$, N conditions must be established. To this end we weight the residual by introducing N weighting functions $\psi_1, \cdots, \psi_N$ and require

$$\int_{\mathcal{D}} R\psi_j \quad = \quad 0, \quad \text{for } j = 1, \cdots, N.$$

This system of equations can be rewritten as

$$\int_{\mathcal{D}} \mathcal{L}w\psi_j \quad = \quad \int_{\mathcal{D}} f\psi_j, \quad j = 1, \cdots, N. \tag{17.1.3}$$

(17.1.2) implies

$$\int_{\mathcal{D}} \mathcal{L}w\psi_j \quad = \quad \int_{\mathcal{D}} \left(\sum_i c_i \mathcal{L}\varphi_i \right) \psi_j \quad = \quad \sum_i c_i \underbrace{\int_{\mathcal{D}} L\varphi_i\psi_j}_{:=a_{ij}}.$$

The integrals a_{ij} constitute a matrix $\mathbf{A}$. The $r_j := \int_{\mathcal{D}} f\psi_j$ set up a vector $\mathbf{r}$ and the coefficients c_j a vector $\mathbf{c} = (c_1, \cdots, c_N)^T$. This allow to rewrite the system of equations in vector notation as

$$\mathbf{Ac} \quad = \quad \mathbf{r}.$$

Example 17.1.1.

1. Galerkin Method: Choose $\psi_j := \varphi_j$. Then $a_{ij} = \int \mathcal{L}\varphi_i\varphi_j$.
2. Collocation Method: Choose $\psi_j := \delta(x - x_j)$. Then (17.1.3) can be rewritten as
$$\mathcal{L}w(x_j) \quad = \quad f(x_j).$$
3. Least Square Method: Choose
$$\psi_j \quad := \quad \frac{\partial R}{\partial c_j}.$$
4. Subdomain Method: Choose
$$\psi_j(x) \quad := \quad \begin{cases} 1, & x \in \mathcal{D}_k, \\ 0, & x \notin \mathcal{D}_k. \end{cases}$$

□

✎ **Hat Functions**

If the matrix **A** of the linear equations is sparse, then the system can be solved efficiently even when it is large. In order to achieve sparsity we require that $\psi_j \equiv 0$ on most of the subdomains $\mathcal{D}_k$.

Definition 17.1.2 (Hat Functions). Assume we are given $x_0, \cdots, x_m$. For $1 \le i \le m-1$, define

$$\varphi_i(x) \quad := \quad \begin{cases} \dfrac{x - x_{i-1}}{x_i - x_{i-1}}, & x_{i-1} \le x < x_i \\ \dfrac{x_{i+1} - x}{x_{i+1} - x_i}, & x_i \le x < x_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

and for boundary functions

$$\varphi_0(x) \quad := \quad \begin{cases} \dfrac{x_1 - x}{x_1 - x_0}, & x_0 \le x < x_1 \\ 0 & \text{otherwise} \end{cases}$$

$$\varphi_m(x) \quad := \quad \begin{cases} \dfrac{x - x_{m-1}}{x_m - x_{m-1}}, & x_{m-1} \le x \le x_m \\ 0 & \text{otherwise} \end{cases}$$

□

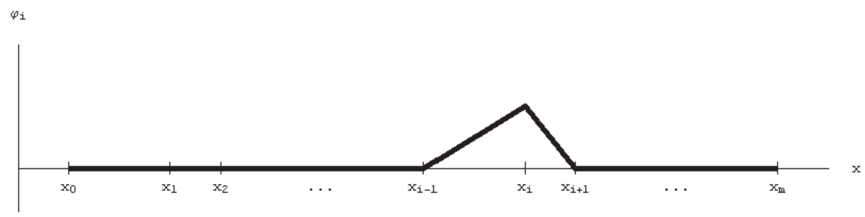


Figure 17.1: Hat function: non-equidistant grid

17.2 The Galerkin method

The differential equation to be solved in this section is a linear parabolic partial differential equation forward in time:

$$\frac{\partial u}{\partial t} = a_0(x)\frac{\partial^2 u}{\partial x^2} + a_1(x)\frac{\partial u}{\partial x} + a_2(x)u + f(x). \tag{17.2.1}$$

Denoting the derivative with respect to time with a dot($\dot{u}$) and the derivative with respect to x with a prime(u'), (17.2.1) reads

$$\dot{u} = a_0 u'' + a_1 u' + a_2 u + f.$$

The BC(boundary conditions) are given by

$$\begin{aligned} p_1 u(x_{\min}) + q_1 u'(x_{\min}) &= r_1 \\ p_2 u(x_{\max}) + q_2 u'(x_{\max}) &= r_2 \end{aligned}$$

and the IC(initial condition) is given by

$$u(t_0, x) = u_0(x).$$

1. Discretization: The domain $[x_{\min}, x_{\max}]$ is subdivided into intervals

$$x_{\min} = x_0, x_1, \cdots, x_N = x_{\max}.$$

   Also we assume all intervals to be of equal length:

$$x_i = x_0 + ih$$

   with h being the step size. At this point, time is not discretized.

2. Interpolation: Define $n_i(x)$ by

$$n_i(x) := \begin{cases} 0, & x \le x_{i-1} \\ \dfrac{x - x_{i-1}}{x_i - x_{i-1}}, & x_{i-1} \le x \le x_i \\ \dfrac{x_{i+1} - x}{x_{i+1} - x_i}, & x_i \le x \le x_{i+1} \\ 0, & x_{i+1} \le x \end{cases}$$

   The interpolation function is given by

$$\tilde{u}(t, x) = \sum_{i=0}^{N} \tilde{u}_i(t) n_i(x). \tag{17.2.2}$$

   Note that the spatial variable x is discretized, while the time variable t is not. The approach of separating time and space is known as the method of Kantorovich.

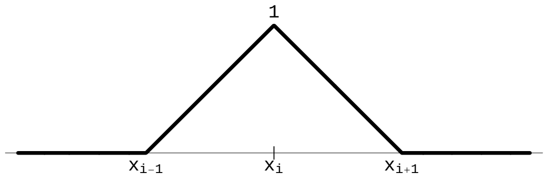


Figure 17.2: $n_i(x)$

3. Elemental formulation: We define the residual as

$$\begin{aligned} R(x,t) &:= a_0\tilde{u}'' + a_1\tilde{u}' + a_2\tilde{u} + f - \dot{\tilde{u}} \\ &= (a_0\tilde{u}')' + (a_1 - a_0')\tilde{u}' + a_2\tilde{u} + f - \dot{\tilde{u}}. \end{aligned}$$

The Galerkin criterion requires that for $k = 0, \cdots, N$

$$\begin{aligned} &\int_{x_{\min}}^{x_{\max}} R(t,x) n_k(x) dx = 0 \\ \Longleftrightarrow \quad &\int_{x_{\min}}^{x_{\max}} \Big\{(a_0\tilde{u}')' + (a_1 - a_0')\tilde{u}' + a_2\tilde{u} + f - \dot{\tilde{u}}\Big\} n_k(x) dx = 0. \end{aligned} \tag{17.2.3}$$

Integrating the first term to eliminate the second order derivative, we have

$$\begin{aligned} \int_{x_{\min}}^{x_{\max}} (a_0\tilde{u}')' n_k dx &= \Big[a_0\tilde{u}' n_k\Big]_{x_{\min}}^{x_{\max}} - \int_{x_{\min}}^{x_{\max}} a_0\tilde{u}' n_k' dx \\ &= a_0\tilde{u}'(x_{\max}) n_k(x_{\max}) - a_0\tilde{u}'(x_{\min}) n_k(x_{\min}) - \int_{x_{\min}}^{x_{\max}} a_0\tilde{u}' n_k' dx. \end{aligned}$$

Inserting the above result in (17.2.3) leads to

$$\begin{aligned} &\int_{x_{\min}}^{x_{\max}} \Big\{a_0\tilde{u}' n_k' - (a_1 - a_0')\tilde{u}' n_k - a_2\tilde{u}\, n_k - f\, n_k + \dot{\tilde{u}}\, n_k\Big\} dx \\ = \quad & a_0\tilde{u}'(x_{\max}) n_k(x_{\max}) - a_0\tilde{u}'(x_{\min}) n_k(x_{\min}), \quad k = 0, \cdots, N. \end{aligned}$$

Now we consider only the LHS. Inserting (17.2.2) into LHS results in

$$\begin{aligned} &\int_{x_{\min}}^{x_{\max}} \left[a_0 n_k' \left(\sum_{i=0}^{N} \tilde{u}_i n_i'\right) - (a_1 - a_0') n_k \left(\sum_{i=0}^{N} \tilde{u}_i n_i'\right) - a_2 n_k \left(\sum_{i=0}^{N} \tilde{u}_i n_i\right) - f n_k + n_k \left(\sum_{i=0}^{N} \dot{\tilde{u}}_i n_i\right)\right] dx \\ = \quad & \sum_{i=0}^{N} \left[\int_{x_{\min}}^{x_{\max}} (a_0 n_k' n_i' - (a_1 - a_0') n_k n_i' - a_2 n_k n_i) dx\right] \tilde{u}_i + \sum_{i=0}^{N} \left[\int_{x_{\min}}^{x_{\max}} n_k n_i dx\right] \dot{\tilde{u}}_i - \int_{x_{\min}}^{x_{\max}} n_k f dx. \end{aligned}$$

Hence we have

$$\begin{aligned} &\sum_{i=0}^{N} \left[\int_{x_{\min}}^{x_{\max}} (a_0 n_k' n_i' - (a_1 - a_0') n_k n_i' - a_2 n_k n_i) dx\right] \tilde{u}_i + \sum_{i=0}^{N} \left[\int_{x_{\min}}^{x_{\max}} n_k n_i dx\right] \dot{\tilde{u}}_i \\ = \quad & \int_{x_{\min}}^{x_{\max}} n_k f dx + a_0\tilde{u}'(x_N)\delta_{kN} - a_0\tilde{u}'(x_0)\delta_{k0}, \quad k = 0, \cdots, N \end{aligned} \tag{17.2.4}$$

where δ_{ij} is the Kronecker delta. For all k with $0 < k < N$ each equation can be written as

$$\sum_{i=0}^{N} \left[\int_{x_{\min}}^{x_{\max}} (a_0 n_k' n_i' - (a_1 - a_0') n_k n_i' - a_2 n_k n_i) dx\right] \tilde{u}_i$$

$$
\begin{aligned}
= \quad & \left[\int_{x_{k-1}}^{x_k} (a_0 n_k' n_{k-1}' - (a_1 - a_0') n_k n_{k-1}' - a_2 n_k n_{k-1}) dx\right] \tilde{u}_{k-1} \\
& + \quad \left[\int_{x_{k-1}}^{x_k} (a_0 n_k' n_k' - (a_1 - a_0') n_k n_k' - a_2 n_k n_k) dx\right] \tilde{u}_k \\
& + \quad \left[\int_{x_k}^{x_{k+1}} (a_0 n_k' n_k' - (a_1 - a_0') n_k n_k' - a_2 n_k n_k) dx\right] \tilde{u}_k \\
& + \quad \left[\int_{x_k}^{x_{k+1}} (a_0 n_k' n_{k+1}' - (a_1 - a_0') n_k n_{k+1}' - a_2 n_k n_{k+1}) dx\right] \tilde{u}_{k+1} = \int_{x_{k-1}}^{x_k} n_k f dx + \int_{x_k}^{x_{k+1}} n_k f dx
\end{aligned}
$$

Collecting all integrations on $[x_k, x_{k+1}]$, we have

$$
\begin{aligned}
k: \quad & \left[\int_{x_k}^{x_{k+1}} (a_0 n_k' n_k' - (a_1 - a_0') n_k n_k' - a_2 n_k n_k) dx\right] \tilde{u}_k \\
& \quad + \quad \left[\int_{x_k}^{x_{k+1}} (a_0 n_k' n_{k+1}' - (a_1 - a_0') n_k n_{k+1}' - a_2 n_k n_{k+1}) dx\right] \tilde{u}_{k+1}, \quad \int_{x_k}^{x_{k+1}} n_k f dx \\
k+1: \quad & \left[\int_{x_k}^{x_{k+1}} (a_0 n_{k+1}' n_k' - (a_1 - a_0') n_{k+1} n_k' - a_2 n_{k+1} n_k) dx\right] \tilde{u}_k \\
& \quad + \quad \left[\int_{x_k}^{x_{k+1}} (a_0 n_{k+1}' n_{k+1}' - (a_1 - a_0') n_{k+1} n_{k+1}' - a_2 n_{k+1} n_{k+1}) dx\right] \tilde{u}_{k+1}, \quad \int_{x_k}^{x_{k+1}} n_{k+1} f dx.
\end{aligned}
$$

Define

$$
\mathbf{N}_k = \begin{pmatrix} \frac{x_{k+1} - x}{x_{k+1} - x_k} \\ \frac{x - x_k}{x_{k+1} - x_k} \end{pmatrix}
$$

We can rewrite (17.2.4) in terms of $\mathbf{N}$:

$$
\int_{x_k}^{x_{k+1}} \left(a_0 \mathbf{N}_k' \mathbf{N}_k'^T - (a_1 - a_0') \mathbf{N}_k \mathbf{N}_k'^T - a_2 \mathbf{N}_k \mathbf{N}_k^T\right) dx, \quad \int_{x_k}^{x_{k+1}} f \mathbf{N}_k dx.
$$

Example 17.2.1. Suppose a_0, a_1 and a_2 are constant.

$$
\begin{aligned}
a_0 \int_{x_k}^{x_{k+1}} \mathbf{N}_k' \mathbf{N}_k'^T dx &= \frac{a_0}{\Delta x^2} \int_{x_k}^{x_{k+1}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \begin{pmatrix} -1 & 1 \end{pmatrix} dx \\
&= \frac{a_0}{\Delta x^2} \int_{x_k}^{x_{k+1}} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} dx \\
&= \frac{a_0}{\Delta x} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \\
a_1 \int_{x_k}^{x_{k+1}} \mathbf{N}_k \mathbf{N}_k'^T dx &= \frac{a_1}{\Delta x^2} \int_{x_k}^{x_{k+1}} \begin{pmatrix} x_{k+1} - x \\ x - x_k \end{pmatrix} \begin{pmatrix} -1 & 1 \end{pmatrix} dx \\
&= \frac{a_1}{\Delta x^2} \int_{x_k}^{x_{k+1}} \begin{pmatrix} x - x_{k+1} & x_{k+1} - x \\ x_k - x & x - x_k \end{pmatrix} dx \\
&= \frac{a_1}{2} \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}, \\
a_2 \int_{x_k}^{x_{k+1}} \mathbf{N}_k \mathbf{N}_k^T dx &= \frac{a_2}{\Delta x^2} \int_{x_k}^{x_{k+1}} \begin{pmatrix} x_{k+1} - x \\ x - x_k \end{pmatrix} \begin{pmatrix} x_{k+1} - x & x - x_k \end{pmatrix} dx
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{a_2}{\Delta x^2}\int_{x_k}^{x_{k+1}} \begin{pmatrix} (x_{k+1}-x)^2 & (x_{k+1}-x)(x-x_k) \\ (x-x_k)(x_{k+1}-x) & (x-x_k)^2 \end{pmatrix} dx \\
&= \frac{a_2\Delta x}{6}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.
\end{aligned}
$$

- $$\begin{pmatrix} \frac{a_0}{\Delta x}+\frac{a_1}{2}-\frac{a_2\Delta x}{3} & -\frac{a_0}{\Delta x}-\frac{a_1}{2}-\frac{a_2\Delta x}{6} \\ -\frac{a_0}{\Delta x}+\frac{a_1}{2}-\frac{a_2\Delta x}{6} & \frac{a_0}{\Delta x}-\frac{a_1}{2}-\frac{a_2\Delta x}{3} \end{pmatrix} := \begin{pmatrix} \alpha & \gamma \\ \beta & \omega \end{pmatrix}$$
  $$\rightarrow \begin{pmatrix} \alpha & \gamma & 0 & \cdots & & & \\ \beta & \alpha+\omega & \gamma & \cdots & & & \\ 0 & \beta & \alpha+\omega & \ddots & & & \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ & & & \ddots & \alpha+\omega & \gamma & 0 \\ & & & \cdots & \beta & \alpha+\omega & \gamma \\ & & & \cdots & 0 & \beta & \omega \end{pmatrix} := \mathbf{A}$$
- $$\frac{\Delta x}{6}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$
  $$\rightarrow \frac{\Delta x}{6}\begin{pmatrix} 2 & 1 & 0 & \cdots & & & \\ 1 & 4 & 1 & \cdots & & & \\ 0 & 1 & 4 & \ddots & & & \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ & & & \ddots & 4 & 1 & 0 \\ & & & \cdots & 1 & 4 & 1 \\ & & & \cdots & 0 & 1 & 2 \end{pmatrix} := \mathbf{B}.$$
- $$\frac{1}{\Delta x}\begin{pmatrix} \int_{x_k}^{x_{k+1}}(x_{k+1}-x)f(x)dx \\ \int_{x_k}^{x_{k+1}}(x-x_k)f(x)dx \end{pmatrix}$$
  $$\rightarrow \frac{1}{\Delta x}\begin{pmatrix} \int_{x_0}^{x_1}(x_1-x)f(x)dx \\ \int_{x_0}^{x_1}(x-x_0)f(x)dx + \int_{x_1}^{x_2}(x_2-x)f(x)dx \\ \int_{x_1}^{x_2}(x-x_1)f(x)dx + \int_{x_2}^{x_3}(x_3-x)f(x)dx \\ \vdots \\ \int_{x_{N-2}}^{x_{N-1}}(x-x_{N-2})f(x)dx + \int_{x_{N-1}}^{x_N}(x_N-x))f(x) \\ \int_{x_{N-1}}^{x_N}(x-x_{N-1})f(x)dx \end{pmatrix} := \mathbf{K}$$
- $$\mathbf{A\tilde{u}}+\mathbf{B\dot{\tilde{u}}} = \mathbf{K} + \begin{pmatrix} -a_0\tilde{u}'(x_0) \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ a_0\tilde{u}'(x_N) \end{pmatrix}, \quad \tilde{\mathbf{u}}(t) = \begin{pmatrix} \tilde{u}_0(t) \\ \tilde{u}_1(t) \\ \tilde{u}_2(t) \\ \vdots \\ \tilde{u}_{N-2}(t) \\ \tilde{u}_{N-1}(t) \\ \tilde{u}_N(t) \end{pmatrix}$$

□

Example 17.2.2. Consider a differential system of the form

$$\mathbf{A\tilde{u}}+\mathbf{B\dot{\tilde{u}}} = \mathbf{K}, \quad \tilde{\mathbf{u}}(t) = \Big(\widetilde{u}_0(t), \cdots, \widetilde{u}_N(t)\Big)^T.$$

Let

$$\tilde{\mathbf{u}}^n = \Big(\tilde{u}_0(nh), \cdots, \tilde{u}_N(nh)\Big)^T, \quad h = \Delta t.$$

With the θ-method, we have

$$\tilde{\mathbf{u}}^{n+1} = \tilde{\mathbf{u}}^n + h\Big((1-\theta)\dot{\tilde{\mathbf{u}}}^n + \theta\dot{\tilde{\mathbf{u}}}^{n+1}\Big).$$

Multiplying by B leads to

$$\begin{aligned}
\mathbf{B}\tilde{\mathbf{u}}^{n+1} &= \mathbf{B}\tilde{\mathbf{u}}^n + h\Big((1-\theta)\mathbf{B}\dot{\tilde{\mathbf{u}}}^n + \theta\mathbf{B}\dot{\tilde{\mathbf{u}}}^{n+1}\Big) \\
&= \mathbf{B}\tilde{\mathbf{u}}^n + h(1-\theta)(-\mathbf{A}\tilde{\mathbf{u}}^n + \mathbf{K}_n) + h\theta(-\mathbf{A}\tilde{\mathbf{u}}^{n+1} + \mathbf{K}_{n+1}).
\end{aligned}$$

Hence the recurrence formula is given by

$$\Big(\mathbf{B} + h\theta\mathbf{A}\Big)\tilde{\mathbf{u}}^{n+1} = \Big(\mathbf{B} - h(1-\theta)\mathbf{A}\Big)\tilde{\mathbf{u}}^n + h\Big(\theta\mathbf{K}_{n+1} + (1-\theta)\mathbf{K}_n\Big)$$

□

Example 17.2.3. For the initial condition $u(t_0, x) = u_0(x)$, we have

$$u_0(x_j) = \sum_{i=1}^{N} \tilde{u}_i(t_0)n_i(x_j) = \tilde{u}_j(t_0).$$

For the boundary condition, we obtain

$$\begin{aligned}
u(t_n, x_0) &= \sum_{i=1}^{N} \tilde{u}_i(t_n)n_i(x_0) = \tilde{u}_0(t_n), \\
u(t_n, x_N) &= \sum_{i=1}^{N} \tilde{u}_i(t_n)n_i(x_N) = \tilde{u}_N(t_n)
\end{aligned}$$

Also the linear boundary conditions

$$\begin{aligned}
2u(t_n, x_1) &= u(t_n, x_0) + u(t_n, x_2), \\
2u(t_n, x_{N-1} &= u(t_n, x_N) + u(t_n, x_{N-2})
\end{aligned}$$

imply that

$$\begin{aligned}
2\tilde{u}_1(t_n) &= \tilde{u}_0(t_n) + \tilde{u}_2(t_n) \\
2\tilde{u}_{N-1}(t_n) &= \tilde{u}_N(t_n) + \tilde{u}_{N-2}(t_n)
\end{aligned}$$

respectively. □

Example 17.2.4 (Black-Scholes PDE). The log-transformed Black-Scholes PDE is given by

$$-\frac{\partial u}{\partial t} = \frac{1}{2}\sigma^2\frac{\partial^2 u}{\partial x^2} + \left(r - q - \frac{1}{2}\sigma^2\right)\frac{\partial u}{\partial x} - ru.$$

In (17.2.1), a_0, a_1, a_2 and f are constant as

$$a_0 = \frac{1}{2}\sigma^2, \quad a_1 = r - q - \frac{1}{2}\sigma^2, \quad a_2 = -r, \quad f = 0.$$

Chapter 18

The Malliavin Calculus

18.1 Introduction to the Malliavin Calculus

The Greeks(sensitivities) in financial market are typically defined as the partial derivatives of the expected payoff function with respect to underlying model parameters. In general, finite difference approximations are heavily used to simulate Greeks by means of Monte Carlo procedures.

To simplify the discussion, let us specialize it to the case of the Delta. Then one has to compute a Monte Carlo estimator of $u(x)$ and a Monte Carlo estimator for $u(x+\epsilon)$ for some small ϵ; the Delta is estimated by $\frac{u(x+\epsilon)-u(x)}{\epsilon}$. If the simulations of the two estimators are drawn independently, then it is proved in Glynn(1989) that the best possible convergence rate is typically $n^{-1/4}$. Replacing the forward finite difference estimator by the central difference $\frac{u(x+\epsilon)-u(x-\epsilon)}{2\epsilon}$ improves the optimal convergence rate to $n^{-1/3}$. However, by using common random numbers for both Monte Carlo estimators, one can achieve the convergence rate $n^{-1/2}$ which is the best that can be expected from ordinary Monte Carlo methods.

However, it is known that the finite difference approximation soon becomes inefficient particularly when payoff functions are complex and discontinuous. This is often the case when we deal with exotic options.

To overcome this difficulty, Broadie and Glasserman(1996) proposed *the likelihood ratio method* which allows to achieve the $n^{-1/2}$ convergence rate. The proposed method is to put the differential of the payoff function inside the expectation operator required to evaluate the sensitivity. An important limitation of this method is that it can only be applied to simple payoff functions and this method applicable only when the density of the random variable involved is explicitly known.

Recently, Fourniè et al.(1999) suggested the use of Malliavin calculus, by means of integration by parts, to shift the differential operator from the expected payoff to the underlying diffusion kernel, introducing a weighting function. The real advantage of using Malliavin calculus is that it is applicable when we deal with random variable whose density is not explicitly known as the case Asian options.

In Malliavin calculus, the Greeks can be computed as the expectation of the original payoff(ϕ) times a weight:

$$Greek = E[e^{-\int_0^T r_s ds}\ \phi(S_T) \times weight].$$

An important advantage is that the Malliavin weight does not depend on the payoff function ϕ. However, Malliavin weight is not unique.

Remark 18.1.1. The success of estimator for delta relies on the fast rate of mean-square of $u(x+\epsilon)$ to $u(x)$. For the digital option, we get the convergence of $u(x+\epsilon)$ to $u(x)$ is only linear in ϵ:

$$E\Big[|u(x+\epsilon)-u(x)|^2\Big] \quad = \quad O(\epsilon).$$

On the contrary, in the case of the call option, it can be shown that

$$E\Big[|u(x+\epsilon)-u(x)|^2\Big] \quad = \quad O(\epsilon^2).$$

18.2 Simple Examples

We consider a stochastic differential equation which is described by

$$dS_t \quad = \quad S_t\Big((r-\delta)dt+\sigma dW_t\Big).$$

We are, in European options, interest in studying how to evaluate the sensitivity with respect to model parameters, e.g., present price S_0, volatility σ, etc., of the expected payoff

$$u(0,x) \quad := \quad E[e^{-rT}\phi(S_T)], \quad S_0 = x.$$

Example 18.2.1 (Delta). Let us calculate the delta. Let $p(z)$ be the PDF of W_T, i.e.

$$p(z) \quad = \quad \frac{1}{\sqrt{2\pi T}}\exp\left(-\frac{z^2}{2T}\right).$$

Since $S_T = x\exp\left(\lambda T+\sigma W_T\right), \lambda = r-\delta-\sigma^2/2$, we have

$$\begin{aligned}
\frac{\partial u}{\partial x} &= \frac{\partial}{\partial x}E[e^{-rT}\phi(S_T)] \\
&= e^{-rT}E\left[\phi'(S_T)\,\frac{\partial S_T}{\partial x}\right] \\
&= e^{-rT}E\left[\phi'(S_T)\,\frac{S_T}{x}\right] \\
&= e^{-rT}\int_{\mathbb{R}}\phi'\left(x\exp\left(\lambda T+\sigma z\right)\right)\,\exp\left(\lambda T+\sigma z\right)\,p(z)dz \\
&= -e^{-rT}\int_{\mathbb{R}}\phi\left(x\exp\left(\lambda T+\sigma z\right)\right)\frac{1}{x\sigma}\,p'(z)dz \qquad \text{(integration by parts)} \\
&= e^{-rT}\int_{\mathbb{R}}\phi\left(x\exp\left(\lambda T+\sigma z\right)\right)\frac{z}{x\sigma T}\,p(z)dz, \qquad \Big(\text{by } p'(z) = -\frac{z}{T}p(z)\Big) \\
&= e^{-rT}\int_{\mathbb{R}}\phi(S_T)\frac{z}{x\sigma T}\,p(z)dz \\
&= e^{-rT}E\left[\phi(S_T)\frac{W_T}{x\sigma T}\right].
\end{aligned}$$

Thus, in this case, the Malliavin weight is $\frac{W_T}{x\sigma T}$. □

Example 18.2.2 (Gamma). Let us calculate the Malliavin weight for gamma.

$$
\begin{aligned}
\frac{\partial^2 u}{\partial x^2} &= e^{-rT}\ \frac{\partial}{\partial x} E\left[\phi(S_T)\frac{W_T}{x\sigma T}\right] \\
&= e^{-rT}\ E\left[\phi'(S_T)\frac{S_T}{x}\ \frac{W_T}{x\sigma T} - \phi(S_T)\frac{W_T}{x^2\sigma T}\right] \\
&= e^{-rT}\int_{\mathbb{R}} \phi'(S_T)\frac{S_T}{x}\ \frac{z}{x\sigma T}p(z)dx \quad - \quad e^{-rT}\ E\left[\phi(S_T)\frac{W_T}{x^2\sigma T}\right] \\
&= -e^{-rT}\int_{\mathbb{R}} \phi(S_T)\frac{1}{x\sigma}\ \frac{1}{x\sigma T}\left(p(z)+zp'(z)\right)dx \quad - \quad e^{-rT}\ E\left[\phi(S_T)\frac{W_T}{x^2\sigma T}\right] \\
&= e^{-rT}\int_{\mathbb{R}} \phi(S_T)\ \frac{1}{x^2\sigma T}\left(-\frac{1}{\sigma}+\frac{z^2}{\sigma T}\right)p(z)dx \quad - \quad e^{-rT}\ E\left[\phi(S_T)\frac{W_T}{x^2\sigma T}\right] \\
&= e^{-rT}E\left[\phi(S_T)\ \frac{1}{x^2\sigma T}\left(-\frac{1}{\sigma}+\frac{W_T^2}{\sigma T}\right)\right] \quad - \quad e^{-rT}\ E\left[\phi(S_T)\frac{W_T}{x^2\sigma T}\right] \\
&= e^{-rT}E\left[\phi(S_T)\ \frac{1}{x^2\sigma T}\left(-\frac{1}{\sigma}+\frac{W_T^2}{\sigma T}-W_T\right)\right]. \qquad (18.2.1)
\end{aligned}
$$

□

Example 18.2.3 (Vega). Next Greek vega is the index that measures sensitivity of the expected payoff with respect to the volatility σ, which can be computed as

$$
\begin{aligned}
\frac{\partial u}{\partial \sigma} &= e^{-rT}\ \frac{\partial}{\partial \sigma}E\left[\phi(S_T)\right] \\
&= e^{-rT}\ E\left[\phi'(S_T)S_T(-\sigma T+W_T)\right] \\
&= e^{-rT}\ \int_{\mathbb{R}} \phi'(S_T)S_T(-\sigma T+z)p(z)dz \\
&= -e^{-rT}\ \int_{\mathbb{R}} \phi(S_T)\left\{\frac{1}{\sigma}p(z)+\left(-T+\frac{z}{\sigma}\right)p'(z)\frac{1}{\sigma}p(z)\right\}dz \\
&= -e^{-rT}\ \int_{\mathbb{R}} \phi(S_T)\left\{\frac{1}{\sigma}p(z)+\left(-T+\frac{z}{\sigma}\right)\left(-\frac{z}{T}\right)p(z)\right\}dz \\
&= e^{-rT}\ \int_{\mathbb{R}} \phi(S_T)\left(-\frac{1}{\sigma}+\frac{z^2}{\sigma T}-z\right)p(z)dz \\
&= e^{-rT}E\left[\phi(S_T)\left(-\frac{1}{\sigma}+\frac{W_T^2}{\sigma T}-W_T\right)\right]. \qquad (18.2.2)
\end{aligned}
$$

Comparing (18.2.1) and (18.2.2), we find the following relationship between gamma(Γ) and vega(V).

$$\Gamma = \frac{V}{x^2\sigma T}$$

□

Example 18.2.4 (Rho & Theta). The last Greek rho is given by as follows:

$$
\begin{aligned}
\frac{\partial u}{\partial r} &= -Te^{-rT}E[\phi(S_T)] + e^{-rT}E[\phi'(S_T)S_T T] \\
&= -T\times \text{price} \ + xT\times \text{delta} \\
&= -Tu + xT\Delta
\end{aligned}
$$

$$= \quad e^{-rT} E\left[\phi(S_T)\left(-T + \frac{W_T}{\sigma}\right)\right].$$

The theta(Θ) can be obtained from the Black-Scholes PDE:

$$\Theta \quad = \quad ru - \frac{1}{2}\sigma^2 x^2 \Gamma - (r-\delta)x\Delta.$$

$\square$

✎ **Greeks for Two Assets**

In this paragraph, we consider stochastic differential equations which are described by

$$dS_t^i \quad = \quad S_t^i((r-\delta_i)dt + \sigma_i dZ_t^i), \quad i = 1, 2,$$

where Z_t^1 and Z_t^2 are Brownian motion with correlation ρ, i.e.

$$dZ_t^1 \cdot dZ_t^2 \quad = \quad \rho\, dt.$$

Suppose we are given a European option's payoff $\phi(h_1, h_2)$ and let $p(z_1, z_2)$ be the joint PDF of Z_T^1 and Z_T^2 i.e.

$$p(z_1, z_2) \quad = \quad \frac{1}{2\pi T\sqrt{(1-\rho^2)}} \exp\left(-\frac{z_1^2 + z_2^2 - 2\rho z_1 z_2}{2T(1-\rho^2)}\right).$$

Let

$$u(0, x, y) \quad := \quad e^{-rT} E[\phi(S_T^1, S_T^2)], \quad S_0^1 = x, S_0^2 = y.$$

Example 18.2.5 (Delta). Let us compute the delta.

$$\begin{aligned}
\frac{\partial u}{\partial x}(0, x, y) &= e^{-rT} E\left[\frac{\partial \phi}{\partial h_1}(S_T^1, S_T^2)\ \frac{\partial S_T^1}{\partial x}\right] \\
&= e^{-rT} E\left[\frac{\partial \phi}{\partial h_1}(S_T^1, S_T^2)\ \frac{S_T^1}{x}\right] \\
&= e^{-rT} \iint_{\mathbb{R}^2} \frac{\partial \phi}{\partial h_1}(S_T^1, S_T^2)\ \frac{S_T^1}{x} p(z_1, z_2) dz_1 dz_2 \\
&= e^{-rT} \iint_{\mathbb{R}^2} \phi(S_T^1, S_T^2)\ \frac{1}{x\sigma_1} \frac{\partial p}{\partial z_1}(z_1, z_2) dz_1 dz_2 \\
&= -e^{-rT} \iint_{\mathbb{R}^2} \phi(S_T^1, S_T^2)\ \frac{1}{x\sigma_1} \frac{\partial p}{\partial z_1}(z_1, z_2) dz_1 dz_2 \\
&= e^{-rT} \iint_{\mathbb{R}^2} \phi(S_T^1, S_T^2)\ \frac{1}{x\sigma_1} \frac{z_1 - \rho z_2}{T(1-\rho^2)} p(z_1, z_2) dz_1 dz_2 \\
&= e^{-rT} E\left[\phi(S_T^1, S_T^2)\ \frac{Z_T^1 - \rho Z_T^2}{x\sigma_1 T(1-\rho^2)}\right].
\end{aligned}$$

Note that we can find independent Brownian motions W_t^1 and W_t^2 such that, for $t > 0$,

$$Z_t^1 \quad = \quad W_t^1,$$

$$Z_t^2 = \rho W_t^1 + \sqrt{1-\rho^2} W_t^2.$$

Then we have

$$\begin{aligned}
\frac{\partial u}{\partial x}(0,x,y) &= e^{-rT} E\left[\phi(S_T^1, S_T^2)\ \frac{Z_T^1 - \rho Z_T^2}{x\sigma_1 T(1-\rho^2)}\right] \\
&= e^{-rT} E\left[\phi(S_T^1, S_T^2)\ \frac{1}{x\sigma_1 T}\left(W_T^1 - \frac{\rho}{\sqrt{1-\rho^2}} W_T^2\right)\right].
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\frac{\partial u}{\partial y}(0,x,y) &= e^{-rT} E\left[\phi(S_T^1, S_T^2)\ \frac{Z_T^2 - \rho Z_T^1}{y\sigma_2 T(1-\rho^2)}\right] \\
&= e^{-rT} E\left[\phi(S_T^1, S_T^2)\ \frac{1}{y\sigma_2 T\sqrt{1-\rho^2}} W_T^2\right].
\end{aligned}$$

□

Example 18.2.6 (Gamma). We now want to get the Malliavin weight for gamma.

$$\begin{aligned}
&\frac{\partial^2 u}{\partial x^2}(0,x,y) \\
= \;& e^{-rT}\frac{\partial}{\partial x} E\left[\phi(S_T^1, S_T^2)\ \frac{Z_T^1 - \rho Z_T^2}{x\sigma_1 T(1-\rho^2)}\right] \\
= \;& e^{-rT} E\left[\frac{\partial \phi}{\partial h_1}(S_T^1, S_T^2)\frac{S_T^1}{x}\frac{Z_T^1 - \rho Z_T^2}{x\sigma_1 T(1-\rho^2)}\right] - e^{-rT} E\left[\phi(S_T^1, S_T^2)\ \frac{Z_T^1 - \rho Z_T^2}{x^2\sigma_1 T(1-\rho^2)}\right] \\
= \;& e^{-rT}\iint_{\mathbb{R}^2}\frac{\partial \phi}{\partial h_1}(S_T^1, S_T^2)\frac{S_T^1}{x}\frac{z_1 - \rho z^2}{x\sigma_1 T(1-\rho^2)} p(z_1,z_2)dz_1dz_2 - e^{-rT} E\left[\phi(S_T^1, S_T^2)\ \frac{Z_T^1 - \rho Z_T^2}{x^2\sigma_1 T(1-\rho^2)}\right] \\
= \;& e^{-rT}\iint_{\mathbb{R}^2}\phi(S_T^1, S_T^2)\frac{1}{x^2\sigma_1^2 T(1-\rho^2)}\left(-1 + \frac{(z_1-\rho z_2)^2}{T(1-\rho^2)}\right) p(z_1,z_2)dz_1dz_2 - e^{-rT} E\left[\phi(S_T^1, S_T^2)\ \frac{Z_T^1 - \rho Z_T^2}{x^2\sigma_1 T(1-\rho^2)}\right] \\
= \;& e^{-rT} E\left[\phi(S_T^1, S_T^2)\frac{1}{x^2\sigma_1^2 T(1-\rho^2)}\left(-1 + \frac{(Z_T^1-\rho Z_T^2)^2}{T(1-\rho^2)}\right)\right] - e^{-rT} E\left[\phi(S_T^1, S_T^2)\ \frac{Z_T^1 - \rho Z_T^2}{x^2\sigma_1 T(1-\rho^2)}\right] \\
= \;& e^{-rT} E\left[\phi(S_T^1, S_T^2)\frac{1}{x^2\sigma_1 T(1-\rho^2)}\left(-\frac{1}{\sigma_1} + \frac{(Z_T^1-\rho Z_T^2)^2}{\sigma_1 T(1-\rho^2)} - (Z_T^1 - \rho Z_T^2)\right)\right].
\end{aligned}$$

Thus the weight for $\frac{\partial^2 u}{\partial x^2}$ is given by

$$\begin{aligned}
&\frac{1}{x^2\sigma_1 T(1-\rho^2)}\left(-\frac{1}{\sigma_1} + \frac{(Z_T^1-\rho Z_T^2)^2}{\sigma_1 T(1-\rho^2)} - (Z_T^1 - \rho Z_T^2)\right) \\
= \;& \frac{1}{x^2\sigma_1 T(1-\rho^2)}\left(-\frac{1}{\sigma_1} + \frac{1-\rho^2}{\sigma_1 T}\left(W_T^1 - \frac{\rho}{\sqrt{1-\rho^2}}W_T^2\right)^2 - (1-\rho^2)\left(W_T^1 - \frac{\rho}{\sqrt{1-\rho^2}}W_T^2\right)\right).
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\frac{\partial^2 u}{\partial y^2}(0,x,y) &= e^{-rT} E\left[\phi(S_T^1, S_T^2)\frac{1}{y^2\sigma_2 T(1-\rho^2)}\left(-\frac{1}{\sigma_2} + \frac{(Z_T^2-\rho Z_T^1)^2}{\sigma_2 T(1-\rho^2)} - (Z_T^2 - \rho Z_T^1)\right)\right] \\
&= e^{-rT} E\left[\phi(S_T^1, S_T^2)\frac{1}{y^2\sigma_2 T(1-\rho^2)}\left(-\frac{1}{\sigma_2} + \frac{(W_T^2)^2}{\sigma_2 T} - \sqrt{1-\rho^2}W_T^2\right)\right].
\end{aligned}$$

Also, we can obtain the cross-gamma:

$$\begin{aligned}
\frac{\partial^2 u}{\partial x \partial y} &= e^{-rT} E\left[\phi(S_T^1, S_T^2)\frac{\rho T(1-\rho^2) + (Z_T^1 - \rho Z_T^2)(Z_T^2 - \rho Z_T^1)}{xy\sigma_1\sigma_2 T^2(1-\rho^2)^2}\right] \\
&= e^{-rT} E\left[\phi(S_T^1, S_T^2)\left(\frac{\rho}{xy\sigma_1\sigma_2 T(1-\rho^2)} + \frac{\big(\sqrt{1-\rho^2}W_T^1 - \rho W_T^2\big)W_T^2}{xy\sigma_1\sigma_2 T^2(1-\rho^2)}\right)\right] \\
&= e^{-rT} E\left[\phi(S_T^1, S_T^2)\left(\frac{\rho}{xy\sigma_1\sigma_2 T(1-\rho^2)} + \frac{W_T^1 W_T^2}{xy\sigma_1\sigma_2 T^2\sqrt{1-\rho^2}} - \frac{\rho\big(W_T^2\big)^2}{xy\sigma_1\sigma_2 T^2(1-\rho^2)}\right)\right].
\end{aligned}$$

□

Example 18.2.7 (Vega).

$$\begin{aligned}
\frac{\partial u}{\partial \sigma_1} &= e^{-rT}\frac{\partial}{\partial \sigma_1} E[\phi(S_T^1, S_T^2)] \\
&= e^{-rT} E\left[\frac{\partial \phi}{\partial h_1}(S_T^1, S_T^2)\,\frac{\partial S_T^1}{\partial \sigma_1}\right] \\
&= e^{-rT} E\left[\frac{\partial \phi}{\partial h_1}(S_T^1, S_T^2)\,S_T\,(-\sigma_1 T + Z_T^1)\right] \\
&= e^{-rT}\iint_{\mathbb{R}^2}\frac{\partial \phi}{\partial h_1}(S_T^1, S_T^2)\,S_T\,(-\sigma_1 T + z_1)p(z_1,z_2)dz_1dz_2 \\
&= -e^{-rT}\iint_{\mathbb{R}^2}\phi(S_T^1, S_T^2)\frac{1}{\sigma_1}\left\{p(z_1,z_2) + (-\sigma_1 T + z_1)\frac{\partial p}{\partial z_1}(z_1,z_2)\right\}dz_1dz_2 \\
&= e^{-rT}\iint_{\mathbb{R}^2}\phi(S_T^1, S_T^2)\left\{-\frac{1}{\sigma_1} + \left(-T + \frac{z_1}{\sigma_1}\right)\frac{z_1 - \rho z_2}{T(1-\rho^2)}\right\}p(z_1,z_2)dz_1dz_2 \\
&= e^{-rT}\iint_{\mathbb{R}^2}\phi(S_T^1, S_T^2)\left\{-\frac{1}{\sigma_1} - \frac{z_1 - \rho z_2}{(1-\rho^2)} + \frac{(z_1 - \rho z_2)z_1}{\sigma_1 T(1-\rho^2)}\right\}p(z_1,z_2)dz_1dz_2 \\
&= e^{-rT} E\left[\phi(S_T^1, S_T^2)\left\{-\frac{1}{\sigma_1} - \frac{Z_T^1 - \rho Z_T^2}{(1-\rho^2)} + \frac{(Z_T^1 - \rho Z_T^2)Z_T^1}{\sigma_1 T(1-\rho^2)}\right\}\right] \\
&= e^{-rT} E\left[\phi(S_T^1, S_T^2)\left\{-\frac{1}{\sigma_1} - \left(W_T^1 - \frac{\rho}{\sqrt{1-\rho^2}}W_T^2\right) + \frac{1}{\sigma_1 T}\left(W_T^1 - \frac{\rho}{\sqrt{1-\rho^2}}W_T^2\right)W_T^1\right\}\right].
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\frac{\partial u}{\partial \sigma_2} &= e^{-rT} E\left[\phi(S_T^1, S_T^2)\Big(-\frac{1}{\sigma_2} - \frac{Z_T^2 - \rho Z_T^1}{(1-\rho^2)} + \frac{(Z_T^2 - \rho Z_T^1)Z_T^2}{\sigma_2 T(1-\rho^2)}\Big)\right] \\
&= e^{-rT} E\left[\phi(S_T^1, S_T^2)\Big(-\frac{1}{\sigma_2} - \frac{W_T^2}{\sqrt{1-\rho^2}} + \frac{W_T^2}{\sigma_2 T}\left(\frac{\rho}{\sqrt{1-\rho^2}}W_T^1 + W_T^2\right)\Big)\right].
\end{aligned}$$

□

Example 18.2.8 (Rho). The rho is given by

$$\begin{aligned}
\frac{\partial u}{\partial r} &= -Te^{-rT}E[\phi(S_T^1, S_T^2)] + e^{-rT}E\left[\frac{\partial \phi}{\partial h_1}(S_T^1, S_T^2)\frac{\partial S_T^1}{\partial r}\right] + e^{-rT}E\left[\frac{\partial \phi}{\partial h_2}(S_T^1, S_T^2)\frac{\partial S_T^2}{\partial r}\right] \\
&= -Te^{-rT}E[\phi(S_T^1, S_T^2)] + e^{-rT}E\left[\frac{\partial \phi}{\partial h_1}(S_T^1, S_T^2)S_T^1 T\right] + e^{-rT}E\left[\frac{\partial \phi}{\partial h_2}(S_T^1, S_T^2)S_T^2 T\right]
\end{aligned}$$

$$= \quad -T\,u + xT\Delta_1 + yT\Delta_2.$$

□

Up to now, we have obtained the Malliavin weights from explicit probability density functions. This method is referred to as *the likelihood ratio method*. The following section gives you some more ideas about the likelihood ratio method.

18.3 Likelihood Ratio Method

Let X be the parameter of the payoff function ϕ and p its density. For example, X equals S_T in the case of a European option and $(S_{t_1}, \cdots, S_{t_n})$ in the case of a discrete Asian option. The price of the option is then given by

$$u \quad = \quad E[e^{-rT}\phi(X)] \quad = \quad \int_0^\infty e^{-rT}\phi(x)p(x)dx.$$

supposing that the payoff function ϕ does not depend on the parameter λ, the derivative of the price of the option with respect to λ is given by

$$\begin{aligned}
\frac{\partial u}{\partial \lambda} &= \int_0^\infty e^{-rT}\phi(x)\frac{\partial p}{\partial \lambda}(x)dx \\
&= \int_0^\infty e^{-rT}\phi(x)\frac{\partial \log p}{\partial \lambda}(x)p(x)dx \\
&= E[e^{-rT}\phi(X)\frac{\partial \log p}{\partial \lambda}(X)].
\end{aligned}$$

Hence the Malliavin weight is given by

$$\frac{\partial \log p(X)}{\partial \lambda}.$$

Example 18.3.1 (Delta). The PDF p of S_T is given by

$$p_x(y) \quad = \quad \frac{1}{\sqrt{2\pi T}\sigma y}\exp\left(-\frac{\left(\log\frac{y}{x} - \left(r - \delta - \frac{1}{2}\sigma^2\right)T\right)^2}{2T\sigma^2}\right), \quad S_0 = x.$$

Then we get the Malliavin weight for delta as follows:

$$\begin{aligned}
weight(\Delta) &= \frac{\partial \log p_x(S_T)}{\partial x} \\
&= \frac{\log\frac{S_T}{x} - (r - \delta - \frac{\sigma^2}{2})T}{x\sigma^2 T} \\
&= \frac{\sigma W_T}{x\sigma^2 T} \\
&= \frac{W_T}{x\sigma T}.
\end{aligned}$$

□

Example 18.3.2 (Gamma). Similarly the weight for gamma is given by

$$\frac{\partial^2 p(S_T)}{\partial x^2}\frac{1}{p(S_T)} = \frac{1}{x^2\sigma T}\left(-\frac{1}{\sigma}+\frac{W_T^2}{\sigma T}-W_T\right).$$

$\square$

This method has however the disadvantage of requiring an explicit expression of the density function. To avoid the need of a closed formula of the density function, Fournié et al.(1999) introduced a new method based on Malliavin calculus.

18.4 Mathematical Preliminary

In this section, we will discuss highly sophisticated mathematics - Malliavin derivatives and Skorohod integrals.

Why we need Malliavin derivatives and Skorohod integrals?

The answer is that the Malliavin weight could be expressed in terms of Malliavin derivative and Skorohod integral. There exists an infinity of weighting function and it can be proved that the weight of minimal total variance is precisely the one given by the likelihood ratio method(see Theorem (18.4.8)).

✎ **Skorohod Integral**

Definition 18.4.1 (Symmetric Square Integrable Function). A real function $g : [0,T]^n \to \mathbb{R}$ is called symmetric if

$$g(x_{\sigma_1}, \cdots x_{\sigma_n}) \quad = \quad g(x_1, \cdots, x_n)$$

for all permutation σ of $(1, 2, \cdots, n)$. If in addition

$$\|g\|^2_{L^2([0,T]^n)} \quad := \quad \int_{[0,T]^n} g^2(x_1, \cdots, x_n) dx_1 \cdots dx_n < \infty$$

we say that $g \in \widehat{L}^2([0,T]^n)$, the space of symmetric square integrable functions on $[0,T]^n$. □

Definition 18.4.2 (Iterated Ito Integral). If f is a deterministic function defined on $S_n (n \geq 1)$ such that

$$\|f\|^2_{L^2(S_n)} \quad := \quad \int_{S_n} f^2(t_1, \cdots, t_n) dt_1 \cdots dt_n < \infty,$$

then we can form the *(n-fold) iterated Ito integral*

$$J_n(f) \quad := \quad \int_0^T \int_0^{t_n} \cdots \int_0^{t_3} \int_0^{t_2} f(t_1, t_2, \cdots, t_{n-1}, t_n) dW(t_1) dW(t_2) \cdots dW(t_{n-1}) dW(t_n),$$

because at each Ito integration with respect to $dW(t_i)$ the integrand is $\mathcal{F}_t$-adapted and square integrable with respect to $dP \times dt_i, 1 \leq i \leq n$.
If $g \in \widehat{L}^2([0,T]^n)$ we define

$$I_n(g) \quad := \quad n! J_n(g).$$

□

Theorem 18.4.3 (The Wiener-Ito Chaos Expansion). Let ψ be an $\mathcal{F}_T$-measurable random variable such that

$$\|\psi\|^2_{L^2(\Omega)} \quad := \quad E_P[\psi^2] < \infty.$$

Then there exists a (unique) sequence $\{f_n\}_{n=0}^{\infty}$ of (deterministic) functions $f_n \in \widehat{L}^2([0,T]^n)$ such that

$$\psi(\omega) \quad = \quad \sum_{n=0}^{\infty} I_n(f) \quad \text{(convergence in } L^2(\Omega)).$$

Moreover, we have the isometry

$$\|\psi\|^2_{L^2(\Omega)} = \sum_{n=0}^{\infty} n!\|f_n\|^2_{L^2([0,T]^n)}.$$

Let $u(t,\omega), \omega \in \Omega, t \in [0,T]$ be a stochastic process such that

$$u(t,\cdot) \text{ is } \mathcal{F}_T\text{-measurable for all } t \in [0,T] \tag{18.4.1}$$

and

$$E[u^2(t,\omega)] < \infty, \quad \text{for all } t \in [0,T]. \tag{18.4.2}$$

Then for each $t \in [0,T]$ we can apply the Wiener-Ito chaos expansion to the random variable $\omega \to (t,\omega)$ and obtain functions $f_{n,t}(t_1,\cdots,t_n) \in \widehat{L}^2(\mathbb{R}^2)$ such that

$$u(t,\omega) = \sum_{n=0}^{\infty} I_n(f_{n,t}(\cdot)).$$

Hence wee may regard f_n as a function of $n+1$ variables $t_1,\cdots,t_n,t$. Since this function is symmetric with respect to its first n variables, its symmetrization $\widetilde{f_n}$ as a function of $n+1$ variables $t_1,\cdots t_n,t$ is given by, with $t_{n+1}=t$,

$$\widetilde{f}(t_1,\cdots,t_{n+1}) = \frac{1}{n+1}\Big[f(t_1,\cdots,t_{n+1}) + \cdots + f(t_1,\cdots,t_{i-1},t_{i+1},t_{n+1},t_i) + \cdots + f(t_2,\cdots,t_{n+1},t_1)\Big],$$

where we only sum over those permutations σ of the indices $(1,\cdots,n+1)$ which interchange the last component with one of the others and leave the rest in place.

Definition 18.4.4 (Skorohod Integral). Suppose $u(t,\omega)$ is a stochastic process satisfying (18.4.1), (18.4.2) and with Wiener-Ito chaos expansion

$$u(t,\omega) = \sum_{n=0}^{\infty} I_n(f_n(\cdot,t)).$$

Then we define the Skorohod integral of u by

$$\delta(u) := \int_0^T u(t,\omega)\delta W(t) := \sum_{n=0}^{\infty} I_{n+1}(\widetilde{f}_n) \quad \text{(when convergent)} \tag{18.4.3}$$

where $\widetilde{f}_n$ is the symmetrization of $f_n(t_1,\cdots,t_n,t)$ as a function of $n+1$ variables $t_1,\cdots,t_n,t$. We say u is Skorohod-integrable and write $u \in \mathrm{Dom}(\delta)$ if the series in (18.4.3) converges in $L^2(P)$. Note that the Skorohod integral coincides with the Ito integral on the space of adapted processes. □

✎ **Malliavin Derivatives**

Let $\{W(t), 0 \le t \le T\}$ be a n-dimensional Brownian motion defined on a complete probability space $(\Omega,\mathcal{F},P)$ and we shall denote by $\{\mathcal{F}_t\}$ the augmentation with respect to P of the filtration generated by W. Let $\mathscr{C}$ be the set of random variables F of the form:

$$F = f\left(\int_0^{\infty} h_1(t)dW(t),\cdots,\int_0^{\infty} h_n(t)dW(t)\right), f \in \mathscr{S}(\mathbb{R}^n)$$

where $\mathscr{S}(\mathbb{R}^n)$ denotes the set of infinitely differentiable and rapidly decreasing [1] functions on $\mathbb{R}^n$ and $h_1, \cdots, h_n \in L^2(\Omega \times \mathbb{R}_+)$. For $F \in \mathscr{C}$, the Malliavin derivative DF of F is defined as the process $\{D_t F, t \geq 0\}$ of $L^2(\Omega \times \mathbb{R}_+)$ with values in $L^2(\mathbb{R}_+)$ which we denote by H:

$$D_t F = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \left(\int_0^\infty h_1(t) dW(t), \cdots, \int_0^\infty h_n(t) dW(t) \right) h_i(t), \quad t \geq 0 \quad a.s.$$

We also define the norm on $\mathscr{C}$

$$\|F\|_{1,2} = (E[F^2])^{1/2} + \left(E\left[\int_0^\infty (D_t F)^2 dt \right] \right)^{1/2}.$$

Then $\mathbb{D}^{1,2}$ denotes the Banach space which is the completion of $\mathscr{C}$ with respect to the norm $\|\|_{1,2}$. The derivative operator D is a closed linear operator defined in $\mathbb{D}^{1,2}$ and its values are in $L^2(\Omega \times \mathbb{R}_+)$.

✎ **First Variation Process(AKA Tangent Process)**

Let $\{X(t), t \geq 0\}$ be an $\mathbb{R}^n$ valued Ito process whose dynamics are driven by the stochastic differential equation

$$\underbrace{dX(t)}_{n\times 1} = \underbrace{b(X(t))}_{n\times 1} dt + \underbrace{\sigma(X(t))}_{n\times n} \underbrace{dW(t)}_{n\times 1},$$

where b and σ are supposed to be continuously differentiable functions with bounded derivatives. Let $\{Y(t), t \geq 0\}_{n\times n}$ be the associated first variation process defined by the stochastic differential equation

$$\underbrace{dY(t)}_{n\times n} = b'(X(t))Y(t)dt + \sum_{i=1}^{n} \sigma_i'(X(t))Y(t)dW_i(t), \quad Y(0) = I_n, \tag{18.4.4}$$

where I_n is the identity matrix of $\mathbb{R}^n$, primes denotes derivatives and σ_i is the i-th column vector of σ. Then the process $\{X(t), t \geq 0\}$ belongs to $\mathbb{D}^{1,2}$ and its Malliavin derivative is given by

$$\underbrace{D_s X(t)}_{n\times n} = Y(t)Y(s)^{-1}\sigma(X(s))\mathbf{1}_{\{s\leq t\}}, \quad s \geq 0 \quad a.s. \tag{18.4.5}$$

Example 18.4.5. We can get the following some examples:

1. $D_t W_T = \mathbf{1}_{\{t\leq T\}}$.
2. $D_t f(W_T) = f'(W_T)\mathbf{1}_{\{t\leq T\}}$.
3. $D_t \left(\int_0^T f(W_s) dW_s \right) = \int_t^T f'(W_s) dW_s + f(W_t)$.
4. When $dS_t = \mu S_t dt + \sigma S_t dW_t$, we have that by (18.4.5)

$$dY_t = \mu Y_t dt + \sigma Y_t dW_t, \quad Y_0 = 1,$$

[1] $\xi : X \to \mathbb{R}$ is rapidly decreasing iff

$$\forall j, k, \quad \lim_{|x|\to\infty} \left| |x|^i \xi^{(k)}(x) \right| = 0.$$

$$
\begin{aligned}
Y_t &= \frac{S_t}{S_0}, \\
D_t S_T &= \frac{S_T/S_0}{S_t/S_0}\,\sigma S_T \mathbf{1}_{\{t\le T\}} \;=\; \sigma S_T \mathbf{1}_{\{t\le T\}}.
\end{aligned} \tag{18.4.6}
$$

5. Chain Rule: $D_t[\psi(F)] = \nabla\psi(F)\cdot D_t F$. For example, when we are given $S_T = S_0\exp(\mu T + \sigma W_T)$,

$$D_t S_T = \sigma S_T \cdot D_t[W_T] = \sigma S_T\,\mathbf{1}_{\{t\le T\}}.$$

6. For any u_t adapted, we have

$$
\begin{aligned}
D_t\left(\int_0^T u_s dW_s\right) &= u_t + \int_t^T D_t u_s dW_s, \\
D_t\left(\int_0^T u_s ds\right) &= \int_t^T D_t u_s ds.
\end{aligned}
$$

7. Product Rule: For F and G,

$$D_t(FG) = GD_t(F) + FD_t(G).$$

□

Example 18.4.6 (Double Asset Option). Suppose that we are given

$$\begin{pmatrix} dX_1 \\ dX_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 X_1 \\ \lambda_2 X_2 \end{pmatrix} dt + \begin{pmatrix} \sigma_1 X_1 & 0 \\ \sigma_2\rho X_2 & \sigma_2\sqrt{1-\rho^2}X_2 \end{pmatrix}\begin{pmatrix} dW_1 \\ dW_2 \end{pmatrix},$$

where $\lambda_1, \lambda_2, \sigma_1, \sigma_2$, and ρ are constant. Then from (18.4.4), the first variation process is given by

$$
\begin{aligned}
\begin{pmatrix} dY_{11} & dY_{12} \\ dY_{21} & dY_{22} \end{pmatrix} &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix} dt + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2\rho \end{pmatrix}\begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix} dW_1 \\
&\quad + \begin{pmatrix} 0 & 0 \\ 0 & \sigma_2\sqrt{1-\rho^2} \end{pmatrix}\begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix} dW_2, \\
\begin{pmatrix} Y_{11}(0) & Y_{12}(0) \\ Y_{21}(0) & Y_{22}(0) \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.
\end{aligned}
$$

Hence we have

$$
\begin{aligned}
dY_{11} &= \lambda_1 Y_{11} dt + \sigma_1 Y_{11} dW_1, \quad Y_{11}(0) = 1, \\
dY_{12} &= \lambda_1 Y_{12} dt + \sigma_1 Y_{12} dW_1, \quad Y_{12}(0) = 0, \\
dY_{21} &= \lambda_2 Y_{21} dt + \sigma_2\rho Y_{21} dW_1 + \sigma_2\sqrt{1-\rho^2}Y_{21} dW_2, \quad Y_{21}(0) = 0, \\
dY_{22} &= \lambda_2 Y_{22} dt + \sigma_2\rho Y_{22} dW_1 + \sigma_2\sqrt{1-\rho^2}Y_{22} dW_2, \quad Y_{22}(0) = 1.
\end{aligned}
$$

The solution is given by

$$Y_{11}(t) = \frac{X_1(t)}{X_1(0)}, \quad Y_{12}(t) = 0, \quad Y_{21}(t) = 0, \quad Y_{22}(t) = \frac{X_2(t)}{X_2(0)}.$$

i.e.

$$\begin{pmatrix} Y_{11}(t) & Y_{12}(t) \\ Y_{21}(t) & Y_{22}(t) \end{pmatrix} = \begin{pmatrix} \frac{X_1(t)}{X_1(0)} & 0 \\ 0 & \frac{X_2(t)}{X_2(0)} \end{pmatrix}.$$

Now, let us calculate Malliavin derivative of $X(t)$.

$$\begin{aligned} D_sX(t) &= Y(t)Y(s)^{-1}\sigma(X(s))\mathbf{1}_{\{s\le t\}} \\ &= \begin{pmatrix} \frac{X_1(t)}{X_1(0)} & 0 \\ 0 & \frac{X_2(t)}{X_2(0)} \end{pmatrix}\begin{pmatrix} \frac{X_1(0)}{X_1(s)} & 0 \\ 0 & \frac{X_2(0)}{X_2(s)} \end{pmatrix}\begin{pmatrix} \sigma_1X_1(s) & 0 \\ \sigma_2\rho X_2(s) & \sigma_2\sqrt{1-\rho^2}X_2(s) \end{pmatrix}\mathbf{1}_{\{s\le t\}} \\ &= \begin{pmatrix} \sigma_1X_1(t) & 0 \\ \sigma_2\rho X_2(t) & \sigma_2\sqrt{1-\rho^2}X_2(t) \end{pmatrix}\mathbf{1}_{\{s\le t\}} \\ &= \sigma(X(t))\mathbf{1}_{\{s\le t\}}. \end{aligned} \tag{18.4.7}$$

(to be continued)

□

Example 18.4.7 (Multi-Asset). In the most case of multi-asset options, we have to resort to Monte Carlo simulation. Thus Malliavin calculus is once more a useful tool.

Suppose we are given

$$\begin{pmatrix} dX_1 \\ dX_2 \\ \vdots \\ dX_n \end{pmatrix} = \begin{pmatrix} \lambda_1X_1 \\ \lambda_2X_2 \\ \vdots \\ \lambda_nX_n \end{pmatrix}dt + \begin{pmatrix} \sigma_1X_1(t) & & & \mathbf{0} \\ & \sigma_2X_2(t) & & \\ & & \ddots & \\ \mathbf{0} & & & \sigma_nX_n(t) \end{pmatrix}\begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \cdots & \epsilon_{1n} \\ \epsilon_{21} & \epsilon_{22} & \cdots & \epsilon_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{n1} & \epsilon_{n2} & \cdots & \epsilon_{nn} \end{pmatrix}\begin{pmatrix} dW_1 \\ dW_2 \\ \vdots \\ dW_n \end{pmatrix}.$$

Let

$$\Phi = \begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \cdots & \epsilon_{1n} \\ \epsilon_{21} & \epsilon_{22} & \cdots & \epsilon_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{n1} & \epsilon_{n2} & \cdots & \epsilon_{nn} \end{pmatrix}, \quad \sigma(X(t)) = \begin{pmatrix} \sigma_1X_1(t) & & & \mathbf{0} \\ & \sigma_2X_2(t) & & \\ & & \ddots & \\ \mathbf{0} & & & \sigma_nX_n(t) \end{pmatrix}\begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \cdots & \epsilon_{1n} \\ \epsilon_{21} & \epsilon_{22} & \cdots & \epsilon_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{n1} & \epsilon_{n2} & \cdots & \epsilon_{nn} \end{pmatrix}.$$

Typically Φ is a lower triangula matrix.

$$\begin{pmatrix} dY_{11} & \cdots & dY_{1n} \\ dY_{21} & \cdots & dY_{2n} \\ \vdots & \ddots & \vdots \\ dY_{n1} & \cdots & dY_{nn} \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & \mathbf{0} \\ & \lambda_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \lambda_n \end{pmatrix}\begin{pmatrix} Y_{11} & \cdots & Y_{1n} \\ Y_{21} & \cdots & Y_{2n} \\ \vdots & \ddots & \vdots \\ Y_{n1} & \cdots & Y_{nn} \end{pmatrix} + \sum_{i=1}^{n}\begin{pmatrix} \epsilon_{1i} & & & \mathbf{0} \\ & \epsilon_{2i} & & \\ & & \ddots & \\ \mathbf{0} & & & \epsilon_{ni} \end{pmatrix}\begin{pmatrix} Y_{11} & \cdots & Y_{1n} \\ Y_{21} & \cdots & Y_{2n} \\ \vdots & \ddots & \vdots \\ Y_{n1} & \cdots & Y_{nn} \end{pmatrix}dW_i$$

Hence we have

$$Y_{ij}(t) = \begin{cases} \frac{X_i(t)}{X_i(0)} & i = j \\ 0 & i \neq j \end{cases}$$

Also we have

$$D_sX(t) = Y(t)Y(s)^{-1}\sigma(X(s))\mathbf{1}_{\{s\le t\}}$$

$$
\begin{aligned}
&= \begin{pmatrix} \frac{X_1(t)}{X_1(0)} & & & \mathbf{0} \\ & \frac{X_2(t)}{X_2(0)} & & \\ & & \ddots & \\ \mathbf{0} & & & \frac{X_n(t)}{X_n(0)} \end{pmatrix} \begin{pmatrix} \frac{X_1(0)}{X_1(s)} & & & \mathbf{0} \\ & \frac{X_2(0)}{X_2(s)} & & \\ & & \ddots & \\ \mathbf{0} & & & \frac{X_n(0)}{X_n(s)} \end{pmatrix} \begin{pmatrix} \sigma_1 X_1(s) & & & \mathbf{0} \\ & \sigma_2 X_2(s) & & \\ & & \ddots & \\ \mathbf{0} & & & \sigma_n X_n(s) \end{pmatrix} \Phi\ \mathbf{1}_{\{s \le t\}} \\
&= \begin{pmatrix} \sigma_1 X_1(t) & & & \mathbf{0} \\ & \sigma_2 X_2(t) & & \\ & & \ddots & \\ \mathbf{0} & & & \sigma_n X_n(t) \end{pmatrix} \Phi\ \mathbf{1}_{\{s \le t\}} \\
&= \sigma(X(t))\ \mathbf{1}_{\{s \le t\}}
\end{aligned}
$$

(to be continued)

□

Theorem 18.4.8. The weighting function with minimal variance denoted by π_0 is the conditional expectation of any weighting function with respect to the filtration $\mathscr{F}_T$

$$\pi_0 \quad = \quad E[weight|\mathscr{F}_T].$$

Proof. Let π be a weighting function. The Greek ratio can be expressed as the expected value of the scalar product of the discounted payoff denoted by ϕ, i.e.

$$Greek \quad = \quad E[\phi \cdot \pi].$$

The variance V of this estimator is given by

$$V \quad = \quad E[(\phi \cdot \pi - Greek)^2].$$

We can introduce the conditional expectation π_0, leading to

$$
\begin{aligned}
V &= E[(\phi \cdot (\pi - \pi_0) + \phi \cdot \pi_0 - Greek)^2] \\
&= E[(\phi \cdot (\pi - \pi_0))^2] + E[(\phi \cdot \pi_0 - Greek)^2] + 2E[(F \cdot (\pi - \pi_0))(F\pi_0 - Greek)].
\end{aligned}
$$

But indeed the last term in the equation above is equal to zero since

$$
\begin{aligned}
E[(F \cdot (\pi - \pi_0))(F\pi_0 - Greek)] &= E\Big[E[(F \cdot (\pi - \pi_0))(F\pi_0 - Greek) \mid \mathscr{F}_T]\Big] \\
&= E\Big[E[F \cdot (\pi - \pi_0) \mid \mathscr{F}_T](F\pi_0 - Greek)\Big] \\
&= 0
\end{aligned}
$$

where we have used first the fact that $(F\pi_0 - Greek)$ and F are $\mathscr{F}_T$ measurable and therefore $E[F \cdot (\pi - \pi_0) \mid \mathscr{F}_T] = 0$. ■

18.5 Malliavin Calculus

Let us begin just introducing some useful properties necessary to our computation.

1. **Chain Rule**: Since D operators on random variables by differentiating functions in the form of partial derivatives, it shares the familiar chain rule property,

$$D_t f(X) = \nabla f(X) D_t X = f'(X) D_t X. \tag{18.5.1}$$

2. **Duality Principle(Integration by Parts)**: Let consider a random variable of the form $F = f(W_{t_1}, \cdots, W_{t_m})$, where f is smooth enough on its arguments and a stochastic process $u = \{u_t\}_{t\in[0,T]}$. The Malliavin derivative operator$D$ and Skorohod integral operator δ are linked by the following duality principle:

$$E\left[\int_0^T (D_t F)\, u_t dt\right] = E[F\delta(u)]. \tag{18.5.2}$$

3. If u_t is an adapted process, then the Skorohod integral coincides with the classical Ito integral: i.e. $\delta(u) = \int_0^T u_t dW_t$. If u_t is not-adapted, one has

$$\delta(Fu) = F\delta(u) - \int_0^T (D_t F) u_t dt. \tag{18.5.3}$$

The property (18.5.3) follows directly from the duality relation (18.5.2) and the product rule of the operator D. Let us assume that F and G are any two smooth random variables, and u_t a generic process, then by the product rule of D one has

$$\begin{aligned} E[GF\delta(u)] &= E\left[\int_0^T D_t(GF) u_t dt\right] \quad \text{by (18.5.2)} \\ &= E\left[\int_0^T G D_t(F) u_t dt\right] + E\left[\int_0^T D_t(G) F u_t dt\right] \\ &= E\left[G\int_0^T D_t(F) u_t dt\right] + E[G\delta(Fu)] \end{aligned}$$

which implies that

$$E[G\delta(Fu)] = E\left[G\left(F\delta(u) - \int_0^T (D_t F) u_t dt\right)\right]$$

for any random variable G. Therefore (18.5.3) must hold almost everywhere.

Example 18.5.1.

1. $\delta(1) = \int_0^T dW_t = W_T$.
2. $\delta(W_T)$: With help of (18.5.3) applied to $u = 1$ and $F = W_T$, we have

$$\delta(W_T) = W_T\ \delta(1) - \int_0^T D_t W_T dt$$

$$
\begin{aligned}
&= W_T^2 - \int_0^T 1dt \\
&= W_T^2 - T.
\end{aligned}
$$

3. $\delta(S_T)$: With $u = 1$ and $F = S_T$, we have

$$
\begin{aligned}
\delta(S_T) &= S_T\ \delta(1) - \int_0^T D_t S_T dt \\
&= S_T W_T - \int_0^T \sigma S_T dt \\
&= S_T W_T - \sigma S_T T \\
&= S_T(W_T - \sigma T). \qquad (18.5.4)
\end{aligned}
$$

4. $\delta\left(e^{\sigma W_T + b}\right)$: Assume σ and b are constant.

$$
\begin{aligned}
\delta\left(e^{\sigma W_T+b}\right) &= e^{\sigma W_T+b} W_T - \int_0^T \sigma e^{\sigma W_T+b}\, dt \\
&= e^{\sigma W_T+b}(W_T - \sigma T), \\
\delta\left(\delta\left(e^{\sigma W_T+b}\right)\right) &= \delta\left(e^{\sigma W_T+b}\right) W_T - \int_0^T \left(\sigma\delta\left(e^{\sigma W_T+b}\right) + e^{\sigma W_T+b}\right) dt \\
&= \delta\left(e^{\sigma W_T+b}\right) W_T - \left(\sigma\delta\left(e^{\sigma W_T+b}\right) + e^{\sigma W_T+b}\right) T \\
&= \delta\left(e^{\sigma W_T+b}\right)(W_T - \sigma T) - Te^{\sigma W_T+b}. \qquad (18.5.5)
\end{aligned}
$$

5. $\delta_1(W_2(T))$: Let $W_1(t)$ and $W_2(t)$ be independent Brownian motions.

$$
\begin{aligned}
\delta_1(W_2(T)) &= W_2(T)\delta_1(1) - \int_0^T D_t^1(W_2(T))dt \\
&= W_2(T)W_1(T) - \int_0^T 0\, dt \\
&= W_2(T)W_1(T).
\end{aligned}
$$

6. For $0 < t_1 < t_2$, let $\Delta W_{t_2} = W_{t_2} - W_{t_1}$. The Skorohod integral $\delta\big(\Delta W_{t_2} \cdot \mathbf{1}_{\{t_1 \le t < t_2\}}\big)$ is given by

$$
\begin{aligned}
\delta(\Delta W_{t_2}\mathbf{1}_{\{t_1 \le t < t_2\}}) &= \Delta W_{t_2}\delta(\mathbf{1}_{\{t_1 \le t < t_2\}}) - \int_{t_1}^{t_2} D_t(\Delta W_{t_2})dt \\
&= (\Delta W_{t_2})^2 - (t_2 - t_1).
\end{aligned}
$$

□

✎ **Malliavin Weight**

Let us assume that we are trying to computer $E[f'(X)Y]$, where X and Y are two random variables(X : $1\times n, Y : n\times n$). Suppose now that we are not able to solve this problem directly, for instant, because the join probability density function of X and Y is unknown. Moreover, the possibility of obtaining the answer through simulation is not a suitable or even feasible way, due to the presence of the derivative on function f. In this scenario we could use the formalism we have presented so far, in order to remove the derivative from f:

$$E[f'(X)Y] = E[f(X)H_{XY}], \tag{18.5.6}$$

in exchange for the appearing of some new random variable H_{XY}.

Let us begin just applying the D operator on $Z := f(X)$. The chain rule (18.5.1) dictates that

$$D_t Z = f'(X) D_t X.$$

Now we need to add Y to this expression. Thus, we multiply the above by $h(s)Y$ where h is an $n \times n$ arbitrary process-maybe depending on X, Y or even on another random variable-to be chosen appropriately. Then we have

$$D_t Z\ h(t)Y = \nabla f(X)\ D_t X\ h(t)Y.$$

Integrating this for $t \in [0, T]$, we have that

$$\int_0^T D_t Z\ h(t) Y dt = \int_0^T \nabla f(X)\ D_t X\ h(t) Y dt = \nabla f(X) Y\ Y^{-1} \int_0^T D_t X\ h(t)\, dt\ Y,$$

then

$$\begin{aligned}
\nabla f(X)\ Y &= \int_0^T D_t Z\ h(t) Y \left(Y^{-1} \int_0^T D_s X\ h(s)\, ds\ Y \right)^{-1} dt \\
&= \int_0^T D_t Z\ h(t) \left(\int_0^T D_s X\ h(s)\, ds \right)^{-1} Y dt,
\end{aligned}$$

and therefore

$$E[\nabla f(X)\ Y] = E\left[\int_0^T (D_t Z) u_t dt \right],$$

with

$$u_t = h(t) \left(\int_0^T D_s X\ h(s)\, ds \right)^{-1} Y.$$

Then, by the duality principle, we have that

$$E[\nabla f(X)\ Y] = E\left[f(X) \delta \left(h(t) \left(\int_0^T D_s X\ h(s)\, ds \right)^{-1} Y \right) \right]$$

and thus we achieve the formula we were looking for, with H_{XY} equal to

$$H_{XY} = \delta \left(h(t) \left(\int_0^T D_s X\ h(s)\, ds \right)^{-1} Y \right). \tag{18.5.7}$$

In particular for h equal to a constant we have that

$$H_{XY} = \delta \left(\left(\int_0^T D_s X\, ds \right)^{-1} Y \right). \tag{18.5.8}$$

✎ **Variations in the Initial Condition**

In the following, we provide an express of the derivatives of the expectation $u(x)$ with respect to the initial condition $x \in \mathbb{R}^n$ in the form of a weighted expectation of the same functional. The payoff ϕ is now a mapping from $(\mathbb{R}^n)^m$ into $\mathbb{R}$ with

$$E[\phi(X(t_1), \cdots, X(t_m))^2] < \infty, \quad X(t_i) \in \mathbb{R}^n$$

for a given integer $m \geq 1$ and $0 < t_1 \leq \cdot \leq t_m \leq T$, where $E^x[\cdot] = E[\cdot|X(0) = x]$. The expectation of interest is

$$u(x) \quad = \quad E^x[\phi(X(t_1), \cdots, X(t_m))].$$

We introduce the set Γ_m defined by

$$\Gamma_m \quad = \quad \left\{ a \in L^2[0, T] \;\Big|\; \int_0^{t_i} a(t)dt = 1, \forall i = 1, \cdots, m \right\}.$$

Theorem 18.5.2. Under the uniform ellipticity conditions of $\sigma(X(t))$, for any $x \in \mathbb{R}^n$ and for any $a \in \Gamma_m$, we have

$$\nabla u(x) \quad = \quad E^x \left[\phi(X(t_1), \cdots, X(t_m)) \int_0^T a(t) \left[\sigma^{-1}(X(t))Y(t) \right]^* dW(t) \right]$$

where $*$ denote the Skorohod integral. In general, this weight is not a function of $X(T)$ alone, hence not optimal. Conditioning with respect to $X(T)$ will yield the optimal weight.

18.6 The European-Style Options

The underlying stock price follows geometric Brownian motion:

$$dS_t = S_t\Big(\lambda dt + \sigma dW_t\Big), \quad \lambda = r - d.$$

For the payoff function ϕ, the discounted expectation is given by

$$u(x) = E^x[e^{-rT}\phi(S_T)].$$

✎ **Delta(Δ)**
Now we computer delta, Δ, the first-order partial differential sensitivity coefficient of the ϕ with respect to $S_0 = x$.

1. Using (18.5.8): We have

$$\begin{aligned}\Delta &= \frac{\partial}{\partial x}\phi(0,x) \\ &= e^{-rT}E\left[\phi'(S_T)\,\frac{\partial S_T}{\partial x}\right] \\ &= e^{-rT}E\left[\phi'(S_T)\,\frac{S_T}{x}\right] \\ &= \frac{e^{-rT}}{x}E\left[\phi'(S_T)\,S_T\right].\end{aligned}$$

   Then, with $X = Y = S_T$ in (18.5.6), we perform the integration by parts applying (18.5.8) to give

$$\begin{aligned}H_{XY} &= \delta\left(\frac{S_T}{\int_0^T D_t S_T dt}\right) \\ &= \delta\left(\frac{S_T}{\int_0^T \sigma S_T dt}\right) \quad \text{by(18.4.6)} \\ &= \delta\left(\frac{S_T}{\sigma T S_T}\right) \\ &= \delta\left(\frac{1}{\sigma T}\right) \\ &= \frac{W_T}{\sigma T}.\end{aligned}$$

   Hence we have

$$\Delta = e^{-rT}E\left[\phi(S_T)\,\frac{W_T}{\sigma x T}\right] = e^{-rT}E\left[\phi(S_T)\,\frac{W_T}{\sigma S_0 T}\right].$$

2. Using Theorem (18.5.2): For the delta, we have to compute an Ito stochastic integral

$$\int_0^T a(t)\frac{Y_t}{\sigma S_t}dW_t$$

where a must satisfy $\int_0^T a(t)dt = 1$. A trivial choice is $a(t) = \frac{1}{T}, \forall 0 \leq t \leq T$. Then we get the following formula:

$$
\begin{aligned}
\Delta &= e^{-rT} E\left[\phi(S_T)\int_0^T \frac{Y_t}{\sigma T S_t} dW_t\right] \\
&= e^{-rT} E\left[\phi(S_T)\int_0^T \frac{S_t/S_0}{\sigma T S_t} dW_t\right] \\
&= e^{-rT} E\left[\phi(S_T)\int_0^T \frac{1}{\sigma T S_0} dW_t\right] \\
&= e^{-rT} E\left[\phi(S_T)\frac{W_T}{\sigma T S_0}\right].
\end{aligned}
$$

Gamma(Γ)

The Greek gamma, Γ, involves a second-order derivative,

$$
\begin{aligned}
\Gamma &= \frac{\partial^2}{\partial x^2} E\left[e^{-rT}\phi(S_T)\right] \\
&= \frac{\partial}{\partial x} E\left[e^{-rT}\phi'(S_T)\frac{S_T}{x}\right] \\
&= E\left[e^{-rT}\phi''(S_T)\left(\frac{S_T}{x}\right)^2\right] + E\left[e^{-rT}\phi'(S_T)\cancelto{0}{\frac{\partial}{\partial x}\left(\frac{S_T}{x}\right)}\right] \\
&= \frac{e^{-rT}}{x^2}\ E\left[\phi''(S_T)S_T^2\right].
\end{aligned}
$$

By (18.5.6) and (18.5.8) with $X = S_T$ and $Y = S_T^2$, we obtain

$$
\begin{aligned}
\Gamma &= \frac{e^{-rT}}{x^2}\ E\left[\phi'(S_T)\delta\left(\frac{S_T^2}{\int_0^T D_t S_T dt}\right)\right] \\
&= \frac{e^{-rT}}{x^2}\ E\left[\phi'(S_T)\delta\left(\frac{S_T^2}{\sigma T S_T}\right)\right] \\
&= \frac{e^{-rT}}{x^2}\ E\left[\phi'(S_T)\delta\left(\frac{S_T}{\sigma T}\right)\right] \\
&= \frac{e^{-rT}}{x^2}\ E\left[\phi'(S_T)\ S_T\left(\frac{W_T}{\sigma T} - 1\right)\right], \quad \text{by (18.5.4).}
\end{aligned}
$$

Then repeated application of (18.5.6) and (18.5.8) with $X = S_T$ and $Y = S_T(W_T/\sigma T - 1)$, the second integration by parts yields

$$
\begin{aligned}
\Gamma &= \frac{e^{-rT}}{x^2}\ E\left[\phi(S_T)\ \delta\left\{\frac{S_T}{\int_0^T D_t S_T dt}\left(\frac{W_T}{\sigma T} - 1\right)\right\}\right] \\
&= \frac{e^{-rT}}{x^2}\ E\left[\phi(S_T)\ \delta\left\{\frac{1}{\sigma T}\left(\frac{W_T}{\sigma T} - 1\right)\right\}\right] \\
&= \frac{e^{-rT}}{x^2}\ E\left[\phi(S_T)\ \frac{1}{\sigma T}\left(\frac{W_T^2 - T}{\sigma T} - W_T\right)\right]
\end{aligned}
$$

$$= \frac{e^{-rT}}{\sigma T x^2} E\left[\phi(S_T)\ \left(\frac{W_T^2}{\sigma T} - W_T - \frac{1}{\sigma}\right)\right].$$

Alternatively we could obtain the gamma from

$$\begin{aligned}
\Gamma &= \frac{\partial}{\partial x}\Delta \\
&= e^{-rT}\frac{\partial}{\partial x}E\left[\phi(S_T)\frac{W_T}{\sigma T x}\right] \\
&= e^{-rT}\left(E\left[\phi'(S_T)\frac{S_T W_T}{\sigma T x^2}\right] - E\left[\phi(S_T)\frac{W_T}{\sigma T x^2}\right]\right) \\
&= \frac{e^{-rT}}{\sigma T x^2}\left(E\left[\phi(S_T)\delta\left(\frac{S_T W_T}{\int_0^T S_T dt}\right)\right] - E\left[\phi(S_T)W_T\right]\right) \\
&= \frac{e^{-rT}}{\sigma T x^2}\left(E\left[\phi(S_T)\delta\left(\frac{W_T}{\sigma T}\right)\right] - E\left[\phi(S_T)W_T\right]\right) \\
&= \frac{e^{-rT}}{\sigma T x^2}\left(E\left[\phi(S_T)\left(\frac{W_T^2 - T}{\sigma T}\right)\right] - E\left[\phi(S_T)W_T\right]\right) \\
&= \frac{e^{-rT}}{\sigma T x^2} E\left[\phi(S_T)\ \left(\frac{W_T^2}{\sigma T} - W_T - \frac{1}{\sigma}\right)\right].
\end{aligned}$$

✎ **Vega($\mathcal{V}$)**

The vega can be computed as

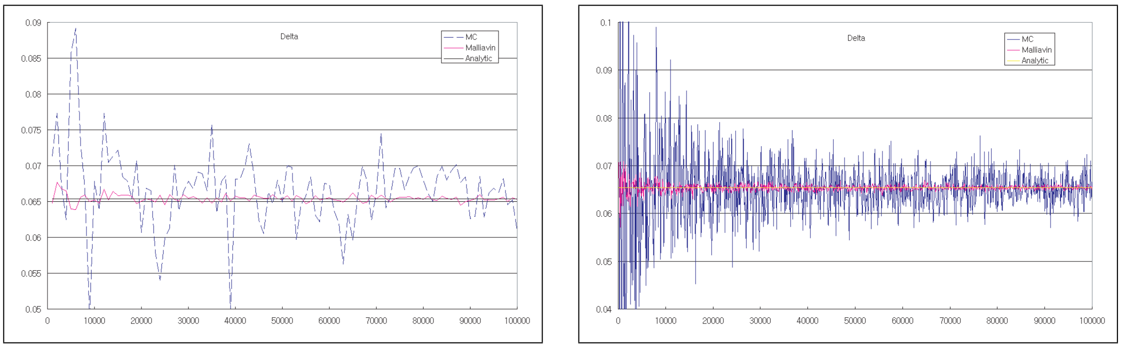
$$\begin{aligned}
\mathcal{V} &= \frac{\partial u}{\partial \sigma} \\
&= e^{-rT}E\left[\phi'(S_T)\frac{\partial S_T}{\partial \sigma}\right] \\
&= e^{-rT}E\left[\phi'(S_T)\ S_T(W_T - \sigma T)\right] \\
&= e^{-rT}E\left[\phi(S_T)\ \delta\left(\frac{S_T(W_T - \sigma T)}{\sigma T S_T}\right)\right] \\
&= e^{-rT}E\left[\phi(S_T)\ \delta\left(\frac{W_T - \sigma T}{\sigma T}\right)\right] \\
&= e^{-rT}E\left[\phi(S_T)\ \left(\frac{W_T^2}{\sigma T} - \frac{1}{\sigma} - W_T\right)\right].
\end{aligned}$$

✎ **Rho(ρ)**

The rho is given by

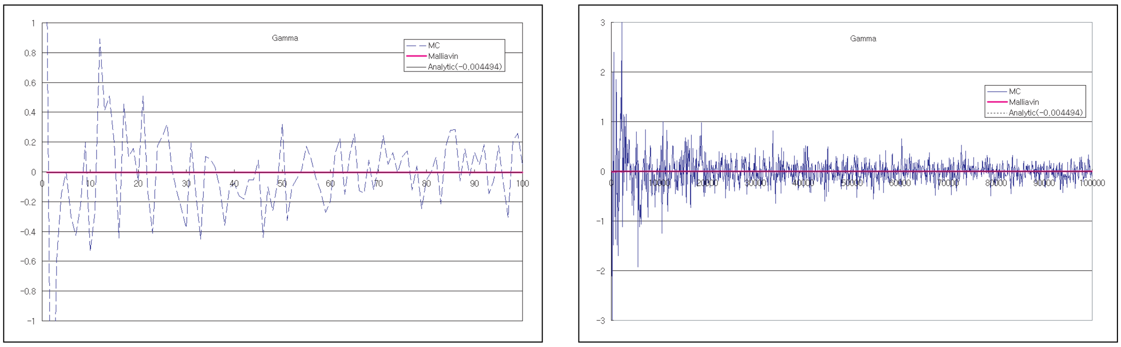
$$\begin{aligned}
\rho &= -Te^{-rT}E\left[\phi(S_T)\right] + e^{-rT}E\left[\phi'(S_T)TS_T\right] \\
&= -Te^{-rT}E\left[\phi(S_T)\right] + e^{-rT}E\left[\phi(S_T)\ \delta\left(\frac{TS_T}{\sigma T S_T}\right)\right] \\
&= e^{-rT}E\left[\phi(S_T)\ \left(-T + \frac{W_T}{\sigma}\right)\right].
\end{aligned}$$

Figure(18.1) and (18.2) demonstrate the efficiency of the Malliavin scheme for the computation of delta and gamma of a digital option.



(a) $1,000 \leq N \leq 10,000$ (b) $1,000 \leq N \leq 100,000$

Figure 18.1: Delta for digital option: $S_0 = 40, K = 40, r = 10\%, \sigma = 20\%, Div = 1\%, T = 0.5$, $N =$ number of simulations.



(a) $1,000 \leq N \leq 10,000$ (b) $1,000 \leq N \leq 100,000$

Figure 18.2: Gamma for digital option: $S_0 = 40, K = 40, r = 10\%, \sigma = 20\%, Div = 1\%, T = 0.5$, $N =$ number of simulations.

Example 18.6.1 (Double Asset Option(continued)). In this example we will calculate delta and gamma of option with two underlying assets. First, let us calculate delta using Theorem(18.5.2). First we have

$$\begin{aligned}
\sigma^{-1}(X(t)) &= \begin{pmatrix} \frac{1}{\sigma_1 X_1(t)} & 0 \\ -\frac{\rho}{\sqrt{1-\rho^2}\sigma_1 X_1(t)} & \frac{1}{\sqrt{1-\rho^2}\sigma_2 X_2(t)} \end{pmatrix} \\
\sigma^{-1}(X(t))Y(t) &= \begin{pmatrix} \frac{1}{\sigma_1 X_1(t)} & 0 \\ -\frac{\rho}{\sqrt{1-\rho^2}\sigma_1 X_1(t)} & \frac{1}{\sqrt{1-\rho^2}\sigma_2 X_2(t)} \end{pmatrix} \begin{pmatrix} \frac{X_1(t)}{X_1(0)} & 0 \\ 0 & \frac{X_2(t)}{X_2(0)} \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{\sigma_1 X_1(0)} & 0 \\ -\frac{\rho}{\sqrt{1-\rho^2}\sigma_1 X_1(0)} & \frac{1}{\sqrt{1-\rho^2}\sigma_2 X_2(0)} \end{pmatrix}.
\end{aligned}$$

For the expectation of the payoff ϕ, if we choose $a(t) = \frac{1}{T}$ we have that

$$\nabla u = E\left[\phi(X_1(T), X_2(T))\frac{1}{T}\int_0^T \left[\sigma^{-1}(X(t))Y(t)\right]^* dW(t)\right]$$

$$
\begin{aligned}
&= E\left[\phi(X_1(T),X_2(T))\frac{1}{T}\int_0^T \begin{pmatrix} \frac{1}{\sigma_1 X_1(0)} & 0 \\ -\frac{\rho}{\sqrt{1-\rho^2}\sigma_1 X_1(0)} & \frac{1}{\sqrt{1-\rho^2}\sigma_2 X_2(0)} \end{pmatrix}^* \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix}\right] \\
&= E\left[\phi(X_1(T),X_2(T))\frac{1}{T}\int_0^T \begin{pmatrix} \frac{1}{\sigma_1 X_1(0)} & 0 \\ -\frac{\rho}{\sqrt{1-\rho^2}\sigma_1 X_1(0)} & \frac{1}{\sqrt{1-\rho^2}\sigma_2 X_2(0)} \end{pmatrix}^T \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix}\right] \\
&= E\left[\phi(X_1(T),X_2(T))\frac{1}{T}\int_0^T \begin{pmatrix} \frac{dW_1(t)}{\sigma_1 X_1(0)} - \frac{\rho dW_2(t)}{\sqrt{1-\rho^2}\sigma_1 X_1(0)}, & \frac{dW_2(t)}{\sqrt{1-\rho^2}\sigma_2 X_2(0)} \end{pmatrix}^T\right] \\
&= E\left[\phi(X_1(T),X_2(T))\frac{1}{T} \begin{pmatrix} \frac{W_1(T)}{\sigma_1 X_1(0)} - \frac{\rho W_2(T)}{\sqrt{1-\rho^2}\sigma_1 X_1(0)}, & \frac{W_2(T)}{\sqrt{1-\rho^2}\sigma_2 X_2(0)} \end{pmatrix}^T\right].
\end{aligned}
$$

Now let us calculate gamma. Let $x = X_1(0), y = X_2(0)$, $w_1 = \frac{1}{x\sigma_1 T}\left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}}W_2(T)\right)$ and $w_2 = \frac{1}{y\sigma_2 T}\frac{1}{\sqrt{1-\rho^2}}W_2(T)$. Then we have

$$
\begin{aligned}
&\nabla\left(\frac{\partial u}{\partial x}\right) \\
&= \left(\frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial y \partial x}\right) \\
&= \nabla E\left[\phi(X_1(T),X_2(T))\ w_1\right] \\
&= E\left[\nabla\phi(X_1(T),X_2(T))\ w_1\right] + E\left[\phi(X_1(T),X_2(T))\ \nabla w_1\right] \\
&= E\left[\phi(X_1(T),X_2(T))\frac{1}{T}\int_0^T \left[\sigma(X(t))^{-1}Y(t)w_1\right]^* dW(t)\right] + E\left[\phi(X_1(T),X_2(T))\ \nabla w_1\right].
\end{aligned}
$$

The last term can be expressed as

$$
\nabla w_1 = \left(-\frac{1}{x^2\sigma_1 T}\left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}}W_2(T)\right), \quad 0 \right) = \left(-\frac{1}{x}w_1, 0\right)
$$

and the integrand in the Skorohod integral

$$
\begin{aligned}
&\frac{1}{T}\left(\sigma^{-1}(X(t))Y(t)w_1\right)^T \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix} \\
&= \frac{1}{T}\begin{pmatrix} \frac{w_1 dW_1(t)}{\sigma_1 X_1(0)} - \frac{\rho w_1 dW_2(t)}{\sqrt{1-\rho^2}\sigma_1 X_1(0)}, & \frac{w_1 dW_2(t)}{\sqrt{1-\rho^2}\sigma_2 X_2(0)} \end{pmatrix}^T.
\end{aligned}
$$

Hence the weight for $\frac{\partial^2 u}{\partial x^2}$ is given by

$$
\begin{aligned}
&\frac{1}{x\sigma_1 T}\left\{\delta_1(w_1) - \delta_2\left(\frac{\rho}{\sqrt{1-\rho^2}}w_1\right) - \sigma_1 T w_1\right\} \\
&= \frac{1}{x^2\sigma_1^2 T^2}\left\{\delta_1\left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}}W_2(T)\right) - \frac{\rho}{\sqrt{1-\rho^2}}\delta_2\left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}}W_2(T)\right) - \sigma_1 T\left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}}W_2(T)\right)\right\} \\
&= \frac{1}{x^2\sigma_1^2 T^2}\left\{\left(W_1^2(T) - T - \frac{\rho}{\sqrt{1-\rho^2}}W_1(T)W_2(T)\right) - \frac{\rho}{\sqrt{1-\rho^2}}\left(W_1(T)W_2(T) - \frac{\rho}{\sqrt{1-\rho^2}}(W_2^2(T) - T)\right)\right\} \\
&\quad -\frac{1}{x^2\sigma_1 T}\left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}}W_2(T)\right) \\
&= \frac{1}{x^2\sigma_1^2 T^2}\left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}}W_2(T)\right)^2 - \frac{1}{x^2\sigma_1 T}\left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}}W_2(T)\right) - \frac{1}{x^2\sigma_1^2 T(1-\rho^2)}. \qquad (18.6.1)
\end{aligned}
$$

Similarly the weight for $\frac{\partial^2 u}{\partial x \partial y}$ is given by

$$
\begin{aligned}
& \frac{1}{y\sigma_2 T\sqrt{1-\rho^2}}\ \delta_2(w_1) \\
= \quad & \frac{1}{y\sigma_2 T\sqrt{1-\rho^2}} \frac{1}{x\sigma_1 T}\ \delta_2\left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}} W_2(T)\right) \\
= \quad & \frac{1}{xy\sigma_1\sigma_2 T^2\sqrt{1-\rho^2}} \left(W_1(T)W_2(T) - \frac{\rho}{\sqrt{1-\rho^2}}\left(W_2^2(T) - T\right)\right). \qquad (18.6.2)
\end{aligned}
$$

Now let us calculate $\nabla\left(\frac{\partial u}{\partial y}\right)$.

$$
\begin{aligned}
& \nabla\left(\frac{\partial u}{\partial y}\right) \\
= \quad & \left(\frac{\partial^2 u}{\partial x \partial y}, \frac{\partial^2 u}{\partial y^2}\right) \\
= \quad & \nabla E\left[\phi(X_1(T), X_2(T))\ w_2\right] \\
= \quad & E\left[\nabla\phi(X_1(T), X_2(T))\ w_2\right] + E\left[\phi(X_1(T), X_2(T))\ \nabla w_2\right] \\
= \quad & E\left[\phi(X_1(T), X_2(T))\frac{1}{T}\int_0^T \left[\sigma(X(t))^{-1}Y(t)w_2\right]^* dW(t)\right] + E\left[\phi(X_1(T), X_2(T))\ \nabla w_2\right].
\end{aligned}
$$

Then we have $\nabla w_2 = \left(0, -\frac{1}{y}w_2\right)$ and

$$
\begin{aligned}
& \frac{1}{T}\left(\sigma^{-1}(X(t))Y(t)w_2\right)^T \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix} \\
= \quad & \frac{1}{T}\begin{pmatrix} \frac{w_2 dW_1(t)}{\sigma_1 X_1(0)} - \frac{\rho w_2 dW_2(t)}{\sqrt{1-\rho^2}\sigma_1 X_1(0)}, & \frac{w_2 dW_2(t)}{\sqrt{1-\rho^2}\sigma_2 X_2(0)} \end{pmatrix}^T.
\end{aligned}
$$

Hence the weight for $\frac{\partial^2 u}{\partial x \partial y}$ is given by

$$
\begin{aligned}
& \frac{1}{x\sigma_1 T}\left(\delta_1(w_2) - \frac{\rho}{\sqrt{1-\rho^2}}\ \delta_2(w_2)\right) \\
= \quad & \frac{1}{x\sigma_1 T}\ \frac{1}{y\sigma_2 T\sqrt{1-\rho^2}}\left(W_1(T)W_2(T) - \frac{\rho}{\sqrt{1-\rho^2}}\left(W_2^2(T) - T\right)\right) \\
= \quad & \frac{1}{xy\sigma_1\sigma_2 T^2\sqrt{1-\rho^2}}\left(W_1(T)W_2(T) - \frac{\rho}{\sqrt{1-\rho^2}}\left(W_2^2(T) - T\right)\right).
\end{aligned}
$$

We have found again the same weight as in (18.6.2). The last weight for $\frac{\partial^2 u}{\partial y^2}$ can be obtained from

$$
\begin{aligned}
& \frac{1}{y\sigma_2 T\sqrt{1-\rho^2}}\delta_2(w_2) - \frac{1}{y}w_2 \\
= \quad & \frac{1}{y\sigma_2 T\sqrt{1-\rho^2}}\frac{1}{y\sigma_2 T\sqrt{1-\rho^2}}\delta_2(W_2(T)) - \frac{1}{y^2\sigma_2 T\sqrt{1-\rho^2}}W_2(T) \\
= \quad & \frac{1}{y^2\sigma_2^2 T^2(1-\rho^2)}(W_2^2(T) - T) - \frac{1}{y^2\sigma_2 T\sqrt{1-\rho^2}}W_2(T)
\end{aligned}
$$

$$= \frac{1}{y^2\sigma_2^2 T^2(1-\rho^2)}\Big(W_2^2(T) - T - \sigma_2 T\sqrt{1-\rho^2}W_2(T)\Big).$$

(to be continued)

□

Example 18.6.2 (Double Asset Option(continued)). In this example we will calculate vega of option with two underlying assets. First let us introduce Z_t^σ as follows.

$$\begin{aligned} Z_t^\sigma &:= \begin{bmatrix} \frac{\partial X_1(t)}{\partial\sigma_1} & \frac{\partial X_1(t)}{\partial\sigma_2} \\ \frac{\partial X_2(t)}{\partial\sigma_1} & \frac{\partial X_2(t)}{\partial\sigma_2} \end{bmatrix} \\ &= \begin{bmatrix} X_1(t)(-\sigma_1 t + W_1(t)) & 0 \\ 0 & X_2(t)(-\sigma_2 t + \rho W_1(t) + \sqrt{1-\rho^2}W_2(t)) \end{bmatrix} \end{aligned}$$

The vega can be obtained from

$$\begin{aligned} \mathcal{V} &\overset{\text{why?}}{=} E\left[\phi(X_1(T),X_2(T))\frac{1}{T}\delta\left(\sigma^{-1}(X(T))Z_T^\sigma\right)\right] \\ &= E\left[\phi\cdot\frac{1}{T}\,\delta\left(\begin{array}{cc} \frac{1}{\sigma_1}(W_1(T)-\sigma_1 T) & 0 \\ -\frac{1}{\sigma_1}\frac{\rho}{\sqrt{1-\rho^2}}(W_1(T)-\sigma_1 T) & \frac{1}{\sigma_2}\frac{1}{\sqrt{1-\rho^2}}(-\sigma_2 T + \rho W_1(T) + \sqrt{1-\rho^2}W_2(T)) \end{array}\right)\right] \end{aligned}$$

Hence we get that

$$\begin{aligned} \frac{\partial u}{\partial\sigma_1} &= E\left[\phi(X_1(T),X_2(T))\;\frac{1}{\sigma_1 T}\;\left(\delta_1\Big(W_1(T)-\sigma_1 T\Big) - \frac{\rho}{\sqrt{1-\rho^2}}\;\delta_2\Big(W_1(T)-\sigma_1 T\Big)\right)\right] \\ &= E\left[\phi(X_1(T),X_2(T))\;\frac{1}{\sigma_1 T}\;\left(\Big(W_1^2(T) - T - \sigma_1 T W_1(T)\Big) - \frac{\rho}{\sqrt{1-\rho^2}}\;\Big(W_1(T)W_2(T) - \sigma_1 T W_2(T)\Big)\right)\right] \\ &= E\left[\phi(X_1(T),X_2(T))\;\left\{-\frac{1}{\sigma_1} - \left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}}W_2(T)\right) + \frac{1}{\sigma_1 T}\left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}}W_2(T)\right)W_1(T)\right\}\right] \end{aligned}$$

and

$$\begin{aligned} \frac{\partial u}{\partial\sigma_2} &= E\left[\phi(X_1(T),X_2(T))\frac{1}{\sigma_2 T\sqrt{1-\rho^2}}\delta_2\left(-\sigma_2 T + \rho W_1(T) + \sqrt{1-\rho^2}W_2(T)\right)\right] \\ &= E\left[\phi(X_1(T),X_2(T))\frac{1}{\sigma_2 T\sqrt{1-\rho^2}}\left(-\sigma_2 T W_2(T) + \rho W_1(T)W_2(T) + \sqrt{1-\rho^2}\left(W_2^2(T) - T\right)\right)\right] \\ &= E\left[\phi(X_1(T),X_2(T))\left\{-\frac{1}{\sigma_2} - \frac{W_2(T)}{\sqrt{1-\rho^2}} + \frac{1}{\sigma_2 T}\left(\frac{\rho}{\sqrt{1-\rho^2}}W_1(T) + W_2(T)\right)W_2(T)\right\}\right] \end{aligned}$$

This results coincides with Example(18.2.7). Verify this!!! □

Remark 18.6.3. Let $h(t) = \sigma^{-1}(X(T))$. Then by (18.5.7) we have the Malliavin weight for vega as

$$\begin{aligned} \mathcal{V} &= E\left[\nabla\phi(X(T))Z_T^\sigma\right] \\ &= E\left[\phi(X(T))\delta\left(\sigma^{-1}(X(T))\left(\int_0^T D_s X_T\sigma^{-1}(X(T))ds\right)^{-1}Z_T^\sigma\right)\right] \end{aligned}$$

$$
\begin{aligned}
&= E\left[\phi(X(T))\delta\left(\sigma^{-1}(X(T))\left(\int_0^T \mathbf{1}\, ds\right)^{-1} Z_T^\sigma\right)\right] \quad \text{by (18.4.7)} \\
&= E\left[\phi(X(T))\frac{1}{T}\delta\left(\sigma^{-1}(X(T))Z_T^\sigma\right)\right].
\end{aligned}
$$

□

Example 18.6.4. For the rho, define Z_t^ρ:

$$
\begin{aligned}
Z_t^\rho &:= \begin{bmatrix} \frac{\partial X_1(t)}{\partial r} \\ \frac{\partial X_2(t)}{\partial r} \end{bmatrix} \\
&= \begin{bmatrix} tX_1(t) \\ tX_2(t) \end{bmatrix}.
\end{aligned}
$$

Then we have

$$
\begin{aligned}
& E\left[\phi(X_1(T), X_2(T))\ \frac{1}{T}\ \delta\Big(\sigma^{-1}(X(T))Z_T^\rho\Big)\right] \\
&= E\left[\phi\ \cdot\ \delta\begin{pmatrix} \frac{1}{\sigma_1} \\ -\frac{\rho}{\sqrt{1-\rho^2}}\frac{1}{\sigma_1} + \frac{1}{\sqrt{1-\rho^2}}\frac{1}{\sigma_2} \end{pmatrix}\right] \\
&= E\left[\phi\ \cdot\ \left\{\delta_1\left(\frac{1}{\sigma_1}\right) + \delta_2\left(-\frac{\rho}{\sqrt{1-\rho^2}}\frac{1}{\sigma_1} + \frac{1}{\sqrt{1-\rho^2}}\frac{1}{\sigma_2}\right)\right\}\right] \\
&= E\left[\phi \cdot \left(\frac{W_1(T)}{\sigma_1} - \frac{\rho W_2(T)}{\sqrt{1-\rho^2}\sigma_1} + \frac{W_2(T)}{\sqrt{1-\rho^2}\sigma_2}\right)\right].
\end{aligned}
$$

Also, the (discounted) rho is given by

$$
\frac{\partial u}{\partial r} = -T\,u + xT\Delta_1 + yT\Delta_2.
$$

□

Example 18.6.5 (Multi-Asset Option-Delta&Gamma(continued from Example 18.4.7)). First, let us calculate $\sigma^{-1}(X(t))Y(t)$:

$$
\begin{aligned}
& \sigma^{-1}(X(t))Y(t) \\
&= \left[\begin{pmatrix} \sigma_1 X_1(t) & & & \mathbf{0} \\ & \sigma_2 X_2(t) & & \\ & & \ddots & \\ \mathbf{0} & & & \sigma_n X_n(t) \end{pmatrix}\Phi\right]^{-1} \begin{pmatrix} \frac{X_1(t)}{X_1(0)} & & & \mathbf{0} \\ & \frac{X_2(t)}{X_2(0)} & & \\ & & \ddots & \\ \mathbf{0} & & & \frac{X_n(t)}{X_n(0)} \end{pmatrix} \\
&= \Phi^{-1}\begin{pmatrix} \frac{1}{\sigma_1 X_1(0)} & & & \mathbf{0} \\ & \frac{1}{\sigma_2 X_2(0)} & & \\ & & \ddots & \\ \mathbf{0} & & & \frac{1}{\sigma_n X_n(0)} \end{pmatrix} = \Phi^{-1}A
\end{aligned}
$$

where

$$A := \begin{pmatrix} \frac{1}{\sigma_1 X_1(0)} & & & \mathbf{0} \\ & \frac{1}{\sigma_2 X_2(0)} & & \\ & & \ddots & \\ \mathbf{0} & & & \frac{1}{\sigma_n X_n(0)} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma_1 x_1} & & & \mathbf{0} \\ & \frac{1}{\sigma_2 x_2} & & \\ & & \ddots & \\ \mathbf{0} & & & \frac{1}{\sigma_n x_n} \end{pmatrix}.$$

From Theorem(18.5.2), if we choose $a(t) = \frac{1}{T}$ we have that

$$\begin{aligned} \nabla u &= E\left[\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)\frac{1}{T}\int_0^T \left[\sigma^{-1}(X(t))Y(t)\right]^* dW(t)\right] \\ &= E\left[\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)\frac{1}{T}\int_0^T \left[\Phi^{-1}A\right]^T dW(t)\right] \\ &= \frac{1}{T}E\left[\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)A\int_0^T (\Phi^{-1})^T dW(t)\right]. \end{aligned}$$

Let

$$(\Phi^{-1})^T = \begin{pmatrix} \varphi_{11} & \cdots & \varphi_{1n} \\ \vdots & \ddots & \vdots \\ \varphi_{n1} & \cdots & \varphi_{nn} \end{pmatrix}.$$

Then we have

$$\nabla u = \frac{1}{T}E\left[\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)\begin{pmatrix} \frac{1}{\sigma_1 X_1(0)}\left(\varphi_{11}W_1(T) + \cdots + \varphi_{1n}W_n(T)\right) \\ \vdots \\ \frac{1}{\sigma_n X_n(0)}\left(\varphi_{n1}W_1(T) + \cdots + \varphi_{nn}W_n(T)\right) \end{pmatrix}\right]$$

Now let us computer the Malliavin weight for gamma. Let w_i be the weight for $\frac{\partial u}{\partial x_i}$, i.e.

$$w_i = \frac{1}{\sigma_i x_i T}\sum_{l=1}^{n} \varphi_{il} W_l(T). \tag{18.6.3}$$

To get the weight for $\frac{\partial^2 u}{\partial x_j \partial x_i}$, we need the j-th component of

$$\begin{aligned} &\nabla E\left[\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)\ w_i\right] \\ = \ &E\left[\nabla\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)\ w_i\right] + E\left[\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)\nabla w_i\right]. \end{aligned} \tag{18.6.4}$$

The ∇w_i in (18.6.4) can be written as

$$\nabla w_i = \left(0, \cdots, 0, -\frac{1}{x_i}w_i, 0, \cdots, 0\right). \tag{18.6.5}$$

The first term in (18.6.4) can be writtne as

$$E\left[\nabla\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)\ w_i\right]$$

$$
\begin{aligned}
&= E\left[\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)\frac{1}{T}\int_0^T \left[\sigma(X(t))^{-1}Y(t)w_i\right]^* dW(t)\right] \\
&= E\left[\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)\frac{1}{T}\int_0^T w_i\left[\Phi^{-1}A\right]^T dW(t)\right] \\
&= E\left[\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)\frac{1}{T}\int_0^T w_i A\left(\Phi^{-1}\right)^T dW(t)\right]. \qquad (18.6.6)
\end{aligned}
$$

The sum of j-th component of (18.6.5) and (18.6.6) is the weight for $\frac{\partial^2 u}{\partial x_j \partial x_i}$. The j-th component of the skorohod integral part in (18.6.6) can be written as follows:

$$
\begin{aligned}
&\frac{1}{\sigma_i\sigma_j x_i x_j T^2}\sum_{k=1}^n \delta_k\left(\varphi_{jk}w_i\right) \\
&= \frac{1}{\sigma_i\sigma_j x_i x_j T^2}\sum_{k=1}^n \varphi_{jk}\delta_k\left(\sum_{l=1}^n \varphi_{il}W_l(T)\right) \\
&= \frac{1}{\sigma_i\sigma_j x_i x_j T^2}\sum_{k=1}^n\sum_{l=1}^n \varphi_{jk}\varphi_{il}\delta_k\Big(W_l(T)\Big). \qquad (18.6.7)
\end{aligned}
$$

Since

$$
\delta_k\Big(W_l(T)\Big) = \begin{cases} W_k^2(T) - T, & k = l \\ W_k W_l, & k \neq l, \end{cases}
$$

(18.6.7) can be written as

$$
\frac{1}{\sigma_i\sigma_j x_i x_j T^2}\sum_{k=1}^n\sum_{l=1}^n \varphi_{jk}\varphi_{il}\Big(W_k W_l - \delta_{kl}T\Big), \qquad (18.6.8)
$$

where δ_{kl} is Kronecker delta.

Example 18.6.6. For the case $n = 2$, we could apply the result of Example(18.6.5). Suppose that Φ is given as

$$
\Phi = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}.
$$

Then we have

$$
(\Phi^{-1})^T = \begin{pmatrix} 1 & -\frac{\rho}{\sqrt{1-\rho^2}} \\ 0 & \frac{1}{\sqrt{1-\rho^2}} \end{pmatrix}.
$$

Now let us calculate the Malliavin weights.

1. $\frac{\partial u}{\partial x_1}$: By (18.6.3) we have

$$
\begin{aligned}
w_1 &= \frac{1}{\sigma_1 x_1 T}\sum_{l=1}^2 \varphi_{1l}W_l(T) \\
&= \frac{1}{\sigma_1 x_1 T}\left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}}W_2(T)\right).
\end{aligned}
$$

2. $\frac{\partial u}{\partial x_2}$: Similarly we have

$$\begin{aligned} w_2 &= \frac{1}{\sigma_2 x_2 T} \sum_{l=1}^{2} \varphi_{1l} W_l(T) \\ &= \frac{1}{\sigma_2 x_2 T} \frac{1}{\sqrt{1-\rho^2}} W_2(T). \end{aligned}$$

3. $\frac{\partial^2 u}{\partial {x_1}^2}$: By (18.6.8), we have

$$\begin{aligned} & \frac{1}{\sigma_1^2 x_1^2 T^2} \sum_{k=1}^{2} \sum_{l=1}^{2} \varphi_{1k}\varphi_{1l}\Big(W_k W_l - \delta_{kl} T\Big) \quad - \quad \frac{1}{x_1} w_1 \\ = \quad & \frac{1}{\sigma_1^2 x_1^2 T^2} \left\{ W_1^2(T) - T - 2\frac{\rho}{\sqrt{1-\rho^2}} W_1(T) W_2(T) + \frac{\rho^2}{1-\rho^2}\Big(W_2^2(T) - T\Big) \right\} \quad - \quad \frac{1}{x_1} w_1 \\ = \quad & \frac{1}{\sigma_1^2 x_1^2 T^2} \left\{ \left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}} W_2(T) \right)^2 - T - \frac{\rho^2}{1-\rho^2} T \right\} \quad - \quad \frac{1}{x_1} w_1 \\ = \quad & \frac{1}{\sigma_1^2 x_1^2 T^2} \left\{ \left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}} W_2(T) \right)^2 - \frac{1}{1-\rho^2} T \right\} \quad - \quad \frac{1}{x_1} w_1 \\ = \quad & w_1^2 \quad - \quad \frac{1}{\sigma_1^2 x_1^2 T(1-\rho^2)} \quad - \quad \frac{1}{x_1} w_1. \end{aligned}$$

This result coincides with (18.6.1).

4. $\frac{\partial^2 u}{\partial x_2 \partial x_1}$: Similary we have

$$\begin{aligned} & \frac{1}{\sigma_1 \sigma_2 x_1 x_2 T^2} \sum_{k=1}^{2} \sum_{l=1}^{2} \varphi_{2k}\varphi_{1l}\Big(W_k W_l - \delta_{kl} T\Big) \\ = \quad & \frac{1}{\sigma_1 \sigma_2 x_1 x_2 T^2} \left\{ \frac{1}{\sqrt{1-\rho^2}} W_1(T) W_2(T) - \frac{\rho}{1-\rho^2}\Big(W_2^2(T) - T\Big) \right\}. \end{aligned}$$

5. $\frac{\partial^2 u}{\partial {x_2}^2}$: Similary we have

$$\begin{aligned} & \frac{1}{\sigma_2^2 x_2^2 T^2} \sum_{k=1}^{2} \sum_{l=1}^{2} \varphi_{2k}\varphi_{2l}\Big(W_k W_l - \delta_{kl} T\Big) \quad - \quad \frac{1}{x_2} w_2 \\ = \quad & \frac{1}{\sigma_2^2 x_2^2 T^2} \left\{ \frac{1}{1-\rho^2}\Big(W_2^2(T) - T\Big) \right\} \quad - \quad \frac{1}{x_2} w_2. \end{aligned}$$

Example 18.6.7 (Multi-Asset Option-Vega(continued from Example 18.6.5)). For the simplicity, define

$$D_n[a_i] \quad := \quad \begin{pmatrix} a_1 & & & \mathbf{0} \\ & a_2 & & \\ & & \ddots & \\ \mathbf{0} & & & a_n \end{pmatrix}.$$

Suppose that

$$Z_t^\sigma = D_n\left[X_i(t)\left(-\sigma_i t + \sum_{k=1}^n \epsilon_{ik} W_k(t)\right)\right].$$

Then we have

$$\begin{aligned}\sigma^{-1}(X(t))Z_t^\sigma &= \Phi^{-1} D_n\left[\frac{1}{\sigma_i X_i(t)}\right] D_n\left[X_i(t)\left(-\sigma_i t + \sum_{k=1}^n \epsilon_{ik} W_k(t)\right)\right] \\ &= \Phi^{-1} D_n\left[\frac{1}{\sigma_i}\left(-\sigma_i t + \sum_{k=1}^n \epsilon_{ik} W_k(t)\right)\right]\end{aligned}$$

Note that the vega is given by

$$\begin{aligned}\mathcal{V} &= E\left[\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)\frac{1}{T}\delta\Big(\sigma^{-1}(X(T))Z_T^\sigma\Big)\right] \\ &= E\left[\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)\frac{1}{T}\int_0^T \Big(\sigma^{-1}(X(T))Z_T^\sigma\Big)^* dW(t)\right] \\ &= E\left[\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)\frac{1}{T}\int_0^T \Big(\sigma^{-1}(X(T))Z_T^\sigma\Big)^T dW(t)\right] \\ &= E\left[\phi\Big(X_1(T), X_2(T), \cdots, X_n(T)\Big)\frac{1}{T}\int_0^T D_n\left[\frac{1}{\sigma_i}\left(-\sigma_i T + \sum_{k=1}^n \epsilon_{ik} W_k(T)\right)\right](\Phi^{-1})^T dW(t)\right]\end{aligned}$$

Hence the i-th component of weight for vega is given by

$$\begin{aligned}&\frac{1}{\sigma_i T}\sum_{j=1}^n \delta_j\left(\left(-\sigma_i T + \sum_{k=1}^n \epsilon_{ik} W_k(T)\right)\varphi_{ij}\right) \\ =\ &\frac{1}{\sigma_i T}\left\{\sum_{j=1}^n \delta_j(-\sigma_i T\varphi_{ij}) + \sum_{j=1}^n\sum_{k=1}^n \epsilon_{ik}\varphi_{ij}\delta_j(W_k(T))\right\} \\ =\ &-\sum_{j=1}^n \varphi_{ij} W_j(T) + \frac{1}{\sigma_i T}\sum_{j=1}^n\sum_{k=1}^n \epsilon_{ik}\varphi_{ij}\Big(W_j(T)W_k(T) - \delta_{jk}T\Big).\end{aligned} \tag{18.6.9}$$

Example 18.6.8 (continued from Example 18.6.6). By (18.6.9), we have the weight for vega as follows:

1. $\frac{\partial u}{\partial \sigma_1}$: We have

$$\begin{aligned}&-\sum_{j=1}^2 \varphi_{1j} W_j(T) + \frac{1}{\sigma_1 T}\sum_{j=1}^2\sum_{k=1}^2 \epsilon_{1k}\varphi_{1j}\Big(W_j(T)W_k(T) - \delta_{jk}T\Big) \\ =\ &-\left(W_1(T) - \frac{\rho}{\sqrt{1-\rho^2}}W_2(T)\right) + \frac{1}{\sigma_1 T}\left(W_1^2(T) - T - \frac{\rho}{\sqrt{1-\rho^2}}W_1(T)W_2(T)\right).\end{aligned}$$

2. $\frac{\partial u}{\partial \sigma_2}$: Similarly we have

$$\begin{aligned}
& -\sum_{j=1}^{2}\varphi_{2j}W_j(T)+\frac{1}{\sigma_2 T}\sum_{j=1}^{2}\sum_{k=1}^{2}\epsilon_{2k}\varphi_{2j}\Big(W_j(T)W_k(T)-\delta_{jk}T\Big) \\
= & -\frac{1}{\sqrt{1-\rho^2}}W_2(T)+\frac{1}{\sigma_2 T}\left(\frac{\rho}{\sqrt{1-\rho^2}}W_1(T)W_2(T)+W_2^2(T)-T\right).
\end{aligned}$$

Example 18.6.9 (Multi-Asset Option-Rho(continued from Example 18.6.7)). For the weight for undiscounted rho, first introduce Z_t^ρ:

$$Z_t^\rho = \begin{pmatrix} tX_1(t) \\ \vdots \\ tX_n(t) \end{pmatrix}.$$

Then we have

$$\begin{aligned}
\sigma^{-1}(X(t))Z_t^\rho &= \Phi^{-1}D_n\left[\frac{1}{\sigma_i X_i(t)}\right]\begin{pmatrix} tX_1(t) \\ \vdots \\ tX_n(t) \end{pmatrix} \\
&= \Phi^{-1}\begin{pmatrix} \frac{t}{\sigma_1} \\ \vdots \\ \frac{t}{\sigma_n} \end{pmatrix}
\end{aligned}$$

The rho is given by

$$\begin{aligned}
& E\left[\phi\Big(X_1(T),X_2(T),\cdots,X_n(T)\Big)\frac{1}{T}\delta\Big(\sigma^{-1}(X(T))Z_T^\rho\Big)\right] \\
= & E\left[\phi\Big(X_1(T),X_2(T),\cdots,X_n(T)\Big)\frac{1}{T}\int_0^T\Big(\sigma^{-1}(X(T))Z_T^\rho\Big)^* dW(t)\right] \\
= & E\left[\phi\Big(X_1(T),X_2(T),\cdots,X_n(T)\Big)\frac{1}{T}\int_0^T\Big(\sigma^{-1}(X(T))Z_T^\rho\Big)^T dW(t)\right] \\
= & E\left[\phi\Big(X_1(T),X_2(T),\cdots,X_n(T)\Big)\frac{1}{T}\int_0^T\left(\frac{T}{\sigma_1},\cdots,\frac{T}{\sigma_n}\right)(\Phi^{-1})^T dW(t)\right] \\
= & E\left[\phi\Big(X_1(T),X_2(T),\cdots,X_n(T)\Big)\frac{1}{T}\int_0^T\left(\frac{T}{\sigma_1},\cdots,\frac{T}{\sigma_n}\right)(\Phi^{-1})^T dW(t)\right] \\
= & E\left[\phi\Big(X_1(T),X_2(T),\cdots,X_n(T)\Big)\sum_{i=1}^{n}\left(\frac{1}{\sigma_i}\sum_{j=1}^{n}\varphi_{ij}W_j(T)\right)\right] \\
= & E\left[\phi\Big(X_1(T),X_2(T),\cdots,X_n(T)\Big)T\sum_{i=1}^{n}x_i w_i\right],
\end{aligned}$$

where w_i is defined in (18.6.3).

18.7 Discrete Path Dependent Options

Let us consider an underlying asset, Ito process $X_t \in \mathbb{R}$ following:

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = x \in \mathbb{R}.$$

We assume that the drift and diffusion coefficients b and σ of the diffusion X_t are continuously differentiable function with bounded Lipschitz in order to ensure the existence of a unique solution. Furthermore we suppose that σ is uniformly elliptic:

$$\exists \epsilon > 0, \quad \xi^T \sigma^T(x)\sigma(x)\xi \geq \epsilon|\xi|^2 \quad \text{for any } \xi, x \in \mathbb{R}^n.$$

This condition is necessary for the existence of the Malliavin weight, because we need to inverse it.

We suppose that the payoff function ϕ depend on a finite number of values $X_{t_1}, \cdots, X_{t_m}$. We denote the payoff

$$\phi(X_{t_1}, \cdots, X_{t_m}).$$

✎ **Conditions On The Weight For the Delta**
We have a necessary and sufficient condition for the generator of the delta.

Theorem 18.7.1. For a Skorohod integrable weight ω,

$$\Delta = e^{-rT}E_x\Big[\phi(X_{t_1}, \cdots, X_{t_m})\ \delta(\omega)\Big] \tag{18.7.1}$$

$$\iff \quad E_x\left[Y_{t_i} - \int_0^{t_i} Y_{t_i}Y_s^{-1}\sigma(s, X_s)\omega(s)ds \;\Big|\; X_{t_1}, \cdots, X_{t_m}\right] = 0, \quad \forall i = 1, \cdots, m. \tag{18.7.2}$$

Proof. Let Y_{t_i} be the first variation process of the $X_{t_i}, i = 1, \cdots, m$. Then we have

$$\Delta = e^{-rT}E\left[\sum_{i=1}^m \partial_i\phi(X_{t_1}, \cdots, X_{t_m})Y_{t_i}\right].$$

Note that

$$\begin{aligned}
& e^{-rT}E\left[\phi(X_{t_1}, \cdots, X_{t_m})\delta(\omega)\right] \\
= \; & e^{-rT}E\left[\int_0^T D_s\phi(X_{t_1}, \cdots, X_{t_m})\omega(s)ds\right] \quad \text{by (18.5.2)} \\
= \; & e^{-rT}E\left[\sum_{i=1}^m \partial_i\phi(X_{t_1}, \cdots, X_{t_m})\int_0^T D_sX_{t_i}\omega(s)ds\right] \\
= \; & e^{-rT}E\left[\sum_{i=1}^m \partial_i\phi(X_{t_1}, \cdots, X_{t_m})\int_0^T Y_{t_i}Y_s^{-1}\sigma(s, X_s)\mathbf{1}_{\{s\leq t_i\}}\omega(s)ds\right] \quad \text{by (18.4.5).}
\end{aligned}$$

Suppose ω satisfies (18.7.1). Then we have

$$e^{-rT}E\left[\sum_{i=1}^m \partial_i\phi(X_{t_1}, \cdots, X_{t_m})\left(Y_{t_i} - \int_0^T Y_{t_i}Y_s^{-1}\sigma(s, X_s)\mathbf{1}_{\{s\leq t_i\}}\omega(s)ds\right)\right] = 0.$$

Since the last equality true for any payoff function ϕ, Equation (18.7.2) must hold. Reciprocally, suppose that ω satisfies (18.7.2). Then exactly, the same argument completes the proof. ■

Example 18.7.2. Suppose that a function $a : [0,T] \to \mathbb{R}$ in $L^2([0,T])$ satisfies

$$\int_0^{t_i} a(s)ds = 1, \quad \forall i = 1, \cdots, m.$$

Define

$$\omega(s) \quad = \quad a(s)\sigma^{-1}(s, X_s)Y_s.$$

The ω satisfies (18.7.2), because, for all i,

$$\begin{aligned}
\int_0^{t_i} Y_{t_i}Y_s^{-1}\sigma(s,X_s)\omega(s)ds &= \int_0^{t_i} Y_{t_i}Y_s^{-1}\sigma(s,X_s)a(s)\sigma^{-1}(s,X_s)Y_s ds \\
&= Y_{t_i}\int_0^{t_i} a(s)ds \\
&= Y_{t_i}
\end{aligned}$$

This example proves Theorem(18.5.2). Note that

$$\int_0^T D_s X_{t_i}\omega(s)ds \overset{\text{by}(18.4.5)}{=} \int_0^T Y_{t_i}Y_s^{-1}\sigma(s,X_s)\mathbf{1}_{\{s\le t_i\}}\omega(s)ds = Y_{t_i}. \tag{18.7.3}$$

□

Example 18.7.3. Suppose that the underlying X_t follows

$$dX_t \quad = \quad \lambda X_t + \sigma X_t dW_t, \quad X_0 \quad = \quad x.$$

1. For a European option with maturity T, Eq(18.7.2) yields

$$\begin{aligned}
E\left[\frac{X_T}{x} - \int_0^T \frac{X_T}{x}\frac{x}{X_s}\sigma X_s\omega(s)ds\right] &= 0 \\
E\left[\frac{X_T}{x} - X_T\int_0^T \sigma\omega(s)ds\right] &= 0.
\end{aligned} \tag{18.7.4}$$

   If we put

$$w(s) \quad := \quad \frac{1}{x\sigma T},$$

   $w(s)$ satisfies (18.7.4). This yields the Malliavin weight

$$\delta(w) \quad = \quad \delta\left(\frac{1}{x\sigma T}\right) \quad = \quad \frac{W_T}{x\sigma T}.$$

2. Suppose a path-dependent option is depend on time $t_1 < t_2 < \cdots < t_m \le T$. Let

$$\begin{aligned}
a(s) &= \frac{1}{t_1}\mathbf{1}_{\{s\le t_1\}}, \\
\omega(s) &= a(s)\sigma^{-1}(s,X_s)Y_s \quad = \quad \frac{1}{x\sigma t_1}\mathbf{1}_{\{s\le t_1\}}.
\end{aligned}$$

Then we have

$$
\begin{aligned}
\Delta &= e^{-rT}E\left[\sum_{i=1}^{m}\partial_i\phi(X_{t_1},\cdots,X_{t_m})Y_{t_i}\right] \\
&= e^{-rT}E\left[\sum_{i=1}^{m}\partial_i\phi(X_{t_1},\cdots,X_{t_m})\int_0^T D_sX_{t_i}\omega(s)ds\right], \quad \text{by (18.7.3)} \\
&= e^{-rT}E\left[\int_0^T D_s\phi(X_{t_1},\cdots,X_{t_m})\omega(s)ds\right] \\
&= e^{-rT}E\left[\phi(X_{t_1},\cdots,X_{t_m})\delta(\omega)\right] \\
&= e^{-rT}E\left[\phi(X_{t_1},\cdots,X_{t_m})\frac{1}{x\sigma t_1}\delta\big(\mathbf{1}_{\{s\le t_1\}}\big)\right] \\
&= e^{-rT}E\left[\phi(X_{t_1},\cdots,X_{t_m})\frac{W_{t_1}}{x\sigma t_1}\right].
\end{aligned}
$$

3. Now let us calculate the weight for gamma.

$$
\begin{aligned}
\Gamma &= \frac{\partial\Delta}{\partial x} \\
&= e^{-rT}E\left[\frac{\partial}{\partial x}\phi(X_{t_1},\cdots,X_{t_m})\frac{W_{t_1}}{x\sigma t_1} - \phi(X_{t_1},\cdots,X_{t_m})\frac{W_{t_1}}{x^2\sigma t_1}\right].
\end{aligned}
$$

The first part can be computed as

$$
\begin{aligned}
& e^{-rT}E\left[\frac{\partial}{\partial x}\phi(X_{t_1},\cdots,X_{t_m})\frac{W_{t_1}}{x\sigma t_1}\right] \\
=\; & \frac{e^{-rT}}{x\sigma t_1}E\left[\frac{\partial}{\partial x}\phi(X_{t_1},\cdots,X_{t_m})W_{t_1}\right] \\
=\; & \frac{e^{-rT}}{x\sigma t_1}E\left[W_{t_1}\sum_{i=1}^{m}\partial_i\phi(X_{t_1},\cdots,X_{t_m})Y_{t_i}\right] \\
=\; & \frac{e^{-rT}}{x\sigma t_1}E\left[W_{t_1}\sum_{i=1}^{m}\partial_i\phi(X_{t_1},\cdots,X_{t_m})\int_0^T D_sX_{t_i}\omega(s)ds\right], \quad \text{by (18.7.3)} \\
=\; & \frac{e^{-rT}}{x\sigma t_1}E\left[W_{t_1}\int_0^T D_s\phi(X_{t_1},\cdots,X_{t_m})\omega(s)ds\right] \\
=\; & \frac{e^{-rT}}{x\sigma t_1}E\left[\phi(X_{t_1},\cdots,X_{t_m})\delta(W_{t_1}\omega)\right] \\
=\; & \frac{e^{-rT}}{x\sigma t_1}E\left[\phi(X_{t_1},\cdots,X_{t_m})\frac{1}{x\sigma t_1}\delta\big(W_{t_1}\big)\right] \\
=\; & \frac{e^{-rT}}{x\sigma t_1}E\left[\phi(X_{t_1},\cdots,X_{t_m})\frac{W_{t_1}^2-t_1}{x\sigma t_1}\right]
\end{aligned}
$$

Thus we have

$$
\Gamma = \frac{e^{-rT}}{x^2\sigma t_1}\,E\left[\phi(X_{t_1},\cdots,X_{t_m})\,\left(\frac{W_{t_1}^2}{\sigma t_1} - W_{t_1} - \frac{1}{\sigma}\right)\right].
$$

□

✎ **Perturbed Process**
Let denote by $\widetilde{b}$ a direction function for the drift term and $\widetilde{\sigma}$ for the volatility term. We assume that for every $\epsilon \in [-1, 1]$,

1. b and $b + \epsilon\widetilde{b}$ are C^1 with bounded derivatives, satisfy the Lipschitz conditions.
2. σ and $\sigma + \epsilon\widetilde{\sigma}$ are C^1 with bounded derivatives, satisfy the Lipschitz conditions and the uniform ellipticity condition.

Define perturbed processes $X_t^{\epsilon,\rho}$ and $X_t^{\epsilon,\sigma}$ of the process X_t by

$$dX_t^{\epsilon,\rho} = \left[b(t, X_t^{\epsilon,\rho}) + \epsilon\widetilde{b}(t, X_t^{\epsilon,\rho})\right]dt + \sigma(t, X_t^{\epsilon,\rho})dW_t, \quad X_0^{\epsilon,\rho} = x$$

and

$$dX_t^{\epsilon,\sigma} = b(t, X_t^{\epsilon,\sigma})dt + \left[\sigma(t, X_t^{\epsilon,\sigma}) + \epsilon\widetilde{\sigma}(t, X_t^{\epsilon,\sigma})\right]dW_t, \quad X_0^{\epsilon,\sigma} = x.$$

We denote variation processes Z_t^ρ and Z_t^σ the Gateau derivative of the drift-perturbed process $X_t^{\epsilon,\rho}$, respectively the volatility-perturbed process $X_t^{\epsilon,\sigma}$ along the direction $\widetilde{b}$, respectively $\widetilde{\sigma}$. The two quantities are defined as the limit in L^2, uniformly with respect to the time t:

$$Z_t^\rho = \lim_{L^2,\epsilon\to 0} \frac{X_t^{\epsilon,\rho} - X_t}{\epsilon} \tag{18.7.5}$$

respectively

$$Z_t^\sigma = \lim_{L^2,\epsilon\to 0} \frac{X_t^{\epsilon,\sigma} - X_t}{\epsilon}. \tag{18.7.6}$$

Then each perturbed process satisfies the following SDE, respectively.

$$\begin{aligned} dZ_t^\rho &= \left[\widetilde{b}(t, X_t) + \partial_2 b(t, X_t)Z_t^\rho\right]dt + \partial_2\sigma(t, X_t)Z_t^\rho dW_t, \quad Z_0^\rho = 0 \quad \text{and} \\ dZ_t^\sigma &= \partial_2 b(t, X_t)Z_t^\sigma dt + \left[\widetilde{\sigma}(t, X_t) + \partial_2\sigma(t, X_t)Z_t^\sigma\right]dW_t, \quad Z_0^\sigma = 0. \end{aligned}$$

Theorem 18.7.4. The processes Z_t^ρ and Z_t^σ can be expressed in terms of the first variation process by

$$Z_t^\rho = \int_0^t Y_t Y_s^{-1}\, \widetilde{b}(s, X_s)\, ds, \qquad \text{and} \tag{18.7.7}$$

$$Z_t^\sigma = \int_0^t Y_t Y_s^{-1}\, \widetilde{\sigma}(s, X_s)\, dW_s - \int_0^t Y_t \partial_2\sigma(s, X_s) Y_s^{-1}\, \widetilde{\sigma}(s, X_s)\, ds. \tag{18.7.8}$$

Example 18.7.5. Suppose we are given

$$\begin{aligned} X_t &= x \exp\left(\left(r - q - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right), \\ \widetilde{b}(t, X_t) &= X_t, \\ \widetilde{\sigma}(t, X_t) &= X_t. \end{aligned}$$

By Theorem(18.7.4), we have

$$
\begin{aligned}
Z_t^\rho &= \int_0^t Y_t Y_s^{-1}\ \widetilde{b}(s, X_s)\ ds \\
&= \int_0^t Y_t \frac{x}{X_s} X_s ds \\
&= xtY_t \\
&= tX_t \\
&= \frac{\partial X_t}{\partial r}, \\
Z_t^\sigma &= \int_0^t Y_t Y_s^{-1}\ \widetilde{\sigma}(s, X_s)\ dW_s - \int_0^t Y_t \partial_2 \sigma(s, X_s) Y_s^{-1}\ \widetilde{\sigma}(s, X_s)\ ds \\
&= W_t X_t - \int_0^t Y_t \sigma \frac{x}{X_s} X_s ds \\
&= W_t X_t - \sigma t X_t \\
&= X_t (W_t - \sigma t) \\
&= \frac{\partial X_t}{\partial \sigma}.
\end{aligned}
$$

✎ **Rho**

For the European options, the undiscounted rho is defined as the Gateau derivative of the perturbed price function in the direction given by $\widetilde{b}$:

$$
\begin{aligned}
\rho &= \frac{\partial}{\partial \epsilon} E[\phi(X_T^{\epsilon,\rho})]\Big|_{\epsilon=0,\widetilde{b}\text{ given}} \\
&= E\left[\phi'(X_T) Z_T^\rho\right] \\
&= E\left[\phi'(X_T) \int_0^T Y_T Y_s^{-1}\ \widetilde{b}(s, X_s)\ ds\right] \quad \text{by (18.7.7)} \\
&= E\left[\phi'(X_T) \int_0^T D_s X_T \sigma^{-1}(s, X_s)\ \widetilde{b}(s, X_s)\ ds\right] \quad \text{by (18.4.5)} \\
&= E\left[\int_0^T \phi'(X_T) D_s X_T \sigma^{-1}(s, X_s)\ \widetilde{b}(s, X_s)\ ds\right] \\
&= E\left[\int_0^T D_s \phi(X_T) \sigma^{-1}(s, X_s)\ \widetilde{b}(s, X_s)\ ds\right] \\
&= E\left[\phi(X_T) \delta\big(\sigma^{-1}(s, X_s)\ \widetilde{b}(s, X_s)\big)\right] \quad \text{by (18.5.2)}.
\end{aligned}
$$

More generally, for $t_1 \leq \cdots \leq t_m = T$,

$$
\begin{aligned}
\rho &= \frac{\partial}{\partial \epsilon} E\left[\phi(X_{t_1}^{\epsilon,\rho}, \cdots, X_{t_m}^{\epsilon,\rho})\right]\Big|_{\epsilon=0,\widetilde{b}\text{ given}} \\
&= E\left[\sum_{i=1}^m \partial_i \phi(X_{t_1}, \cdots, X_{t_m}) Z_{t_i}^\rho\right]
\end{aligned}
$$

$$
\begin{aligned}
&= E\left[\sum_{i=1}^{m} \partial_i \phi(X_{t_1}, \cdots, X_{t_m}) \int_0^{t_i} Y_{t_i} Y_s^{-1}\, \widetilde{b}(s, X_s)\, ds\right] \quad \text{by (18.7.7)} \\
&= E\left[\sum_{i=1}^{m} \partial_i \phi(X_{t_1}, \cdots, X_{t_m}) \int_0^{T} Y_{t_i} Y_s^{-1}\, \widetilde{b}(s, X_s) \mathbf{1}_{\{s \le t_i\}}\, ds\right] \quad \text{by (18.4.5)} \\
&= E\left[\int_0^{T} \sum_{i=1}^{m} \partial_i \phi(X_{t_1}, \cdots, X_{t_m}) Y_{t_i} Y_s^{-1}\, \widetilde{b}(s, X_s) \mathbf{1}_{\{s \le t_i\}}\, ds\right] \\
&= E\left[\int_0^{T} \sum_{i=1}^{m} \partial_i \phi(X_{t_1}, \cdots, X_{t_m}) D_s X_{t_i} \sigma^{-1}(s, X_s)\, \widetilde{b}(s, X_s)\, ds\right] \\
&= E\left[\int_0^{T} D_s \phi(X_{t_1}, \cdots, X_{t_m})\, \sigma^{-1}(s, X_s)\, \widetilde{b}(s, X_s)\, ds\right] \\
&= E\left[\phi(X_{t_1}, \cdots, X_{t_m})\, \delta\Big(\sigma^{-1}(s, X_s)\, \widetilde{b}(s, X_s)\Big)\right] \quad \text{by (18.5.2).}
\end{aligned}
$$

$\delta\Big(\sigma^{-1}(s, X_s)\ \widetilde{b}(s, X_s)\Big)$ is a special case of the following Theorem 18.7.6. In particular, when $\widetilde{b}(s, X_s) = X_s$, we obtain the classical rho.

Theorem 18.7.6. For a Skorohod integrable weight ω,

$$
\rho = E_x\Big[\phi(X_{t_1}, \cdots, X_{t_m})\ \delta(\omega)\Big] \tag{18.7.9}
$$

$$
\Longleftrightarrow \quad E_x\left[Y_{t_i} \int_0^{t_i} \frac{\sigma(s, X_s)\omega(s) - \widetilde{b}(s, X_s)}{Y_s} ds \,\Big|\, X_{t_1}, \cdots, X_{t_m}\right] = 0, \quad \forall i = 1, \cdots, m. \tag{18.7.10}
$$

(up to the discounted factor)

Proof. Let $Z_{t_i}^{\rho}$ be the variation process of the $X_{t_i}, i = 1, \cdots, m$, defined in (18.7.5). Then we have

$$
\begin{aligned}
\rho &= E\left[\sum_{i=1}^{m} \partial_i \phi(X_{t_1}, \cdots, X_{t_m}) Z_{t_i}^{\rho}\right] \\
&= E\left[\sum_{i=1}^{m} \partial_i \phi(X_{t_1}, \cdots, X_{t_m}) \int_0^{t_i} Y_{t_i} Y_s^{-1} \widetilde{b}(s, X_s)\, ds\right] \quad \text{by (18.7.7)} \\
&= E\left[\sum_{i=1}^{m} \partial_i \phi(X_{t_1}, \cdots, X_{t_m}) \int_0^{T} Y_{t_i} Y_s^{-1} \widetilde{b}(s, X_s) \mathbf{1}_{\{s \le t_i\}}\, ds\right].
\end{aligned}
$$

Note that

$$
\begin{aligned}
& E\left[\phi(X_{t_1}, \cdots, X_{t_m}) \delta(\omega)\right] \\
&= E\left[\int_0^{T} D_s \phi(X_{t_1}, \cdots, X_{t_m}) \omega(s) ds\right] \quad \text{by (18.5.2)} \\
&= E\left[\sum_{i=1}^{m} \partial_i \phi(X_{t_1}, \cdots, X_{t_m}) \int_0^{T} D_s X_{t_i} \omega(s) ds\right] \\
&= E\left[\sum_{i=1}^{m} \partial_i \phi(X_{t_1}, \cdots, X_{t_m}) \int_0^{T} Y_{t_i} Y_s^{-1} \sigma(s, X_s) \mathbf{1}_{\{s \le t_i\}} \omega(s) ds\right] \quad \text{by (18.4.5).}
\end{aligned}
$$

Suppose ω satisfies (18.7.9). Then we have

$$e^{-rT}E\left[\sum_{i=1}^{m}\partial_i\phi(X_{t_1},\cdots,X_{t_m})\left(Y_{t_i}\int_0^T Y_s^{-1}\Big(\sigma(s,X_s)-\widetilde{b}(s,X_s)\Big)\mathbf{1}_{\{s\le t_i\}}\omega(s)ds\right)\right]=0.$$

Since the last equality true for any payoff function ϕ, (18.7.10) holds. Reciprocally, suppose that ω satisfies (18.7.10). Then exactly, the same argument completes the proof. ■

Example 18.7.7. For an early redeemable option, let a stopping time τ be the redemption time. Then up to discounted factor the Malliavin weight for rho is given by

$$\frac{W_\tau}{\sigma},$$

since $W_T = W_T - W_\tau + W_\tau$, $W_T - W_\tau$ is independent of X_{t_i} for $t_i \le \tau$ and

$$\begin{aligned}
& E\left[\phi(X_{t_1},\cdots,X_{t_m})\cdot\frac{W_T}{\sigma}\right]\\
=\ & E\left[\phi(X_{t_1},\cdots,X_{t_m})\cdot\left(\frac{W_T-W_\tau}{\sigma}+\frac{W_\tau}{\sigma}\right)\right]\\
=\ & E\left[\phi(X_{t_1},\cdots,X_{t_m})\cdot\left(\frac{W_T-W_\tau}{\sigma}\right)\right]+E\left[\phi(X_{t_1},\cdots,X_{t_m})\cdot\frac{W_\tau}{\sigma}\right]\\
=\ & E\left[\phi(X_{t_1},\cdots,X_\tau)\right]\cdot \cancelto{0}{E\left[\frac{W_T-W_\tau}{\sigma}\right]}+E\left[\phi(X_{t_1},\cdots,X_\tau)\cdot\frac{W_\tau}{\sigma}\right]\\
=\ & E\left[\phi(X_{t_1},\cdots,X_\tau)\cdot\frac{W_\tau}{\sigma}\right].
\end{aligned}$$

Hence the rho can be written as

$$\rho \;=\; E\left[e^{-r\tau}\phi(X_{t_1},\cdots,X_\tau)\cdot\left(\frac{W_\tau}{\sigma}-\tau\right)\right].$$

✎ **Vega**

Theorem 18.7.8. For a Skorohod integrable weight ω,

$$\mathcal{V} \;=\; e^{-rT}E_x\Big[\phi(X_{t_1},\cdots,X_{t_m})\ \delta(\omega)\Big] \tag{18.7.11}$$

$$\Longleftrightarrow\quad E_x\left[Y_{t_i}\int_0^{t_i}\frac{\sigma(s,X_s)\omega(s)+\partial_2\sigma(s,X_s)\widetilde{\sigma}(s,X_s)}{Y_s}ds-\int_0^{t_i}\frac{\widetilde{\sigma}(s,X_s)Y_{t_i}}{Y_s}dW_s\ \Big|\ X_{t_1},\cdots,X_{t_m}\right]=0,$$

$$\forall i=1,\cdots,m. \tag{18.7.12}$$

$$\Longleftrightarrow\quad Z_{t_i}^{\sigma}=\int_0^T D_sX_{t_i}\omega(s)ds,\quad \forall i=1,\cdots,m. \tag{18.7.13}$$

Proof. Let Z_{ti}^{σ} be the variation process of the $X_{t_i}, i=1,\cdots,m$, defined in (18.7.6). Then we have

$$\mathcal{V} \;=\; e^{-rT}E\left[\sum_{i=1}^{m}\partial_i\phi(X_{t_1},\cdots,X_{t_m})Z_{t_i}^{\sigma}\right]$$

$$
\begin{aligned}
&= e^{-rT}E\left[\sum_{i=1}^m \partial_i\phi(X_{t_1},\cdots,X_{t_m})\left(\int_0^{t_i} Y_{t_i}Y_s^{-1}\,\widetilde{\sigma}(s,X_s)\,dW_s - \int_0^{t_i} Y_{t_i}\partial_2\sigma(s,X_s)Y_s^{-1}\,\widetilde{\sigma}(s,X_s)\,ds\right)\right] \quad \text{by (18.7.8)}\\
&= e^{-rT}E\left[\sum_{i=1}^m \partial_i\phi(X_{t_1},\cdots,X_{t_m})\left(\int_0^{T} Y_{t_i}Y_s^{-1}\,\widetilde{\sigma}(s,X_s)\mathbf{1}_{\{s\le t_i\}}\,dW_s - \int_0^{T} Y_{t_i}\partial_2\sigma(s,X_s)Y_s^{-1}\,\widetilde{\sigma}(s,X_s)\mathbf{1}_{\{s\le t_i\}}\,ds\right)\right].
\end{aligned}
$$

Note that

$$
\begin{aligned}
& E\left[\phi(X_{t_1},\cdots,X_{t_m})\delta(\omega)\right]\\
= \;& E\left[\int_0^T D_s\phi(X_{t_1},\cdots,X_{t_m})\omega(s)ds\right] \quad \text{by (18.5.2)}\\
= \;& E\left[\sum_{i=1}^m \partial_i\phi(X_{t_1},\cdots,X_{t_m})\int_0^T D_sX_{t_i}\omega(s)ds\right]\\
= \;& E\left[\sum_{i=1}^m \partial_i\phi(X_{t_1},\cdots,X_{t_m})\int_0^T Y_{t_i}Y_s^{-1}\sigma(s,X_s)\mathbf{1}_{\{s\le t_i\}}\omega(s)ds\right] \quad \text{by (18.4.5)}.
\end{aligned}
$$

By the same argument in the proof of Theorem(18.7.6) we can complete the proof. ■

Example 18.7.9. Founie et al.(1999) proved the weighting function could be written in the case of adapted processes as some Ito integral. Let us define

$$
\widetilde{\Gamma}_m = \left\{\widetilde{a}\in L^2[0,T]\;\Big|\;\int_{t_{i-1}}^{t_i}\widetilde{a}(t)dt = 1, \quad \text{for } i=1,\cdots m\right\}
$$

For some $a(t)\in\widetilde{\Gamma}_m$, let

$$
w(t) = \frac{Y_t}{\sigma(t,X_t)}\widetilde{a}(t)\sum_{i=1}^m\left(\frac{Z^\sigma_{t_i}}{Y_{t_i}} - \frac{Z^\sigma_{t_{i-1}}}{Y_{t_{i-1}}}\right)\mathbf{1}_{\{t_{i-1}\le t\le t_i\}}.
$$

Suppose that $X_{t_1},\cdots,X_{t_m}$ are given, then we have

$$
\begin{aligned}
\int_0^T D_sX_{t_i}\omega(s)ds &= \int_0^T Y_{t_i}Y_s^{-1}\sigma(s,X_s)\mathbf{1}_{\{s\le t_i\}}\omega(s)ds\\
&= Y_{t_i}\sum_{k=1}^i\left(\int_{t_{k-1}}^{t_k}\widetilde{a}(s)\left(\frac{Z^\sigma_{t_k}}{Y_{t_k}} - \frac{Z^\sigma_{t_{k-1}}}{Y_{t_{k-1}}}\right)ds\right)\\
&= Y_{t_i}\sum_{k=1}^i\left(\left(\frac{Z^\sigma_{t_k}}{Y_{t_k}} - \frac{Z^\sigma_{t_{k-1}}}{Y_{t_{k-1}}}\right)\int_{t_{k-1}}^{t_k}\widetilde{a}(s)ds\right)\\
&= Y_{t_i}\sum_{k=1}^i\left(\frac{Z^\sigma_{t_k}}{Y_{t_k}} - \frac{Z^\sigma_{t_{k-1}}}{Y_{t_{k-1}}}\right)\\
&= Z^\sigma_{t_i}.
\end{aligned}
$$

This means that $\omega(t)$ satisfies condition(18.7.13). If we choose

$$
\widetilde{a}(t) = \sum_{i=1}^m\frac{1}{t_i - t_{i-1}}\,\mathbf{1}_{\{t_{i-1}\le t<t_i\}},
$$

$\widetilde{\sigma}(t, X_t) = \widetilde{\theta} X_t$ and $\sigma(t, X_t) = \sigma X_t$, where $\widetilde{\theta}$ and σ are constant, then we have, by Theorem(18.7.4),

$$\begin{aligned}
Z_t^{\sigma} &= \int_0^t Y_t Y_s^{-1}\, \widetilde{\sigma}(s, X_s)\, dW_s - \int_0^t Y_t \partial_2 \sigma(s, X_s) Y_s^{-1}\, \widetilde{\sigma}(s, X_s) \\
&= x\widetilde{\theta} Y_t W_t - x\widetilde{\theta}\sigma t Y_t \\
&= x\widetilde{\theta}(W_t - \sigma t) Y_t.
\end{aligned}$$

Also we have

$$\begin{aligned}
w(t) &= \frac{Y_t}{\sigma X_t}\widetilde{a}(t) \sum_{i=1}^m \left(x\widetilde{\theta}(W_{t_i} - \sigma t_i) - x\widetilde{\theta}(W_{t_{i-1}} - \sigma t_{i-1}) \right) \mathbf{1}_{\{t_{i-1} \le t \le t_i\}} \\
&= \frac{\widetilde{\theta}}{\sigma} \sum_{i=1}^m \widetilde{a}(t) \left(W_{t_i} - W_{t_{i-1}} - \sigma(t_i - t_{i-1}) \right) \mathbf{1}_{\{t_{i-1} \le t < t_i\}} \\
&= \frac{\widetilde{\theta}}{\sigma} \sum_{i=1}^m \frac{W_{t_i} - W_{t_{i-1}} - \sigma(t_i - t_{i-1})}{t_i - t_{i-1}} \mathbf{1}_{\{t_{i-1} \le t < t_i\}}.
\end{aligned}$$

Hence by Theorem(18.7.8), the weight for vega is given by

$$\begin{aligned}
\delta(w) &= \frac{\widetilde{\theta}}{\sigma} \sum_{i=1}^m \frac{1}{t_i - t_{i-1}} \delta\left(W_{t_i} - W_{t_{i-1}} - \sigma(t_i - t_{i-1}) \mathbf{1}_{\{t_{i-1} \le t < t_i\}} \right) \\
&= \frac{\widetilde{\theta}}{\sigma} \sum_{i=1}^m \frac{(W_{t_i} - W_{t_{i-1}})^2 - (t_i - t_{i-1}) - \sigma(t_i - t_{i-1})(W_{t_i} - W_{t_{i-1}})}{t_i - t_{i-1}} \\
&= \widetilde{\theta} \sum_{i=1}^m \left\{ \frac{(\Delta W_{t_i})^2}{\sigma \Delta t_i} - \Delta W_{t_i} - \frac{1}{\sigma} \right\},
\end{aligned}$$

where $\Delta W_{t_i} = W_{t_i} - W_{t_{i-1}}$ and $\Delta t_i = t_i - t_{i-1}$. When $\widetilde{\theta} = 1$, we can obtain the classical vega. Note that

$$E\left[\frac{(\Delta W_{t_i})^2}{\sigma \Delta t_i} - \Delta W_{t_i} - \frac{1}{\sigma} \right] = \frac{\Delta t_i}{\sigma \Delta t_i} - 0 - \frac{1}{\sigma} = 0.$$

For an early redeemable option, let a stopping time τ be the redemption time. Then the vega is given by

$$\mathcal{V} = E\left[e^{-r\tau} \phi(X_{t_1}, \cdots, X_{\tau}) \sum_{t_i \le \tau} \left\{ \frac{(\Delta W_{t_i})^2}{\sigma \Delta t_i} - \Delta W_{t_i} - \frac{1}{\sigma} \right\} \right].$$

□

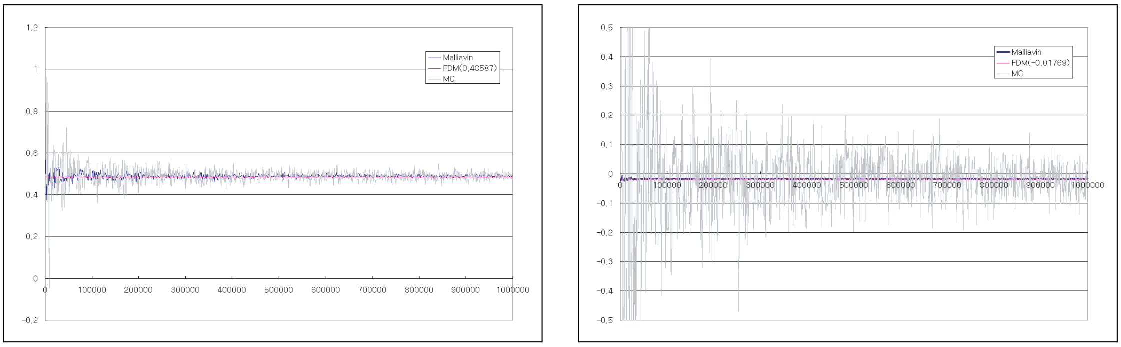
Example 18.7.10. For option with double assets, the vegas are given by

$$\begin{aligned}
\mathcal{V}_1 &= E\left[e^{-r\tau} \phi(\boldsymbol{X}_{t_1}, \cdots, \boldsymbol{X}_{\tau}) \sum_{t_i \le \tau} \left\{ \frac{\Delta W_{t_i}^1}{\sigma_1 \Delta t_i} \left(\Delta W_{t_i}^1 - \frac{\rho}{\sqrt{1-\rho^2}} \Delta W_{t_i}^2 \right) - \left(\Delta W_{t_i}^1 - \frac{\rho}{\sqrt{1-\rho^2}} \Delta W_{t_i}^2 \right) - \frac{1}{\sigma_1} \right\} \right], \\
\mathcal{V}_2 &= E\left[e^{-r\tau} \phi(\boldsymbol{X}_{t_1}, \cdots, \boldsymbol{X}_{\tau}) \sum_{t_i \le \tau} \left\{ \frac{\Delta W_{t_i}^2}{\sigma_2 \Delta t_i} \left(\frac{\rho}{\sqrt{1-\rho^2}} \Delta W_{t_i}^1 + \Delta W_{t_i}^2 \right) - \frac{\Delta W_{t_i}^2}{\sqrt{1-\rho^2}} - \frac{1}{\sigma_2} \right\} \right].
\end{aligned}$$

Also, the rho is given by

$$\text{Rho} \quad = \quad E\left[e^{-r\tau}\phi(\boldsymbol{X}_{t_1},\cdots,\boldsymbol{X}_{\tau})\left(\frac{W_\tau^1}{\sigma_1}-\frac{\rho W_\tau^2}{\sqrt{1-\rho^2}\sigma_1}+\frac{W_\tau^2}{\sqrt{1-\rho^2}\sigma_2}-\tau\right)\right]$$

In Figure(18.3), we present simulation results of gamma for 'Multi-Chance Early Redemption Note'.



(a) Delta (b) Gamma

Figure 18.3: Numerical results of Multi-Chance Early Redemption Note: redeemable period= 0.5, maturity= 3,$K = 100, \sigma = 30\%, r = 5\%$, coupon rate= $4.5\% \times 2$.

18.8 Applications of Malliavin Calculus to Barrier Options

✎ **When Malliavin Calculus Fails(source: P.K. Friz(2003))**
Consider a simple knock-out barrier option with payoff

$$\Phi(\omega) = \begin{cases} 0, & \text{if } \tau(\omega) < T \\ \phi(S_T(\omega)), & \text{otherwise} \end{cases}$$

where τ is the first hitting time of the knock-out region. For Malliavin calculus to work, we need $\Phi(\omega)$ smooth in the *sense of Malliavin* such that $D_t\Phi$ exists. We won't go into the maths, but the intuition of Malliavin's smoothness is the small perturbations in the path ω must not change $\Phi(\omega)$ too much. But there are paths extremely close to each other, one of which hits the barrier while the other doesn't. We see that $\Phi(\omega)$ is not smooth and Malliavin Calculus doesn't work. On the other hand, there is a smooth joint low for (S_T, τ) i.e. a smooth probability distribution, say $f(\cdot,\cdot)$, on $\mathbb{R}^2$ for the 2-dimensional random variable (S_T, τ). In general, we don't know f but by the logarithmic derivation trick, the existence of some optimal weight π_0 for computing Greeks is ensured.

Example 18.8.1. Let $X_t = x + W_t$ be a linear Brownian motion starting at $x \in \mathbb{R}$. Suppose we are interested in rewriting

$$\partial_x E\left[f\left(\inf_{0\le t\le 1} X_t\right)\right] = E\left[f\left(\inf_{0\le t\le 1} X_t\right) H\right]$$

for some square integrable random variable H and for some bounded function f. Using the explicit law[2] of $\inf_{0\le t\le 1} X_t$, we have

$$\begin{aligned}
\partial_x E\left[f\left(\inf_{0\le t\le 1} X_t\right)\right] &= \partial_x \int_{-\infty}^{x} f(y)\frac{2}{\sqrt{2\pi}}\exp\left(-\frac{(y-x)^2}{2}\right)\,dy \\
&= f(x)\frac{2}{\sqrt{2\pi}} + \int_{-\infty}^{x} f(y)\frac{2}{\sqrt{2\pi}}(y-x)\exp\left(-\frac{(y-x)^2}{2}\right)\,dy \\
&= f(x)\frac{2}{\sqrt{2\pi}} + E\left[f\left(\inf_{0\le t\le 1} X_t\right)\left(\inf_{0\le t\le 1} X_t - x\right)\right].
\end{aligned}$$

Hence this basic computation shows that finding a weight H if $f(x) \neq 0$ turns out to be hard. □

E. Gobet and A.K. Higa(2003) introduced the concept of dominating processes to computer barrier option's price via Malliavin calculus. But it is quite difficult to control the dominating processes.

In the followings, we propose a new method easy to computer the value of barrier options using Malliavin calculus. This new approach is based on Babsiri and Noel(1998).

The European barrier option's payoff are dependent on whether the barrier is touched during option's life time or not. But it doesn't matter when the barrier is touched. This feature is very important in the proposed method. One touch option, for example, cannot be evaluated via the proposed method. For the sake of simplicity, we are going to consider up & in barrier options.

✎ **Up & In Options**
In general, up & in type barrier option's payoff is given by

$$\Phi = \begin{cases} \phi(S_T), & M_T^S \ge H, \\ 0, & M_T^S < H, \end{cases}$$

[2] See Example(6.1.6).

where H is the barrier and $M_T^S = \max_{0\le t\le T} S(t)$. Note that we can computer the conditional probability

$$p_H := P(M_T^S \ge H \mid S_T) = \begin{cases} 1, & S_T \ge H \\ \exp\left(\frac{2\log\frac{S_T}{H}\ \log\frac{H}{S_0}}{\sigma^2 T}\right), & S_T < H \end{cases}$$

from Example(8.1.8). Let $x = S_0$. The value of the up & in barrier option is given by

$$\begin{aligned} u(x) &= e^{-rT} E\left[\phi(S_T)\cdot \mathbf{1}_{\{M_T^S \ge H\}}\right] \\ &= e^{-rT} E\left[E\left[\phi(S_T)\cdot \mathbf{1}_{\{M_T^S \ge H\}} | S_T\right]\right] \\ &= e^{-rT} E\left[\phi(S_T) E\left[\mathbf{1}_{\{M_T^S \ge H\}} | S_T\right]\right] \\ &= e^{-rT} E\left[\phi(S_T)\cdot P(M_T^S \ge H \mid S_T)\right] \\ &= e^{-rT} E\left[\phi(S_T)\cdot p_H\right]. \end{aligned}$$

We are more specifically interested in computing some derivatives, so called Greeks.

1. Delta: For the delta we have, by (18.5.8),

$$\begin{aligned} \Delta &= \frac{\partial u}{\partial x} \\ &= e^{-rT} E\left[\phi'(S_T)\frac{S_T}{x}\ p_H + \phi(S_T)\frac{\partial p_H}{\partial x}\right] \\ &= e^{-rT} E\left[\phi(S_T)\ \frac{1}{x}\delta\left(\frac{S_T p_H}{\int_0^T D_s S_T ds}\right) + \phi(S_T)\frac{\partial p_H}{\partial x}\right] \\ &= e^{-rT} E\left[\phi(S_T)\ \delta\left(\frac{p_H}{x\sigma T}\right) + \phi(S_T)\frac{\partial p_H}{\partial x}\right]. \end{aligned}$$

Since

$$\frac{\partial p_H}{\partial x} = \begin{cases} 0, & S_T \ge H \\ p_H \frac{2}{x\sigma^2 T}\left(2\log\frac{H}{x} - \mu T - \sigma W_T\right), & S_T < H \end{cases}$$

and by Example(18.4.5) and (18.5.3), if $S_T < H$

$$\begin{aligned} \delta(p_H) &= p_H\delta(1) - \int_0^T D_t p_H\ dt \\ &= p_H W_T - \int_0^T \frac{\partial p_H}{\partial W_T}\mathbf{1}_{\{t\le T\}} dt \\ &= p_H W_T - \int_0^T p_H \frac{2\log\frac{H}{x}}{\sigma T}\mathbf{1}_{\{t\le T\}} dt \\ &= p_H\left(W_T - \frac{2\log\frac{H}{x}}{\sigma}\right). \end{aligned}$$

Hence we have

$$\delta(p_H) = \begin{cases} W_T (= p_H W_T), & S_T \ge H \\ p_H\left(W_T - \frac{2\log\frac{H}{x}}{\sigma}\right), & S_T < H \end{cases}$$

Thus we have

$$\Delta = e^{-rT}E\left[\phi(S_T)p_H\left\{\frac{W_T}{x\sigma T}+\frac{2}{x\sigma^2 T}\left(\log\frac{H}{x}-\mu T-\sigma W_T\right)\cdot\mathbf{1}_{\{S_T<H\}}\right\}\right].$$

2. Gamma: Let π_Δ be the Malliavin weight for delta, i.e.

$$\pi_\Delta = p_H\left\{\frac{W_T}{x\sigma T}+\frac{2}{x\sigma^2 T}\left(\log\frac{H}{x}-\mu T-\sigma W_T\right)\cdot\mathbf{1}_{\{S_T<H\}}\right\}.$$

We have

$$\begin{aligned}\Gamma &= e^{-rT}\frac{\partial\Delta}{\partial x}\\ &= e^{-rT}\frac{\partial}{\partial x}E\left[\phi(S_T)\ \pi_\Delta\right]\\ &= e^{-rT}E\left[\phi'(S_T)\ \frac{S_T}{x}\pi_\Delta+\phi(S_T)\frac{\partial\pi_\Delta}{\partial x}\right]\\ &= e^{-rT}E\left[\phi(S_T)\ \delta\left(\frac{\pi_\Delta}{x\sigma T}\right)+\phi(S_T)\frac{\partial\pi_\Delta}{\partial x}\right]\\ &= e^{-rT}E\left[\phi(S_T)\cdot\pi_\Gamma\right],\end{aligned}$$

where π_Γ is the Malliavin weight for gamma, i.e.

$$\pi_\Gamma = \delta\left(\frac{\pi_\Delta}{x\sigma T}\right)+\frac{\partial\pi_\Delta}{\partial x}.$$

(a) $S_T\geq H$: In this case we have

$$\pi_\Gamma = \frac{1}{x^2\sigma T}\left(\frac{W_T^2}{\sigma T}-W_T-\frac{1}{\sigma}\right).$$

(b) $S_T<H$: First computer $\delta(\pi_\Delta)$:

$$\begin{aligned}\delta(\pi_\Delta) &= \pi_\Delta\delta(1)-\int_0^T D_t\pi_\Delta\,dt\\ &= \pi_\Delta W_T-\int_0^T\frac{\partial\pi_\Delta}{\partial W_T}\mathbf{1}_{\{t\leq T\}}\,dt\\ &= \pi_\Delta W_T-\int_0^T\left(\pi_\Delta\frac{2\log\frac{H}{x}}{\sigma T}-p_H\frac{1}{x\sigma T}\right)\mathbf{1}_{\{t\leq T\}}\,dt\\ &= \pi_\Delta W_T-\pi_\Delta\frac{2\log\frac{H}{x}}{\sigma}+p_H\frac{1}{x\sigma}\\ &= \pi_\Delta\left(W_T-\frac{2\log\frac{H}{x}}{\sigma}\right)+p_H\frac{1}{x\sigma}.\end{aligned}$$

Now let us calculate $\frac{\partial\pi_\Delta}{\partial x}$:

$$\frac{\partial\pi_\Delta}{\partial x} = \frac{\partial}{\partial x}\left[p_H\left\{\frac{W_T}{x\sigma T}+\frac{2}{x\sigma^2 T}\left(\log\frac{H}{x}-\mu T-\sigma W_T\right)\right\}\right]$$

$$
\begin{aligned}
&= \frac{\partial p_H}{\partial x}\cdot\left\{\frac{W_T}{x\sigma T}+\frac{2}{x\sigma^2 T}\left(\log\frac{H}{x}-\mu T-\sigma W_T\right)\right\}+p_H\cdot\frac{\partial}{\partial x}\left\{\frac{W_T}{x\sigma T}+\frac{2}{x\sigma^2 T}\left(\log\frac{H}{x}-\mu T-\sigma W_T\right)\right\}\\
&= \pi_\Delta\cdot\frac{2}{x\sigma^2 T}\left(2\log\frac{H}{x}-\mu T-\sigma W_T\right)+p_H\cdot\left\{-\frac{W_T}{x^2\sigma T}-\frac{2}{x^2\sigma^2 T}\left(\log\frac{H}{x}-\mu T-\sigma W_T+1\right)\right\}\\
&= \pi_\Delta\cdot\frac{2}{x\sigma^2 T}\left(2\log\frac{H}{x}-\mu T-\sigma W_T\right)+p_H\cdot\left\{\frac{W_T}{x^2\sigma T}-\frac{2}{x^2\sigma^2 T}\left(\log\frac{H}{x}-\mu T+1\right)\right\}.
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
\pi_\Gamma &= \pi_\Delta\frac{1}{x\sigma T}\left(W_T-\frac{2\log\frac{H}{x}}{\sigma}\right)+p_H\frac{1}{x^2\sigma^2 T}\\
&\quad+\pi_\Delta\cdot\frac{2}{x\sigma^2 T}\left(2\log\frac{H}{x}-\mu T-\sigma W_T\right)+p_H\cdot\left\{\frac{W_T}{x^2\sigma T}-\frac{2}{x^2\sigma^2 T}\left(\log\frac{H}{x}-\mu T+1\right)\right\}\\
&= \pi_\Delta\cdot\frac{1}{x\sigma^2 T}\left(2\log\frac{H}{x}-2\mu T-\sigma W_T\right)+p_H\cdot\left\{\frac{W_T}{x^2\sigma T}-\frac{1}{x^2\sigma^2 T}\left(2\log\frac{H}{x}-2\mu T+1\right)\right\}.
\end{aligned}
$$

(c) Hence we can conclude that

$$
\pi_\Gamma = \begin{cases} \frac{1}{x^2\sigma T}\left(\frac{W_T^2}{\sigma T}-W_T-\frac{1}{\sigma}\right), & S_T\geq H\\ \pi_\Delta\cdot\frac{1}{x\sigma^2 T}\left(2\log\frac{H}{x}-2\mu T-\sigma W_T\right)+p_H\cdot\left\{\frac{W_T}{x^2\sigma T}-\frac{1}{x^2\sigma^2 T}\left(2\log\frac{H}{x}-2\mu T+1\right)\right\}, & S_T<H. \end{cases}
$$

Note that we can computer the Malliavin weights more compactly. Consider that the case of $S_T<H$. Since

$$
\pi_\Delta = \delta\left(\frac{p_H}{x\sigma T}\right)+\frac{\partial p_H}{\partial x},
$$

we have that

$$
\begin{aligned}
\pi_\Gamma &= \delta\left(\frac{\pi_\Delta}{x\sigma T}\right)+\frac{\partial\pi_\Delta}{\partial x}\\
&= \frac{1}{x\sigma T}\delta\left(\delta\left(\frac{p_H}{x\sigma T}\right)+\frac{\partial p_H}{\partial x}\right)+\frac{\partial}{\partial x}\left(\delta\left(\frac{p_H}{x\sigma T}\right)+\frac{\partial p_H}{\partial x}\right)\\
&= \frac{1}{x^2\sigma^2 T^2}\delta(\delta(p_H))+\frac{1}{x\sigma T}\delta\left(\frac{\partial p_H}{\partial x}\right)+\frac{1}{x\sigma T}\frac{\partial\delta(p_H)}{\partial x}-\frac{1}{x^2\sigma T}\delta(p_H)+\frac{\partial^2 p_H}{\partial x^2}\\
&= \frac{1}{x^2\sigma^2 T^2}\delta(\delta(p_H))+\frac{2}{x\sigma T}\frac{\partial\delta(p_H)}{\partial x}-\frac{1}{x^2\sigma T}\delta(p_H)+\frac{\partial^2 p_H}{\partial x^2}.
\end{aligned}
$$

3. Vega: We can obtain Vega as follows.

$$
\begin{aligned}
\mathcal{V} &= e^{-rT}\frac{\partial}{\partial\sigma}E\left[\phi(S_T)p_H\right]\\
&= e^{-rT}E\left[\phi'(S_T)S_T(W_T-\sigma T)p_H+\phi(S_T)\frac{\partial p_H}{\partial\sigma}\right]\\
&= e^{-rT}E\left[\phi(S_T)\delta\left(\frac{(W_T-\sigma T)p_H}{\sigma T}\right)+\phi(S_T)\frac{\partial p_H}{\partial\sigma}\right]\\
&= e^{-rT}E\left[\phi(S_T)\cdot\pi_{\mathcal{V}}\right],
\end{aligned}
$$

where $\pi_{\mathcal{V}}$ is the Malliavin weight for Vega, i.e.

$$
\pi_{\mathcal{V}} = \delta\left(\frac{(W_T-\sigma T)p_H}{\sigma T}\right)+\frac{\partial p_H}{\partial\sigma}
$$

$$= \frac{1}{\sigma T}\delta(W_T p_H) - \delta(p_H) + \frac{\partial p_H}{\partial \sigma}.$$

Note that if $S_T < H$, we have

$$\begin{aligned}
\frac{\partial p_H}{\partial \sigma} &= p_H \frac{2\log\frac{H}{x}}{T}\frac{\partial}{\partial \sigma}\left(\frac{1}{\sigma^2}\log\frac{S_T}{H}\right) \\
&= p_H \frac{2\log\frac{H}{x}}{T}\left(-\frac{2}{\sigma^3}\log\frac{S_T}{H} + \frac{W_T - \sigma T}{\sigma^2}\right) \\
&= p_H \frac{2\log\frac{H}{x}}{\sigma^2 T}\left(-\frac{2}{\sigma}\log\frac{S_T}{H} + W_T - \sigma T\right).
\end{aligned}$$

(a) $S_T \geq H$: In this case, we have

$$\begin{aligned}
\pi_{\mathcal{V}} &= \frac{1}{\sigma T}\delta(W_T p_H) - \delta(p_H) + \frac{\partial p_H}{\partial \sigma} \\
&= \frac{1}{\sigma T}\delta(W_T) - \delta(1) + 0 \\
&= \frac{1}{\sigma T}(W_T^2 - T) - W_T \\
&= \frac{W_T^2}{\sigma T} - \frac{1}{\sigma} - W_T.
\end{aligned}$$

(b) $S_T < H$: By Example(18.8.2), we have

$$\begin{aligned}
\pi_{\mathcal{V}} &= \frac{1}{\sigma T}\delta(W_T p_H) - \delta(p_H) + \frac{\partial p_H}{\partial \sigma} \\
&= \frac{W_T\delta(p_H) - T p_H}{\sigma T} - \delta(p_H) + p_H \frac{2\log\frac{H}{x}}{\sigma^2 T}\left(-\frac{2}{\sigma}\log\frac{S_T}{H} + W_T - \sigma T\right).
\end{aligned}$$

Thus we have

$$\pi_{\mathcal{V}} = \begin{cases} \frac{W_T^2}{\sigma T} - \frac{1}{\sigma} - W_T, & S_T \geq H \\ \frac{W_T\delta(p_H) - T p_H}{\sigma T} - \delta(p_H) + p_H \frac{2\log\frac{H}{x}}{\sigma^2 T}\left(-\frac{2}{\sigma}\log\frac{S_T}{H} + W_T - \sigma T\right), & S_T < H \end{cases}$$

4. Rho: By differentiation with respect to r, we have

$$\begin{aligned}
\text{Rho} &= -Te^{-rT}E[\phi(S_T)p_H] + e^{-rT}E[\phi'(S_T)TS_T p_H + \phi(S_T)\frac{\partial p_H}{\partial r}] \\
&= -Te^{-rT}E\left[\phi(S_T)p_H] + e^{-rT}E[\phi(S_T)\delta\left(\frac{p_H}{\sigma}\right) + \phi(S_T)\frac{\partial p_H}{\partial r}\right] \\
&= e^{-rT}E\left[\phi(S_T)\left(-Tp_H + \frac{1}{\sigma}\delta(p_H) + \frac{\partial p_H}{\partial r}\right)\right].
\end{aligned}$$

Thus the Malliavin weight for rho is given by

$$\pi_{\text{rho}} = -Tp_H + \frac{1}{\sigma}\delta(p_H) + \frac{\partial p_H}{\partial r}$$

(a) $S_T \geq H$: In this case, we get

$$\pi_{\text{rho}} = -T + \frac{W_T}{\sigma}. \tag{18.8.1}$$

$$\begin{aligned}
p_H &= \exp\left(\frac{2\log\frac{S_T}{H}\ \log\frac{H}{x}}{\sigma^2 T}\right)\\
\delta(p_H) &= p_H\left(W_T - \frac{2\log\frac{H}{x}}{\sigma}\right)\\
\delta(\delta(p_H)) &= \delta(p_H)\left(W_T - \frac{2\log\frac{H}{x}}{\sigma}\right) - p_H T\\
\frac{\partial\delta(p_H)}{\partial x} &= \frac{\partial p_H}{\partial x}\left(W_T - \frac{2\log\frac{H}{x}}{\sigma}\right) + p_H\frac{2}{x\sigma} \quad = \quad \delta\left(\frac{\partial p_H}{\partial x}\right)\\
\frac{\partial p_H}{\partial x} &= p_H\frac{2}{x\sigma^2 T}\left(2\log\frac{H}{x} - \mu T - \sigma W_T\right)\\
\frac{\partial^2 p_H}{\partial x^2} &= \frac{\partial p_H}{\partial x}\frac{2}{x\sigma^2 T}\left(2\log\frac{H}{x} - \mu T - \sigma W_T\right) - p_H\frac{2}{x^2\sigma^2 T}\left(2\log\frac{H}{x} - \mu T - \sigma W_T + 2\right)\\
\frac{\partial p_H}{\partial \sigma} &= p_H\frac{2\log\frac{H}{x}}{\sigma^2 T}\left(-\frac{2}{\sigma}\log\frac{S_T}{H} + W_T - \sigma T\right)\\
\frac{\partial p_H}{\partial r} &= p_H\frac{2\log\frac{H}{x}}{\sigma^2}
\end{aligned}$$

Table 18.1: Summary

(b) $S_T < H$: Since

$$\frac{\partial p_H}{\partial r} = p_H\frac{2\log\frac{H}{x}}{\sigma^2},$$

we have

$$\pi_{\text{rho}} = -Tp_H + p_H\left(\frac{W_T}{\sigma} - \frac{2\log\frac{H}{x}}{\sigma^2}\right) + p_H\frac{2\log\frac{H}{x}}{\sigma^2} \tag{18.8.2}$$

$$= p_H\left(-T + \frac{W_T}{\sigma}\right). \tag{18.8.3}$$

Hence by (18.8.1) and (18.8.3) we have

$$\pi_{\text{rho}} = p_H\left(-T + \frac{W_T}{\sigma}\right).$$

Example 18.8.2. By (18.5.3) we have the following equations.

1. $\delta(\delta(p_H))$:

$$\begin{aligned}
\delta(\delta(p_H)) &= \delta(p_H)W_T - \int_0^T\left\{\frac{\partial p_H}{\partial W_T}\left(W_T - \frac{2\log\frac{H}{x}}{\sigma}\right) + p_H\right\}dt\\
&= \delta(p_H)W_T - \int_0^T\left\{p_H\frac{2\log\frac{H}{x}}{\sigma T}\left(W_T - \frac{2\log\frac{H}{x}}{\sigma}\right) + p_H\right\}dt
\end{aligned}$$

$$
\begin{aligned}
&= \delta(p_H)W_T - \int_0^T \left\{ \delta(p_H)\frac{2\log\frac{H}{x}}{\sigma T} + p_H \right\} dt \\
&= \delta(p_H)\left(W_T - \frac{2\log\frac{H}{x}}{\sigma}\right) - p_H T.
\end{aligned}
$$

Note that this result is trivial from (18.5.5).

2. $\delta(p_H W_T)$:

$$
\begin{aligned}
\delta(p_H W_T) &= W_T\delta(p_H) - \int_0^T p_H dt \\
&= W_T\delta(p_H) - Tp_H.
\end{aligned}
$$

□

Part VI

Appendex

Appendix A

The Multivariate Normal Distribution

Suppose we are given the $p \times 1$ vector $\boldsymbol{\mu}$ and the $p \times p$ matrix $\boldsymbol{\Sigma}$. We shall assume the symmetric matrix $\boldsymbol{\Sigma}$ is positive definite. A p-dimensional normal density for the random vector $\mathbf{X} = [X_1, X_2, \cdots, X_p]^t$ has the form

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}\sqrt{|\boldsymbol{\Sigma}|}} e^{-(\mathbf{x}-\boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})/2}$$

where $\mathbf{x} = (x_1, x_2, \cdots, x_p)^t, -\infty < x_i < \infty, i = 1, 2, \cdots, p$. We shall denote this p-dimensional normal density by $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Example A.1 (Bivariate Normal Density). Let us evaluate the $p = 2$ variate normal density in terms of the individual parameters

$$E[X_1] = \mu_1, \quad E[X_2] = \mu_2, \quad \mathrm{Var}[X_1] = \sigma_{11}, \quad \mathrm{Var}[X_2] = \sigma_{22},$$

and

$$\rho_{12} = \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} = \mathrm{Corr}(X_1, X_2). \tag{A.0.1}$$

The inverse of the covariance

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$$

is

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{\sigma_{11}\sigma_{22} - \sigma_{12}^2} \begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{bmatrix}.$$

Introducing the correlation coefficient ρ_{12} in Eq(A.0.1), we obtain

$$\sigma_{11}\sigma_{22} - \sigma_{12}^2 = \sigma_{11}\sigma_{22}(1 - \rho_{12}^2)$$

and

$$(\mathbf{x}-\boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) = [x_1 - \mu_1, x_2 - \mu_2] \frac{1}{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)} \begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{bmatrix} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}$$

$$
\begin{aligned}
&= \frac{1}{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}\left\{\sigma_{22}(x_1-\mu_1)^2+\sigma_{11}(x_2-\mu_2)^2-2\rho_{12}\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}(x_1-\mu_1)(x_2-\mu_2)\right\} \\
&= \frac{1}{1-\rho_{12}^2}\left\{\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2+\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2-2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right\}.
\end{aligned}
$$

Thus we have the PDF for a bivariate normal distribution as follows:

$$
\begin{aligned}
& f(x_1,x_2) \\
= & \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}}\exp\left\{-\frac{1}{2(1-\rho_{12}^2)}\left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2+\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2-2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right]\right\}.
\end{aligned}
$$

Theorem A.2. If $\mathbf{X}$ is distributed as $N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$, the q linear combinations

$$
\mathbf{AX} = \begin{bmatrix} a_{11} & \cdots & a_{1p} \\ a_{21} & \cdots & a_{2p} \\ & \vdots & \\ a_{q1} & \cdots & a_{qp} \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} = \begin{bmatrix} a_{11}X_1+\cdots+a_{1p}X_p \\ a_{21}X_1+\cdots+a_{2p}X_p \\ \vdots \\ a_{q1}X_1+\cdots+a_{qp}X_p \end{bmatrix}
$$

are distributed as $N_q(\mathbf{A}\boldsymbol{\mu},\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^t)$.

Theorem A.3 (Conditional Density). Let

$$
\mathbf{X} = \begin{bmatrix} \underset{(q\times 1)}{\mathbf{X}_1} \\ ---\\ \underset{(p-q\times 1)}{\mathbf{X}_2} \end{bmatrix}
$$

be distributed as $N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$ with

$$
\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ --- \\ \boldsymbol{\mu}_2 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right]
$$

and $|\boldsymbol{\Sigma}_{22}|>0$. Then the conditional distribution of $\mathbf{X}_1$, given that $\mathbf{X}_2=\mathbf{x}_2$, is normal and has

$$
\text{Mean} = \boldsymbol{\mu}_1+\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2)
$$

and

$$
\text{Covariance} = \boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}.
$$

Note that the covariance does not depend on the value $\mathbf{x}_2$ of the conditioning variable.

Proof. Let

$$
\underset{(p\times p)}{\mathbf{A}} = \left[\begin{array}{c|c} \underset{(q\times q)}{\mathbf{I}} & \underset{q\times(p-q)}{-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}} \\ \hline \underset{(p-q)\times p}{\mathbf{0}} & \underset{(p-q)\times(p-q)}{\mathbf{I}} \end{array}\right].
$$

Then

$$
\mathbf{A}(\mathbf{X}-\boldsymbol{\mu}) = \left[\begin{array}{c} \mathbf{X}_1-\boldsymbol{\mu}_1 \\ \hline \mathbf{X}_2-\boldsymbol{\mu}_2 \end{array}\right] = \left[\begin{array}{c} \mathbf{X}_1-\boldsymbol{\mu}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2-\boldsymbol{\mu}_2) \\ \hline \mathbf{X}_2-\boldsymbol{\mu}_2 \end{array}\right]
$$

is jointly normal with covariance matrix $\mathbf{A\Sigma A}^t$ given by

$$\left[\begin{array}{c|c} \mathbf{I} & -\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1} \\ \hline \mathbf{0} & \mathbf{I} \end{array}\right]\left[\begin{array}{c|c} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \hline \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{array}\right]\left[\begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline -(\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1})^t & \mathbf{I} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{\Sigma}_{11}-\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{\Sigma}_{22} \end{array}\right].$$

To get the covariance matrix, we made use of the fact that

$$\mathbf{\Sigma}_{12}^t \quad = \quad \mathbf{\Sigma}_{21} \quad \text{and} \quad (\mathbf{\Sigma}_{22}^{-1})^t \quad = \quad (\mathbf{\Sigma}_{22}^t)^{-1} \quad = \quad \mathbf{\Sigma}_{22}^{-1}.$$

Since $\mathbf{X}_1-\boldsymbol{\mu}_1-\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(\mathbf{X}_2-\boldsymbol{\mu}_2)$ and $\mathbf{X}_2-\boldsymbol{\mu}_2$ have zero covariance, they are independent. Given that $\mathbf{X}_2=\mathbf{x}_2$, the conditional distribution of $\mathbf{X}_1-\boldsymbol{\mu}_1-\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2)$ is the same as the unconditional distribution of $\mathbf{X}_1-\boldsymbol{\mu}_1-\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(\mathbf{X}_2-\boldsymbol{\mu}_2)$. Hence, given that $\mathbf{X}_2=\mathbf{x}_2$, $\mathbf{X}_1$ is distributed as

$$N_q(\boldsymbol{\mu}_1+\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2), \mathbf{\Sigma}_{11}-\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21})$$

■

Appendix B

The Brownian Motion

Consider a sequence $\{X_i\}_{i\in\mathbb{N}}$ of i.i.d random variables with mean zero and finite variance σ^2. For each $t \geq 0$ and $n \in \mathbb{N}$, define process $Z_n(t)$ by the partial sums,

$$Z_n(t) \quad = \quad \frac{1}{\sigma\sqrt{n}} \sum_{j=1}^{[nt]} X_j, \tag{B.0.1}$$

where $[nt]$ is the smallest integer less or equal to nt.

Theorem B.1. Let Z_n be defined by (B.0.1) and $0 < t_1 < \cdots < t_k < \infty$. Then

$$(Z_n(t_1), \cdots, Z_n(t_k)) \quad \xrightarrow{\mathrm{d}} \quad (W(t_1), \cdots, W(t_k)) \qquad \text{as } n \to \infty$$

where $\{W_t\}$ is standard Brownian motion.

Proof. We will show the case $k = 2$. Let $s = t_1$ and $t = t_2$. we have to show that

$$\frac{1}{\sigma\sqrt{n}} \left(\sum_{j=1}^{[ns]} X_j, \sum_{j=1}^{[nt]} X_j \right) \quad \xrightarrow{\mathrm{d}} \quad (W(s), W(t)) \qquad \text{as } n \to \infty.$$

It is equivalent to proving that

$$\frac{1}{\sigma\sqrt{n}} \left(\sum_{j=1}^{[ns]} X_j, \sum_{j=[ns]+1}^{[nt]} X_j \right) \quad \xrightarrow{\mathrm{d}} \quad (W(s), W(t) - W(s)) \qquad \text{as } n \to \infty.$$

We employ the method of characteristic functions and obtain that

$$\begin{aligned} & \lim_{n\to\infty} E\left[\exp\left(\frac{iu}{\sigma\sqrt{n}} \sum_{j=1}^{[ns]} X_j + \frac{iv}{\sigma\sqrt{n}} \sum_{j=[ns]+1}^{[nt]} X_j \right) \right] \\ = \quad & \lim_{n\to\infty} E\left[\exp\left(\frac{iu}{\sigma\sqrt{n}} \sum_{j=1}^{[ns]} X_j \right) \right] \cdot \lim_{n\to\infty} E\left[\exp\left(\frac{iv}{\sigma\sqrt{n}} \sum_{j=[ns]+1}^{[nt]} X_j \right) \right] \end{aligned}$$

$$
\begin{aligned}
&= \lim_{n\to\infty} E\left[\exp\left(\frac{iu\sqrt{s}}{\sigma\sqrt{ns}}\sum_{j=1}^{[ns]} X_j\right)\right]\cdot \lim_{n\to\infty} E\left[\exp\left(\frac{iv\sqrt{t}}{\sigma\sqrt{nt}}\sum_{j=[ns]+1}^{[nt]} X_j\right)\right] \\
&= \exp\left(-\frac{su^2}{2}\right)\cdot\exp\left(-\frac{(t-s)v^2}{2}\right).
\end{aligned}
$$

■

Definition B.2 (Brownian Motion). We shall say a sequence of random variables, W_t, for $t \geq 0$, is a Brownian motion if $W_0 = 0$, and for every t and s, with $s < t$, we have that

$$W_t - W_s \quad \sim \quad N(0, t-s)$$

and $W_t - W_s$ is independent of W_r for $r \leq s$.

Appendix C

The Girsanov Theorem

Example C.1 (Exponential Martingale). Suppose that the Novikov condition

$$E\left[\exp\left(\frac{1}{2}\int_0^t \gamma^2(s,\omega)ds\right)\right] < \infty$$

holds. Define

$$\begin{aligned} M_t &= \exp\Big\{-\int_0^t \gamma(s,\omega)dW_s - \frac{1}{2}\int_0^t \gamma^2(s,\omega)ds\Big\}, \quad \gamma^2 = \gamma\cdot\gamma, \\ &= \exp\Big\{Z_t - \frac{1}{2}\langle Z\rangle_t\Big\}, \quad \text{where } Z_t = -\int_0^t \gamma(s,\omega)dW_s. \end{aligned}$$

Then, by Ito formula, we get

$$\begin{aligned} dM_t &= M_t(-\gamma_t dW_t - \frac{1}{2}\gamma_t^2 dt) + \frac{1}{2}M_t\gamma_t^2 dt \\ &= -M_t\gamma_t dW_t, \\ d(\log M_t) &= \frac{dM_t}{M_t} - \frac{1}{2}(dM_t)^2 \\ &= -\gamma_t dW_t - \frac{1}{2}\gamma_t^2 dt, \\ \log M_t &= -\int_0^t \gamma_s dW_s - \frac{1}{2}\int_0^t \gamma_s^2 ds, \\ M_t &= \exp\Big(-\int_0^t \gamma_s dW_s - \frac{1}{2}\int_0^t \gamma_s^2 ds\Big). \end{aligned} \tag{C.0.1}$$

Hence M_t is a (exponential) martingale. The following theorem show that M_t can be a Radon-Nikodym derivative $\frac{dQ}{dP}$.

Theorem C.2 (Girsanov). Assume that $\gamma(s,\omega)$ satisfies Novikov's condition:

$$E_P\left[\exp\left(\frac{1}{2}\int_0^t \gamma^2(s,\omega)ds\right)\right] < \infty.$$

Put

$$M_t = \exp\Big\{-\int_0^t \gamma(s,\omega)dW_s - \frac{1}{2}\int_0^t \gamma^2(s,\omega)ds\Big\} \quad t \le T \le \infty.$$

Define the measure Q by

$$dQ(\omega) = M_T(\omega)dP(\omega).$$

Then we have that

$$\widetilde{W}_t := W_t + \int_0^t \gamma(s,\omega)ds$$

is Brownian motion with respect to the probability measure Q.

Proof.

1. For each t, define a new measure Q_t as following: for $A \in \mathcal{F}_s, s \leq t$

$$Q_t(A) = \int_A M_t dP.$$

   Since M_t is a martingale, for $A \in \mathcal{F}_s$

$$\begin{aligned} Q_t(A) &= \int_A M_t dP \\ &= \int_A E[M_t|\mathcal{F}_s]dP, \quad \text{by def. of conditional expectation} \\ &= \int_A M_s dP. \end{aligned}$$

   Thus Q is defined independent of t. i.e. Q_t and Q_s coincide on $\mathcal{F}_s$, for $s \leq t$. $\{Q_t\}$ is a consistent family of probability measures.

2. $M_t\widetilde{W}_t$ is a P-martingale.

$$\begin{aligned} d(M_t\widetilde{W}_t) &= \widetilde{W}_t dM_t + M_t d\widetilde{W}_t + dM_t \cdot d\widetilde{W}_t \\ &= -\widetilde{W}_t M_t \gamma_t dW_t + M_t(dW_t + \gamma dt) + (-M_t\gamma_t dW_t) \cdot (dW_t + \gamma dt) \\ &= M_t(-\widetilde{W}_t\gamma_t + 1)dW_t. \end{aligned}$$

3. $\widetilde{W}_t$ is a Q-martingale. For $A \in \mathcal{F}_s$

$$\begin{aligned} \int_A \widetilde{W}_t dQ &= \int_A \widetilde{W}_t M_t dP \\ &= \int_A \widetilde{W}_s M_s dP \\ &= \int_A \widetilde{W}_s dQ. \end{aligned}$$

   By definition of the conditional expectation, $E_Q[\widetilde{W}_t|\mathcal{F}_s] = \widetilde{W}_s$.

4. Since $\widetilde{W}_t$ is a martingale and $d\langle\widetilde{W}\rangle_t = dt$, $\widetilde{W}_t$ is Q-Brownian motion.

■

Example C.3. Let

$$M_t = \exp\left(-\mu W_t - \frac{1}{2}\mu^2 t\right),$$

and a probability measure Q by

$$Q(A) = \int_A M_t dP.$$

The process $\widetilde{W}_t = \mu t + W_t$ is a standard Brownian motion under the probability measure Q. Note that

$$\begin{aligned} dQ &= M_t dP = \exp\left(-\mu W_t - \frac{1}{2}\mu^2 t\right) dP = \exp\left(-\mu \widetilde{W}_t + \frac{1}{2}\mu^2 t\right) dP, \\ dP &= \exp\left(\mu W_t + \frac{1}{2}\mu^2 t\right) dQ = \exp\left(\mu \widetilde{W}_t - \frac{1}{2}\mu^2 t\right) dQ, \quad t \le T. \end{aligned}$$

Example C.4. Suppose we are given a probability measure P and a Brownian motion W_t under P. For a martingale process M_t with

$$dM_t = \mu M_t dW_t, \quad M_0 = 1, \tag{C.0.2}$$

define a new measure Q by

$$Q(A) = E_P[\mathbf{1}_A M_t], \quad A \in \mathcal{F}_t.$$

Note that (C.0.2) implies that

$$M_t = \exp\left(-\frac{1}{2}\mu^2 t + \mu W_t\right).$$

Since W_t is distributed as $N(0,t)$, we therefore have that

$$\begin{aligned} Q(W_t < x) &= E_P\left[\mathbf{1}_{\{W_t<x\}} M_t\right] \\ &= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{x} \exp\left(-\frac{1}{2}\mu^2 t + \mu z\right) \exp\left(-\frac{z^2}{2t}\right) dz \\ &= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{x} \exp\left(-\frac{1}{2t}(z-\mu t)^2\right) dz \\ &= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{x-\mu t} \exp\left(-\frac{1}{2t}y^2\right) dy \\ &= P(W_t < x - \mu t). \end{aligned}$$

We have shown that the probability that $W_t < x$ in the new measure Q is equal to the probability that $W_t + \mu t < x$ in the old measure P. In other words, the process W_t in Q is equivalent to $W_t + \mu t$ in P. Hence we can see that W_t is a Brownian motion with drift μ in Q, i.e.

$$W_t = \widetilde{W}_t + \mu t,$$

for the Brownian motion $\widetilde{W}_t$ in Q.

Example C.5. Suppose $Y(t) = \left[\begin{smallmatrix} Y_1(t) \\ Y_2(t) \end{smallmatrix}\right] \in \mathbb{R}^2$ is given by

$$\begin{aligned} dY_1(t) &= 2dt + dW_1(t) + dW_2(t) \\ dY_2(t) &= 4dt + dW_1(t) - dW_2(t). \end{aligned}$$

i.e.

$$dY(t) = \begin{bmatrix} 2 \\ 4 \end{bmatrix} dt + \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} dW(t); \quad W(t) = \begin{bmatrix} W_1(t) \\ W_2(t) \end{bmatrix}.$$

The equation

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

has the unique solution $u_1 = 3, u_2 = -1$. Hence we put

$$dQ = \exp\Big(- \big(u_1 W_1(t) + u_2 W_2(t)\big) - \frac{1}{2}(u_1^2 + u_2^2)t \Big),$$

and

$$d\widetilde{W}(t) = \begin{bmatrix} 3 \\ -1 \end{bmatrix} dt + dW(t).$$

We conclude that $\widetilde{W}(t)$ is a Brownian motion w.r.t. the probability measure Q and

$$dY(t) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} d\widetilde{W}(t).$$

Appendix D

Transforming Variables

Suppose that two processes are given by

$$\begin{aligned} dS_1 &= rS_1 dt + \sigma_1 S_1 dW_1, \\ dS_2 &= rS_2 dt + \sigma_2 S_2 dW_2, \end{aligned}$$

where the instantaneous correlation between the Wiener processes, dW_1 and dW_2, is ρ. This gives us

$$\begin{aligned} d\log S_1 &= (r - \frac{1}{2}\sigma_1^2)dt + \sigma_1 dW_1, \\ d\log S_2 &= (r - \frac{1}{2}\sigma_2^2)dt + \sigma_2 dW_2. \end{aligned}$$

✎ **Transforming Variables**

The covariance matrix Σ of random variables X_1 and X_2, is given by

$$\Sigma := \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

1. Let

$$P = \begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ 0 & \frac{1}{\sigma_2} \end{bmatrix}.$$

Then we have

$$P^t \Sigma P = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} := \Sigma'$$

2. The correlation matrix Σ' is diagonalized as follows. Let

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Then

$$Q^t \Sigma' Q = \begin{bmatrix} 1+\rho & 0 \\ 0 & 1-\rho \end{bmatrix}.$$

3. Thus we have

$$
\begin{aligned}
Q^t P^t \Sigma P Q &= \begin{bmatrix} 1+\rho & 0 \\ 0 & 1-\rho \end{bmatrix}, \\
\left(\sqrt{2}\,\sigma_1\sigma_2 PQ\right)^t \Sigma \left(\sqrt{2}\,\sigma_1\sigma_2 PQ\right) &= \begin{bmatrix} 2(1+\rho)\sigma_1^2\sigma_2^2 & 0 \\ 0 & 2(1-\rho)\sigma_1^2\sigma_2^2 \end{bmatrix}, \qquad \text{(D.0.1)} \\
PQ &= \begin{bmatrix} \frac{1}{\sigma_1\sqrt{2}} & \frac{1}{\sigma_1\sqrt{2}} \\ \frac{1}{\sigma_2\sqrt{2}} & -\frac{1}{\sigma_2\sqrt{2}} \end{bmatrix}, \\
\sqrt{2}\,\sigma_1\sigma_2 PQ &= \begin{bmatrix} \sigma_2 & \sigma_2 \\ \sigma_1 & -\sigma_1 \end{bmatrix}.
\end{aligned}
$$

4. Put

$$
\begin{aligned}
\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} &:= \sqrt{2}\,\sigma_1\sigma_2 \Big(PQ\Big)^t \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \\
&= \begin{bmatrix} \sigma_2 & \sigma_1 \\ \sigma_2 & -\sigma_1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}.
\end{aligned}
$$

The covariance matrix of $[Y_1, Y_2]^t$ is given by (D.0.1).

Appendix E

The Linear Stochastic Differential Equations

Theorem E.1. For the solution $\mathbf{X}_t$ of the linear stochastic differential equation

$$d\mathbf{X}_t = \Big(\mathbf{A}(t)\mathbf{X}_t + \mathbf{a}(t)\Big)dt + \mathbf{B}(t)d\mathbf{W}_t, \mathbf{X}_{t_0} = \mathbf{c},$$

we have, under the assumption $E|\mathbf{c}|^2 < \infty$,

$$\mathbf{m}_t := E\mathbf{X}_t = \mathbf{\Phi}(t)\left(E\mathbf{c} + \int_{t_0}^{t} \mathbf{\Phi}(s)^{-1}\mathbf{a}(s)ds\right).$$

Therefore, $\mathbf{m}_t$ is the solution of the deterministic linear differential equation.

$$\dot{\mathbf{m}}_t = \mathbf{A}(t)\mathbf{m}_t + \mathbf{a}(t), \quad m_{t_0} = E\mathbf{c}.$$

$$\begin{aligned}\mathbf{K}(s,t) &= E(\mathbf{X}_s - E\mathbf{X}_s)(\mathbf{X}_t - E\mathbf{X}_t)^t \\ &= \mathbf{\Phi}(t)\left(E(\mathbf{c} - E\mathbf{c})(\mathbf{c} - E\mathbf{c})^t + \int_{t_0}^{\min(s,t)} \mathbf{\Phi}(s)^{-1}\mathbf{B}(s)\mathbf{B}(s)^t\left(\mathbf{\Phi}(t)^{-1}\right)^t ds\right)\mathbf{\Phi}(t)^t \qquad \text{(E.0.1)}\end{aligned}$$

In particular, the covariance matrix of the components of $\mathbf{X}_t$

$$\mathbf{K}(t) = \mathbf{K}(t,t) = E(\mathbf{X}_t - E\mathbf{X}_t)(\mathbf{X}_t - E\mathbf{X}_t)^t$$

is the unique symmetric nonnegative-definite solution of the matrix equation

$$\dot{\mathbf{K}}(t) = \mathbf{A}(t)\mathbf{K}(t) + \mathbf{K}(t)\mathbf{A}(t)^t + \mathbf{B}(t)\mathbf{B}(t)^t, \quad \mathbf{K}(t_0) = E(\mathbf{c} - E\mathbf{c})(c - E\mathbf{c})^t.$$

Proof.

$$\begin{aligned}\dot{\mathbf{m}}_t &= \mathbf{\Phi}'(t)\left(E\mathbf{c} + \int_{t_0}^{t} \mathbf{\Phi}(s)^{-1}\mathbf{a}(s)ds\right) + \mathbf{\Phi}(t)\mathbf{\Phi}(t)^{-1}\mathbf{a}(t) \\ &= \mathbf{A}(t)\mathbf{\Phi}(t)\left(E\mathbf{c} + \int_{t_0}^{t} \mathbf{\Phi}(s)^{-1}\mathbf{a}(s)ds\right) + \mathbf{a}(t)\end{aligned}$$

$$= \quad \mathbf{A}(t)\mathbf{m}(t) + \mathbf{a}(t).$$

Let

$$\Gamma(t) = E(\mathbf{c} - E\mathbf{c})(\mathbf{c} - E\mathbf{c})^t + \int_{t_0}^{\min(s,t)} \mathbf{\Phi}(s)^{-1}\mathbf{B}(s)\mathbf{B}(s)^t \left(\mathbf{\Phi}(t)^{-1}\right)^t ds.$$

Then $\mathbf{K}(t,t) = \mathbf{\Phi}(t)\Gamma(t)\mathbf{\Phi}(t)^t$. If we differentiate Eq. (E.0.1) with respect to t, we get that

$$\begin{aligned}
\dot{\mathbf{K}}(t) &= \mathbf{\Phi}'(t)\Gamma(t)\mathbf{\Phi}(t)^t + \mathbf{\Phi}(t)\Gamma'(t)\mathbf{\Phi}(t)^t + \mathbf{\Phi}(t)\Gamma(t)\mathbf{\Phi}'(t)^t \\
&= \mathbf{A}(t)\mathbf{\Phi}(t)\Gamma(t)\mathbf{\Phi}(t)^t + \mathbf{\Phi}(t)\mathbf{\Phi}(t)^{-1}\mathbf{B}(t)\mathbf{B}(t)^t \left(\mathbf{\Phi}(t)^{-1}\right)^t \mathbf{\Phi}(t)^t + \mathbf{\Phi}(t)\Gamma(t)\mathbf{\Phi}(t)^t\mathbf{A}(t)^t \\
&= \mathbf{A}(t)\mathbf{K}(t) + \mathbf{B}(t)\mathbf{B}(t)^t + \mathbf{K}(t)\mathbf{A}(t)^t.
\end{aligned}$$

■

Appendix F

The Infinitesimal Operator

Definition F.1. Let $X(t)$ be any Markov process with stationary transition probabilities and $f(x)$ be a function(sufficiently smooth). Define

$$\mathcal{A}f(x) = \lim_{t\to 0} \frac{E_x\big[f(X(t))\big] - f(x)}{t}.$$

Denote the resulting function by $Af(x)$. A is called **infinitesimal operator** and summarizes the behavior of the transition probabilities as $t \to 0$. The set of functions $f : \mathbb{R}^n \to \mathbb{R}$ such that the limit exists at x is denoted by $\mathcal{D}_A(x)$, while $\mathcal{D}_A$ denotes the set of functions for which the limits exists for all $x \in \mathbb{R}^n$.

Theorem F.2. Let X_t be the Ito diffusion

$$d\mathbf{X}_t = \mathbf{b}(X_t)dt + \boldsymbol{\sigma}(\mathbf{X}_t)d\mathbf{B}_t.$$

i.e.

$$d\begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} dt + \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1m} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nm} \end{bmatrix} \begin{bmatrix} dB_1 \\ \vdots \\ dB_m \end{bmatrix}$$

If $f \in C_0^2(\mathbb{R}^n)$, i.e. $f \in C^2(\mathbb{R}^n)$ and f has compact support, then $f \in \mathcal{D}_A$ and

$$\mathcal{A}f(x) = \sum_i b_i(x)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j}(\boldsymbol{\sigma}\boldsymbol{\sigma}^T)_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Proof. See Oksendal(2003) page 117. ■

In generally, for a given SDE

$$d\mathbf{X}_t = \boldsymbol{\mu}(t, \mathbf{X}_t)dt + \boldsymbol{\sigma}(t, \mathbf{X}_t)d\mathbf{W}_t,$$

the infinitesimal operator of $\mathbf{X}$, is given by

$$\mathcal{A}f(t, \mathbf{x}) = \sum_{i=1}^{n} \mu_i(t, \mathbf{x})\frac{\partial f}{\partial x_i}(t, \mathbf{x}) + \frac{1}{2}\sum_{i,j=1}^{n}\left(\boldsymbol{\sigma}\boldsymbol{\sigma}^t\right)_{ij}(t, \mathbf{x})\frac{\partial^2 f}{\partial x_i \partial x_j}(t, \mathbf{x}),$$

and the adjoint operator $\mathcal{A}^*$ is defined by

$$\mathcal{A}^* f(t,\mathbf{x}) \quad = \quad -\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\Big(\mu_i(t,\mathbf{x}) f(t,\mathbf{x})\Big) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}\bigg(\Big(\boldsymbol{\sigma}\boldsymbol{\sigma}^t\Big)_{ij}(t,\mathbf{x}) f(t,\mathbf{x})\bigg),$$

for any function $f \in C^2(\mathbb{R}^n)$.

The Feynman-Kac formula(1-dim) is given as follows.

Theorem F.3 (Feynman-Kac). Assume that F is a solution to the boundary value problem

$$\begin{cases} \dfrac{\partial F}{\partial t} + \mu(t,x)\dfrac{\partial F}{\partial x} + \dfrac{1}{2}\sigma^2(t,x)\dfrac{\partial^2 F}{\partial x^2}(t,x) - rF(t,x) = 0, \\ F(T,x) = \Phi(x). \end{cases}$$

Assume furthermore that the process $\sigma(s,X_s)\frac{\partial F}{\partial x}(s,X_s)$ is in L^2, where X is defined by a SDE

$$dX_s \quad = \quad \mu(s,X_s)ds + \sigma(s,X_s)dW_s, \quad X_t = x(s \ge t).$$

Then F has the representation

$$F(t,x) \quad = \quad e^{-r(T-t)} E_{t,x}\Big[\Phi(X_T)\Big].$$

✎ **Feynman-Kac formula**
The n-dimensional version can be proved as followings:

1. Define the stochastic process $\mathbf{X}$ on the time interval $[t,T]$ as the solution to the SDE

$$d\mathbf{X}_s \quad = \quad \mu(s,\mathbf{X}_s)ds + \boldsymbol{\sigma}(s,\mathbf{X}_s)d\mathbf{W}_s, \quad \mathbf{X}_t = \mathbf{x}$$

2. Assume that $F : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}$ is a solution to the boundary value problem

$$\begin{cases} \dfrac{\partial F}{\partial t} + \displaystyle\sum_{i=1}^{n} \mu_i \dfrac{\partial F}{\partial x_i} + \dfrac{1}{2}\sum_{i,j=1}^{n} \left(\boldsymbol{\sigma}\boldsymbol{\sigma}^t\right)_{ij} \dfrac{\partial^2 F}{\partial x_i \partial x_j} - rF = 0, \\ F(T,\mathbf{x}) = \Phi(\mathbf{x}). \end{cases} \tag{F.0.1}$$

   This equation equivalent to

$$\frac{\partial F}{\partial t}(t,\mathbf{x}) + \mathcal{A}F(t,\mathbf{x}) - rF(t,\mathbf{x}) = 0$$

3. Applying the Ito lemma to the process $e^{-rs}F(s,\mathbf{X}(s))$ gives us

$$\begin{aligned} e^{-rT}F(T,\mathbf{X}_T) \quad &= \quad e^{-rt}F(t,\mathbf{X}_t) + \int_t^T e^{-rs}\left\{\frac{\partial F}{\partial t}(s,\mathbf{X}_s) + \mathcal{A}F(s,\mathbf{X}_s) - rF(s,\mathbf{X}_s)\right\} ds \\ &\quad + \sum_{i=1}^{n}\int_t^T e^{-rs}\frac{\partial F}{\partial x_i}\boldsymbol{\sigma}_i d\mathbf{W}_s \\ &= \quad e^{-rt}F(t,\mathbf{X}_t) + \sum_{i=1}^{n}\int_t^T e^{-rs}\frac{\partial F}{\partial x_i}\boldsymbol{\sigma}_i d\mathbf{W}_s \end{aligned}$$

   where $\boldsymbol{\sigma}_i$ is the ith row of the matrix $\boldsymbol{\sigma}$.

4. Assume that the process
$$\sigma_{ij}\frac{\partial F}{\partial x_i}$$
is in L^2.

5. If we take expected value, the stochastic integral will vanish. The initial value $X_t = \mathbf{x}$ and boundary condition $F(T, \mathbf{x}) = \Phi(\mathbf{x})$ will eventually leave us with the formula
$$e^{-rT}E_{t,\mathbf{x}}\Big[F(T, \mathbf{X}_T)\Big] = e^{-rt}F(t, \mathbf{x}).$$
This gives us that
$$F(t, \mathbf{x}) = e^{-r(T-t)}E_{t,\mathbf{x}}\Big[\Phi(\mathbf{X}_T)\Big]$$

F.1 Kolmogorov Backward Equations

Theorem F.4. Let X be a solution of to SDE
$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t. \tag{F.1.1}$$
Then the transition probability $p(s, y; t, B) = P\Big(X_t \in B|X_s = y), (t \geq s)$ are given as the solution to the equation
$$\left(\frac{\partial p}{\partial s} + \mathcal{A}p\right)(s, y; t, B) = 0, \quad p(t, y; t, B) = \mathbf{1}_B(y).$$
Note that when $\tau = t - s$,
$$\frac{\partial p}{\partial \tau} = \mathcal{A}p.$$

For the transition density function, we can prove the following corresponding theorem.

Theorem F.5. Let X be a solution of Eq(F.1.1). Assume that the measure $P(s, y; t, dx)$ has a density $p(s, y; t, x)$. Then we have
$$\left(\frac{\partial p}{\partial s} + \mathcal{A}p\right)(s, y; t, B) = 0, \quad p(s, y; t, x) \to \delta_x, \text{ as } s \to t.$$

The reason that equations are called backward equations is that the differential operator is working on the backward variables (s, y). When the differential operator is working on the forward variables (t, x), the equations is called forward equations.

F.2 Kolmogorov Forward Equations

Theorem F.6. Assume that the solution X of Eq(F.1.1) has a transition density $p(s, y; t, x)$. Then p will satisfy the Kolmogorov forward equation
$$\frac{\partial p}{\partial t}(s, y; t, x) = \mathcal{A}^*p(s, y; t, x), \quad p(s, y; t, x) \to \delta_y, \text{ as } t \downarrow s.$$
This equation is also known as the Fokker-Planck equation.

Appendix G

Tridiagonal Matrix

Definition G.1 (Strictly diagonally dominant). The $n \times n$ matrix A is said to be strictly diagonally dominant when

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|$$

holds for each $i = 1, 2, \cdots, n$.

Theorem G.2. A strictly diagonally dominant matrix A is non-singular. Moreover, in this case, Gaussian elimination can be performed on any linear system of the form $A\mathbf{x} = \mathbf{b}$ to obtain its unique solution without row or column exchanges, and the computations are stable with respect to the growth of roundoff errors.

Proof. Suppose $A\mathbf{x} = 0$ has a non trivial solution $\mathbf{x} = (x_i)$. Let k be an index for which

$$0 < |x_k| = \max_{1 \leq j \leq n} |x_j|$$

Since $\sum_{j=1}^{n} a_{ij} x_j = 0$ for each $i = 1, 2, \cdots, n$, we have, when $i = k$,

$$\begin{aligned}
a_{kk} x_k &= -\sum_{\substack{j=1 \\ j \neq k}}^{n} a_{kj} x_j, \\
|a_{kk}||x_k| &\leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |a_{kj}||x_j| \\
|a_{kk}| &\leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |a_{kj}| \frac{|x_j|}{|x_k|} \leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |a_{kj}|
\end{aligned}$$

This inequality contradicts the strict diagonal dominance of A. Consequently, the only solution of $A\mathbf{x} = 0$ is $\mathbf{x} = 0$.
For the remainder of the proof, see Burden(1997) page 404. ■

Theorem G.3. If A is strictly diagonally dominant, then for any choice of $\mathbf{x}^{(0)}$, both the Jacobi and Gauss-Seidel methods give sequences $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ that converge to the unique solution of $A\mathbf{x} = \mathbf{b}$.

Proof. See Burden(1997) page 451. ■

✎ **LU Decomposition**
For a given matrix A, once the LU decomposition such that

$$A = LU,$$

where L is a lower triangular matrix and U is upper triangular, the linear system

$$A\mathbf{x} = \mathbf{y}$$

can be solve as following way.

For the solution $\mathbf{z}$ of equation

$$L\mathbf{z} = \mathbf{y},$$

the solution $\mathbf{x}$ can be obtained from equation

$$U\mathbf{x} = \mathbf{z}.$$

G.1 An LU solver for tridiagonal systems

Suppose we are given the following tridiagonal matrix A:

$$A = \begin{bmatrix} a_1 & c_1 & 0 & 0 & \cdots & 0 & 0 \\ b_1 & a_2 & c_2 & 0 & \cdots & 0 & 0 \\ 0 & b_2 & a_3 & c_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & a_{n-1} & c_{n-1} \\ 0 & 0 & 0 & 0 & \cdots & b_{n-1} & a_n \end{bmatrix}$$

We find that if we write a tridiagonal matrix in the LU form

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ l_1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & l_2 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & l_{n-1} & 1 \end{bmatrix} \begin{bmatrix} d_1 & u_1 & 0 & \cdots & 0 & 0 \\ 0 & d_2 & u_2 & \cdots & 0 & 0 \\ 0 & 0 & d_3 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & u_{n-2} & 0 \\ 0 & 0 & 0 & \cdots & d_{n-1} & u_{n-1} \\ 0 & 0 & 0 & \cdots & 0 & d_n \end{bmatrix},$$

the d_i's u_i's and l_i's are given by

$$\begin{aligned} d_1 &= a_1, \qquad d_i = a_i - \frac{c_{i-1}b_{i-1}}{d_{i-1}}, \quad 2 \le i \le n, \\ u_i &= c_i, \qquad l_i = \frac{b_i}{d_i}, \quad 1 \le i \le n-1. \end{aligned}$$

Also the z_i's and x_i's are given by

$$\begin{aligned} z_1 &= y_1, \qquad z_i = y_i - l_{i-1}\, z_{i-1} = y_i - \frac{b_{i-1}}{d_{i-1}}\, z_{i-1}, \quad 2 \le i \le n, \\ x_n &= \frac{z_n}{d_n}, \qquad x_i = \frac{z_i - c_i\, x_{i+1}}{d_i}, \quad i = n-1, n-2, \cdots, 1. \end{aligned}$$

Remark G.1.1. We are obtaining the variables in the following order.

$$\begin{aligned} &d_1 \to d_2 \to \cdots \to d_n \to \\ &z_1 \to z_2 \to \cdots \to z_n \to \\ &x_n \to x_{n-1} \to \cdots \to x_1. \end{aligned}$$

In fact, we need not solve u_i's and l_i's explicitly.

Appendix H

Principal Components Analysis

Let the random vector $\mathbf{X} = \left[X_1, X_2, \cdots, X_p\right]^t$ ($X_i \in \mathbb{R}$, random variable) which has the covariance matrix $\mathbf{\Sigma} = (\sigma_{ij})$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$. For $a_{ij} \in \mathbb{R}$, consider the linear combinations

$$\begin{aligned} Y_1 &= \mathbf{a}_1^t\mathbf{X} = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p \\ Y_2 &= \mathbf{a}_2^t\mathbf{X} = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p \\ &\vdots \\ Y_p &= \mathbf{a}_p^t\mathbf{X} = a_{p1}X_1 + a_{p2}X_2 + \cdots + a_{pp}X_p. \end{aligned}$$

Let

$$A = (a_{ij}) = \begin{bmatrix} \mathbf{a}_1^t \\ \mathbf{a}_2^t \\ \vdots \\ \mathbf{a}_p^t \end{bmatrix}.$$

Then $\mathbf{Y} = A\mathbf{X}$.

1. First principal component = linear combination $a_1^t\mathbf{X}$ that maximizes $\mathrm{Var}(a_1^t\mathbf{X})$ subject to $\|\mathbf{a}_1\| = 1$
2. Second principal component = linear combination $a_2^t\mathbf{X}$ that maximizes $\mathrm{Var}(a_2^t\mathbf{X})$ subject to $\|\mathbf{a}_2\| = 1$ and $\mathrm{Cov}(a_1^t\mathbf{X}, a_2^t\mathbf{X}) = 0$.
3. ith principal component = linear combination $a_i^t\mathbf{X}$ that maximizes $\mathrm{Var}(a_i^t\mathbf{X})$ subject to $\|\mathbf{a}_i\| = 1$ and $\mathrm{Cov}(a_k^t\mathbf{X}, a_i^t\mathbf{X}) = 0$ for $k < i$.

Put

$$A = \begin{bmatrix} \text{the first eigenvector} \\ \text{the second eigenvector} \\ \vdots \\ \text{the } p\text{th eigenvector} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^t \\ \mathbf{a}_2^t \\ \vdots \\ \mathbf{a}_p^t \end{bmatrix},$$

and $\mathbf{Y} = A\mathbf{X}$. Then

$$\begin{aligned}
\mathbf{\Sigma} &= A^t D A, \quad \text{where } D = \text{diag}(\lambda_1, \cdots, \lambda_p), \\
&= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_p \end{bmatrix} DA \\
&= \begin{bmatrix} \lambda_1 \mathbf{a}_1 & \lambda_2 \mathbf{a}_2 & \cdots & \lambda_p \mathbf{a}_p \end{bmatrix} \begin{bmatrix} \mathbf{a}_1^t \\ \mathbf{a}_2^t \\ \vdots \\ \mathbf{a}_p^t \end{bmatrix}, \\
&= \lambda_1 \mathbf{a}_1 \mathbf{a}_1^t + \lambda_2 \mathbf{a}_2 \mathbf{a}_2^t + \cdots + \lambda_p \mathbf{a}_p \mathbf{a}_p^t, \\
\text{Var}(\mathbf{Y}) &= A \mathbf{\Sigma} A^t = D.
\end{aligned}$$

Hence we find that

$$\begin{aligned}
\sigma_{ij} &= \left(\sum_{k=1}^{p} \lambda_k \mathbf{a}_k \mathbf{a}_k^t \right)_{ij} \\
&= \sum_{k=1}^{p} \lambda_k a_{ki} a_{k_j}.
\end{aligned}$$

For example, the sum of squares of the first component of each eigenvector is σ_{11} up to eigenvalues. Note that[1]

$$\begin{aligned}
\sum_{i=1}^{p} \text{Var} X_i &= \sigma_{11} + \sigma_{22} + \cdots + \sigma_{pp} \\
&= \text{tr}(\mathbf{\Sigma}) \\
&= \text{tr}(A \mathbf{\Sigma} A^t) \\
&= \text{tr}(D) \\
&= \lambda_1 + \lambda_2 + \cdots + \lambda_p \\
&= \sum_{i=1}^{p} \text{Var} Y_i \\
&= \text{total population variance.}
\end{aligned}$$

Example H.0.2 (미국 50개주 범죄 자료를 통한 주성분 분석).

State	murder	rape	robbery	assault	burglary	larceny	auto
WEST VIRGINIA	6	13.2	42.2	90.9	597.4	1341.7	163.3
NORTH DAKOTA	0.9	9	13.3	43.8	446.1	1843	144.7
SOUTH DAKOTA	2	13.5	17.9	155.7	570.5	1704.4	147.5
MISSISSIPPI	14.3	19.6	65.7	189.1	915.6	1239.9	144.4
KENTUCKY	10.1	19.1	81.1	123.3	872.2	1662.1	245.4
PENNSYLVANIA	5.6	19	130.3	128	877.5	1624.1	333.2
ARKANSAS	8.8	27.6	83.2	203.4	972.6	1862.1	183.4
NEBRASKA	3.9	18.1	64.7	112.7	760	2316.1	249.1
ALABAMA	14.2	25.2	96.8	278.3	1135.5	1881.9	280.7
WISCONSIN	2.8	12.9	52.2	63.7	846.9	2614.2	220.7

[1] $\text{tr}(AB) = \text{tr}(BA)$

NEW HAMPSHIRE	3.2	10.7	23.2	76	1041.7	2343.9	293.4
TENNESSEE	10.1	29.7	145.8	203.9	1259.7	1776.5	314
IOWA	2.3	10.6	41.2	89.8	812.5	2685.1	219.9
NORTH CAROLINA	10.6	17	61.3	318.3	1154.1	2037.8	192.1
VIRGINIA	9	23.3	92.1	165.7	986.2	2521.2	226.7
VERMONT	1.4	15.9	30.8	101.2	1348.2	2201	265.2
MONTANA	5.4	16.7	39.2	156.8	804.9	2773.2	309.2
WYOMING	5.4	21.9	39.7	173.9	811.6	2772.2	282
MAINE	2.4	13.5	38.7	170	1253.1	2350.7	246.9
IDAHO	5.5	19.4	39.6	172.5	1050.8	2599.6	237.6
OKLAHOMA	8.6	29.2	73.8	205	1288.2	2228.1	326.8
MINNESOTA	2.7	19.5	85.9	85.8	1134.7	2559.3	343.1
INDIANA	7.4	26.5	123.2	153.5	1086.2	2498.7	377.4
GEORGIA	11.7	31.1	140.5	256.5	1351.1	2170.2	297.9
LOUISIANA	15.5	30.9	142.9	335.5	1165.5	2469.9	337.7
KANSAS	6.6	22	100.7	180.5	1270.4	2739.3	244.3
MISSOURI	9.6	28.3	189	233.5	1318.3	2424.2	378.4
OHIO	7.8	27.3	190.5	181.1	1216	2696.8	400.4
UTAH	3.5	20.3	68.8	147.3	1171.6	3004.6	334.5
CONNECTICUT	4.2	16.8	129.5	131.8	1346	2620.7	593.2
ILLINOIS	9.9	21.8	211.3	209	1085	2828.5	528.6
SOUTH CAROLINA	11.9	33	105.9	485.3	1613.6	2342.4	245.1
NEW JERSEY	5.6	21	180.4	185.1	1435.8	2774.5	511.5
NEW MEXICO	8.8	39.1	109.6	343.4	1418.7	3008.6	259.5
MASSACHUSETTS	3.1	20.8	169.1	231.6	1532.2	2311.3	1140.1
RHODE ISLAND	3.6	10.5	86.5	201	1489.5	2844.1	791.4
TEXAS	13.3	33.8	152.4	208.2	1603.1	2988.7	397.6
MARYLAND	8	34.8	292.1	358.9	1400	3177.7	428.5
WASHINGTON	4.3	39.6	106.2	224.8	1605.6	3386.9	360.3
ALASKA	10.8	51.6	96.8	284	1331.7	3369.8	753.3
MICHIGAN	9.3	38.9	261.9	274.6	1522.7	3159	545.5
OREGON	4.9	39.9	124.1	286.9	1636.4	3506.1	388.9
NEW YORK	10.7	29.4	472.6	319.1	1728	2782	745.8
DELAWARE	6	24.9	157	194.2	1682.6	3678.4	467
HAWAII	7.2	25.5	128	64.1	1911.5	3920.4	489.4
FLORIDA	10.2	39.6	187.9	449.1	1859.9	3840.5	351.4
COLORADO	6.3	42	170.7	292.9	1935.2	3903.2	477.1
CALIFORNIA	11.5	49.4	287	358	2139.4	3499.8	663.5
ARIZONA	9.5	34.2	138.2	312.3	2346.1	4467.4	439.5
NEVADA	15.8	49.1	323.1	355	2453.1	4212.6	559.2

Table H.1: 미국 50개 주의 범죄 자료

Table (H.1)는 초기 자료를 나타내고, Table (H.2)는 평균 0, 분산 1이 되도록 조정한 자료을 나타낸다.

State	murder	rape	robbery	assault	burglary	larceny	auto

NORTH DAKOTA	-1.692	-1.555	-1.254	-1.671	-1.956	-1.141	-1.204
SOUTH DAKOTA	-1.408	-1.137	-1.202	-0.555	-1.668	-1.332	-1.189
WEST VIRGINIA	-0.373	-1.165	-0.927	-1.201	-1.606	-1.832	-1.108
IOWA	-1.33	-1.407	-0.938	-1.212	-1.109	0.019	-0.815
WISCONSIN	-1.201	-1.193	-0.814	-1.472	-1.029	-0.079	-0.811
NEW HAMPSHIRE	-1.098	-1.397	-1.142	-1.35	-0.579	-0.451	-0.435
NEBRASKA	-0.917	-0.71	-0.672	-0.984	-1.23	-0.489	-0.664
VERMONT	-1.563	-0.914	-1.056	-1.098	0.13	-0.648	-0.581
MAINE	-1.304	-1.137	-0.967	-0.412	-0.09	-0.442	-0.675
KENTUCKY	0.687	-0.617	-0.487	-0.878	-0.971	-1.39	-0.683
PENNSYLVANIA	-0.477	-0.626	0.07	-0.831	-0.958	-1.443	-0.229
MONTANA	-0.529	-0.84	-0.961	-0.544	-1.126	0.14	-0.353
MINNESOTA	-1.227	-0.579	-0.432	-1.252	-0.364	-0.154	-0.178
MISSISSIPPI	1.773	-0.57	-0.661	-0.221	-0.87	-1.972	-1.205
IDAHO	-0.503	-0.589	-0.956	-0.387	-0.558	-0.099	-0.724
WYOMING	-0.529	-0.356	-0.955	-0.373	-1.111	0.139	-0.494
ARKANSAS	0.351	0.173	-0.463	-0.079	-0.738	-1.115	-1.004
UTAH	-1.02	-0.505	-0.626	-0.638	-0.278	0.459	-0.222
VIRGINIA	0.402	-0.226	-0.362	-0.455	-0.707	-0.207	-0.78
NORTH CAROLINA	0.816	-0.812	-0.711	1.067	-0.319	-0.873	-0.959
KANSAS	-0.218	-0.347	-0.265	-0.307	-0.05	0.094	-0.689
CONNECTICUT	-0.839	-0.83	0.061	-0.793	0.125	-0.07	1.115
INDIANA	-0.011	0.071	-0.01	-0.577	-0.476	-0.238	-0.001
OKLAHOMA	0.299	0.322	-0.569	-0.063	-0.009	-0.611	-0.262
RHODE ISLAND	-0.994	-1.416	-0.425	-0.103	0.457	0.238	2.14
TENNESSEE	0.687	0.369	0.246	-0.074	-0.074	-1.233	-0.328
ALABAMA	1.747	-0.05	-0.309	0.668	-0.362	-1.087	-0.501
NEW JERSEY	-0.477	-0.44	0.637	-0.261	0.333	0.142	0.693
OHIO	0.092	0.146	0.752	-0.301	-0.176	0.035	0.118
GEORGIA	1.101	0.499	0.186	0.451	0.137	-0.69	-0.412
ILLINOIS	0.635	-0.366	0.987	-0.023	-0.478	0.217	0.781
MISSOURI	0.558	0.238	0.735	0.221	0.061	-0.34	0.005
HAWAII	-0.063	-0.022	0.044	-1.468	1.433	1.721	0.578
WASHINGTON	-0.813	1.289	-0.203	0.135	0.725	0.986	-0.089
DELAWARE	-0.373	-0.078	0.372	-0.171	0.903	1.387	0.463
MASSACHUSETTS	-1.123	-0.459	0.509	0.202	0.556	-0.496	3.943
LOUISIANA	2.083	0.48	0.213	1.239	-0.292	-0.277	-0.206
NEW MEXICO	0.351	1.242	-0.164	1.318	0.293	0.465	-0.61
TEXAS	1.514	0.75	0.32	-0.031	0.72	0.437	0.104
OREGON	-0.658	1.317	0	0.754	0.797	1.15	0.059
SOUTH CAROLINA	1.152	0.675	-0.206	2.733	0.744	-0.453	-0.685
MARYLAND	0.144	0.843	1.902	1.472	0.25	0.698	0.264
MICHIGAN	0.48	1.224	1.56	0.631	0.534	0.672	0.869
ALASKA	0.868	2.404	-0.309	0.725	0.092	0.962	1.943
COLORADO	-0.296	1.512	0.528	0.814	1.488	1.697	0.515
ARIZONA	0.532	0.787	0.16	1.007	2.438	2.474	0.32
FLORIDA	0.713	1.289	0.722	2.372	1.313	1.611	-0.135

NEW YORK	0.842	0.341	3.945	1.075	1.008	0.153	1.904
CALIFORNIA	1.049	2.2	1.844	1.463	1.96	1.141	1.479
NEVADA	2.161	2.172	2.253	1.433	2.685	2.123	0.939

Table H.2: 표준화된 미국 50개 주의 범죄 자료

공분산행렬	murder	rape	robbery	assault	burglary	larceny	auto
murder(X_1)	14.652						
rape(X_2)	24.513	113.454					
robbery(X_3)	161.940	551.386	7649.359				
assault(X_4)	246.385	782.537	4835.477	9849.660			
burglary(X_5)	632.262	3247.314	23860.063	26466.077	183277.582		
larceny(X_6)	280.359	4699.649	28077.753	28838.816	243691.995	516404.581	
auto(X_7)	50.431	711.492	9890.572	5241.179	45730.870	61109.810	36653.372

Table H.3: 공분산 행렬

상관계수 행렬	murder	rape	robbery	assault	burglary	larceny	auto
murder	1						
rape	0.601	1					
robbery	0.483	0.591	1				
assault	0.648	0.740	0.557	1			
burglary	0.385	0.712	0.637	0.622	1		
larceny	0.101	0.613	0.446	0.404	0.792	1	
auto	0.068	0.348	0.590	0.275	0.557	0.444	1

Table H.4: 상관계수 행렬

eigen vector	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7
murder	0.300	-0.629	0.178	-0.232	0.538	0.259	0.267
rape	0.431	-0.169	-0.244	0.062	0.188	-0.773	-0.296
robbery	0.396	0.0422	0.495	-0.557	-0.519	-0.114	-0.003
assault	0.396	-0.343	-0.069	0.629	-0.506	0.172	0.191
burglary	0.440	0.203	-0.209	-0.057	0.101	0.535	-0.648
larceny	0.357	0.402	-0.539231	-0.23489	0.030	0.039	0.601
auto	0.295	0.502	0.568384	0.419238	0.369	-0.057	0.147

Table H.5: Principal components(상관계수)

7개의 범죄요소들의 1차결합으로 새로운 요소 7개를 구성한다.

$$
\begin{aligned}
Y_1 &= 0.300X_1 + 0.431X_2 + 0.396X_3 + 0.396X_4 + 0.440X_5 + 0.357X_6 + 0.295X_7, \\
Y_2 &= \cdots \\
&\vdots
\end{aligned}
$$

eigenvalue	proportion	cumulative
4.11495951	0.5879	0.5879
1.23872183	0.177	0.7648
0.72581663	0.1037	0.8685
0.31643205	0.0452	0.9137
0.25797446	0.0369	0.9506
0.22203947	0.0317	0.9823
0.12405606	0.0177	1

Table H.6: Eigenvalues of the Correlation Matrix

새롭게 구성된 요소중 가장 큰 eigenvalue에 해당하는 Y_1은 다음과 같이 분해될 수 있다.

$$
\begin{aligned}
Y_1 &= 0.300X_1 + 0.431X_2 + 0.396X_3 + 0.396X_4 + 0.440X_5 + 0.357X_6 + 0.295X_7 \\
&= 0.300 \begin{bmatrix} -1.69 \\ -1.408 \\ \vdots \end{bmatrix} + 0.431 \begin{bmatrix} -1.555 \\ -1.137 \\ \vdots \end{bmatrix} + 0.396 \begin{bmatrix} -1.254 \\ -1.201 \\ \vdots \end{bmatrix} + 0.396 \begin{bmatrix} -1.670 \\ -0.554 \\ \vdots \end{bmatrix} \\
&\quad +0.440 \begin{bmatrix} -1.955 \\ -1.668 \\ \vdots \end{bmatrix} + 0.357 \begin{bmatrix} -1.141 \\ -1.331 \\ \vdots \end{bmatrix} + 0.295 \begin{bmatrix} -1.203 \\ -1.189 \\ \vdots \end{bmatrix}
\end{aligned}
$$

□

Appendix I

C++ Program Codes

I.1 Solving Tridiagonal Matrix

```
void tridiagonal_solver(double a,double b, double c,double* y,double* x, int N,
                                 double a0,double c0,double bn_2,double an_1)
{
    //  a0      c0      0       0       0
    //  b       a       c       0       0
    //  0       b       a       c       0
    //  0       0       b       a       c
    //  0       0       0       bn_2    an_1

    double *d=new double[N];
    double *z=new double[N];


    // u
    d[0] = a0;  // a0
    d[1] = a-c0*b/d[0];  // c0
    for(int i=2;i<N-1;i++){
        d[i] = a-c*b/d[i-1];
    }

    d[N-1] = an_1-c*bn_2/d[N-2]; // an_1, bn_1 이기 때문에

    // solve: Lz=y
    z[0] = y[0];
    for(i=1;i<N-1;i++){
        z[i] = y[i] - (b/d[i-1])*z[i-1];
    }
    z[N-1] = y[N-1] -(bn_2/d[N-2])*z[N-2]; //b[N-2] 적용.
```

```
    // solve Ux = z
    x[N-1] = z[N-1]/d[N-1];
    for(i=N-2;i>0;i--){
        x[i] = (z[i] - c*x[i+1])/d[i];
    }
    x[0] = (z[0]-c0*x[1])/d[0];  // c[0] 적용.


    delete[] d;
    delete[] z;
}
```

I.2 Explicit FDM

1. Boundary condition:

$$\frac{\partial^2 f}{\partial S^2} = 0.$$

```
double Explicit_FDM(double S_min, double S_max,double Expiry,double Interest,
                double CostOfCarry,double Volatility,int Nx,double* price,
                double (*payoff)(double),double current,int flag)
{

    double halfVolSquare=0.5*Volatility*Volatility;
    double dS = (S_max-S_min)/Nx;

    // 수렴 조건 check!!!
    assert(dS <= Volatility*Volatility*S_min/(Interest-CostOfCarry) );


    int NearestGridPoint = static_cast<int>( (current-S_min)/dS);

    double dummy = (current - NearestGridPoint*dS-S_min)/dS;
    double dt = dS*dS/(Volatility*Volatility*S_max*S_max);

    int Nt = static_cast<int>(Expiry/dt) + 1;
    dt = Expiry/Nt;

    double* VOld = new double[Nx+1];
    double* VNew = price;  // price의 메모리를 활용한다.
    double* Delta = new double[Nx+1];
    double* Gamma = new double[Nx+1];
    double* S = new double[Nx+1];
    double* Ssquare = new double[Nx+1];

    for(int i=0;i<=Nx;i++){
        S[i] = S_min + i*dS;
        Ssquare[i] = S[i]*S[i];
        VOld[i] = payoff(S[i]);
    }


    for(i=1;i<=Nt;i++){
        for(int j=1;j<Nx;j++){
            Delta[j] = ( VOld[j+1] - VOld[j-1] ) / (2*dS);
            Gamma[j] = ( VOld[j+1] - 2*VOld[j] + VOld[j-1] ) / (dS*dS);
            VNew[j] = VOld[j] + dt * ( halfVolSquare*Ssquare[j]*Gamma[j]
                    + (Interest-CostOfCarry)*S[j]*Delta[j] - Interest*VOld[j] );
        }
        VNew[0] = MAX(0, 2*VNew[1] - VNew[2]);
```

```
        VNew[Nx] = MAX(0, 2*VNew[Nx-1] - VNew[Nx-2]);


        double* temp=VNew;
        VNew = VOld;
        VOld=temp;

    }// end of time loop

    if(price != VOld){
        for(i=0;i<=Nx;i++){
            price[i] = VOld[i];
        }
    }


    double return_value;
    switch(flag)
    {
        case 0:
            return_value = (1-dummy)*price[NearestGridPoint]
                         + dummy*price[NearestGridPoint+1];
            break;
        case 1:
            return_value =  (1-dummy)*(price[NearestGridPoint+1]
                                  -price[NearestGridPoint-1])/(2*dS)
                    + dummy*(price[NearestGridPoint+2]
                             -price[NearestGridPoint])/(2*dS);
            break;
        case 2:
            return_value =  (1-dummy)*(price[NearestGridPoint+1]
              -2*price[NearestGridPoint]+price[NearestGridPoint-1])/(dS*dS)
                    + dummy*(price[NearestGridPoint+2]
              -2*price[NearestGridPoint+1]+price[NearestGridPoint])/(dS*dS);
            break;
        case 3:
            return_value =  (1-dummy) * ( VOld[NearestGridPoint]
                                   - VNew[NearestGridPoint])/dt
                    + dummy*(VOld[NearestGridPoint+1]
                               - VNew[NearestGridPoint+1])/dt;
    }


    delete[] Delta;
    delete[] Gamma;
    delete[] S;
    delete[] Ssquare;
    if(price==VOld){
        delete[] VNew;
    }
```

```
    else delete[] VOld;

    return return_value;
}
```

I.3 Projected SOR

1. The obstacle problem is given by

$$Q\mathbf{x} \geq \mathbf{y}, \quad \mathbf{x} \geq \mathbf{g}, \quad (Q\mathbf{x}-\mathbf{y})(\mathbf{x}-\mathbf{g}) = \mathbf{0},$$

   where

$$Q = \begin{bmatrix} a & c & 0 & 0 & \cdots & 0 & 0 \\ b & a & c & 0 & \cdots & 0 & 0 \\ 0 & b & a & c & \cdots & 0 & 0 \\ 0 & 0 & b & a & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & a & c \\ 0 & 0 & 0 & 0 & \cdots & b & a \end{bmatrix}$$

2. For $\mathbf{x}[0..N-1], \mathbf{y}[0..N-1]$,

$$b\,\mathbf{x}_{i-1} + a\mathbf{x}_i + c\mathbf{x}_{i+1} = \mathbf{y}_i \quad \Longrightarrow \quad \mathbf{x}_i = \frac{\mathbf{y}_i - b\mathbf{x}_{i-1} - c\mathbf{x}_{i+1}}{a}, i = 1, \cdots N-2.$$

3. Since $\mathbf{x}_0$ and $\mathbf{x}_{N-1}$ is obtained from $\mathbf{g}_0$ and $\mathbf{g}_{N-1}$ respectively, we don't use $\mathbf{y}_0$ and $\mathbf{y}_{N-1}$ in computation.

4. We use the Gauss-Seidel method.

```
void PSOR_tridiagonal_solve(double a,double b,double c,double* y,
             double* g,double* x,int N,double omega,double error)
{
    // Projected SOR
    // Q=tridiagonal(b,a,c),  Qx>=y, x>=g, (Qx-y)(x-g)=0.
    // omega: relaxation parameter.

    double err;
    error *= error;

    double temp;

    do{
        err=0.0;
        x[0] = g[0];
        x[N-1] = g[N-1];  // 양 경계값은 boundary 값으로 고정


        for(int i=1;i<N-1;i++){
            temp = (y[i]-b*x[i-1]-c*x[i+1])/a;

            // American style 이므로 체크해야 함.
            temp = MAX(g[i],  x[i] + omega*(temp-x[i]) );
```

```
            err += (x[i]-temp)*(x[i]-temp);
            x[i] = temp;
        }
    }while(  err > error);
}
```

I.4 Callable Equity Swap

```
CGreeks CallableEquitySwap(double S_min, double S_max,CDate t0, CDate t1,double Interest,
                double CostOfCarry,double Volatility,int Nx,int Nt,double* DerivativeValue,
                double (*payoff)(double),CProvision& provision,double current_price,double theta)
{

    if(theta <0 || theta > 1) return CGreeks();


    double X_min = log(S_min);
    double X_max = log(S_max);
    double dx = (X_max-X_min)/Nx;
    double Time_to_Maturity = (t1-t0)/365.0;
    double dt=1.0/(Nt*365);



    // dt에 대한 수렴 조건
    if(theta < 0.5){
        // dx에 대한 수렴 조건-theta와 무관.
        dx = Volatility*Volatility/fabs(Interest-CostOfCarry-Volatility*Volatility/2);
        int NN = static_cast<int>((X_max - X_min)/dx) + 1;
        if(Nx > NN) dx = (X_max - X_min)/Nx;
        else{
            dx = (X_max - X_min)/NN;
            Nx=NN;
        }

        // dt에 대한 수렴 조건.
        dt = 1.0/( (1-2*theta)*(Volatility*Volatility/(dx*dx) +Interest) );
        NN =  static_cast<int>(1.0/dt/365) +1;
        if(Nt >= NN) dt = 1.0/(Nt*365);
        else{
            dt = 1.0/(NN*365);
            Nt = NN;
        }
    }


    double A = dt/(2*dx*dx)*Volatility*Volatility
             - dt/(2*dx)*(Interest-CostOfCarry-Volatility*Volatility/2);
    double B =  -2*dt/(2*dx*dx)*Volatility*Volatility - Interest*dt;
    double C = dt/(2*dx*dx)*Volatility*Volatility
              + dt/(2*dx)*(Interest-CostOfCarry-Volatility*Volatility/2);


    double* VNew = new double[Nx+1];
    double* VOld = new double[Nx+1];


    for(int i=0;i<=Nx;i++){
        VNew[i] = payoff( exp( X_min+ i*dx ));
    }


    for(;t0<=t1;t1--){
```

```
////////////////////////////////////////////////////////////////////////
// Step 1.
// 각각의 하루에 대한 loop
for(int i=1;i<=Nt;i++){

    // Step 1-1.
    // Theta를 구하기 위해 loop의 마지막 단계에서의 값을 저장해 놓는다.
    // Theta를 구하지 않는다면, 이 단계은 필요없다.
    if(t0==t1 && i==Nt){
        for(int j=0;j<=Nx;j++){
            DerivativeValue[j] = VNew[j];  // DerivativeValue에 임시 보관
        }
    }


    ////////////////////////////////////////////////////////////////////////
    // step 1-2.
    for(int j=1;j<Nx;j++){
        VOld[j] =(1-theta)*A*VNew[j-1]+( 1+(1-theta)*B )*VNew[j]+(1-theta)*C*VNew[j+1];
    }


    ////////////////////////////////////////////////////////////////////////
    // Step 2. callable date인지에 따라.
    if( provision.Is_Callable_Date(t1) ){
        double obstacle;
        bool flag= provision.Get_CallablePrice(t1,obstacle);
        assert(flag);

        PSOR_solve_special(1-theta*B,-theta*A,-theta*C,VOld+1,VNew+1,Nx-1,obstacle,
        1-theta*B-2*theta*A,-theta*C+theta*A,-theta*A+theta*C,1-theta*B-2*theta*C);
    }
    else{
        tridiagonal_solve_special(1-theta*B,-theta*A,-theta*C, VOld+1,VNew+1, Nx-1,
        1-theta*B-2*theta*A,-theta*C+theta*A,-theta*A+theta*C,1-theta*B-2*theta*C);

    }


    ////////////////////////////////////////////////////////////////////////
    // Step 1-3. boundary value 계산.
    VNew[0] = MAX(0, 2*VNew[1] - VNew[2]);
    VNew[Nx] = MAX(0, 2*VNew[Nx-1] - VNew[Nx-2]);


}// end of each day loop( i )


// Step 2. coupon과 같은 cash flow 반영하기.
// t1+ 값을 t1- 로 바꾸기(cash flow 적용)
if(provision.Is_Coupon_Date(t1)){
    double coupon = provision.Get_Coupon(t1);
    for(int j=0;j<=Nx;j++){
        VNew[j] += coupon;
    }
}
```

```
    }// end of loop(all days)



        // x ( = log S) 변수를 이용하여 Greek을 구한다.

    for(i=0;i<=Nx;i++){
        VOld[i] = DerivativeValue[i];
        DerivativeValue[i] = VNew[i];
    }


    int NearestGridPoint = static_cast<int>( (log(current_price)-X_min)/dx);
    double dummy = (log(current_price) - NearestGridPoint*dx-X_min)/dx;
    CGreeks greeks;

    // price
    greeks.Value = (1-dummy)*DerivativeValue[NearestGridPoint]
                   + dummy*DerivativeValue[NearestGridPoint+1];

    // delta: Chain rule 이용.
    greeks.Delta
     =( (1-dummy)*(DerivativeValue[NearestGridPoint+1]-DerivativeValue[NearestGridPoint-1])/(2*dx)
     + dummy*(DerivativeValue[NearestGridPoint+2]-DerivativeValue[NearestGridPoint])/(2*dx) ) / current_price;
    // gamma: Chain rule 이용.
    greeks.Gamma
    =(  ( (1-dummy)*(DerivativeValue[NearestGridPoint+1]-2*DerivativeValue[NearestGridPoint]
                     +DerivativeValue[NearestGridPoint-1])/(dx*dx)
           + dummy*(DerivativeValue[NearestGridPoint+2]-2*DerivativeValue[NearestGridPoint+1]
                     +DerivativeValue[NearestGridPoint])/(dx*dx))
     -greeks.Delta*current_price)/(current_price*current_price)    ;
    // theta
    greeks.Theta =
    (1-dummy) * ( VOld[NearestGridPoint] - DerivativeValue[NearestGridPoint])/dt
     + dummy*(VOld[NearestGridPoint+1] - DerivativeValue[NearestGridPoint+1])/dt;

    LogTran_Interpolation(S_min,S_max,VNew,DerivativeValue,Nx);


    delete[] VNew;
    delete[] VOld;

    return greeks;

}
```

Appendix J

Normal Distribution Table

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

Appendix K

Notations

general mathematical symbols:

B_t	Bank account process
$\chi^2(r)$	chi-square distribution with degrees of freedom r
E_Q	the expectation w.r.t. the measure Q
$\mathscr{F}_t$	the σ-algebra
$\mathbb{N}$	the set of natural number
$N(\mu, \sigma^2)$	normal distribution with mean μ and variance σ^2
$\Pr(A), P(A)$	the probability that event A occurs
$N(x)$	$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}} dx$
r	riskless interest rate
q	continuous dividend rate
$\mathbb{R}$	the set of real number
W_t	Brownian motion
$\mathbb{Z}$	the set of integer number
$:=, =:$	equal to by definition
$(A)^+$	$\max(A, 0)$
$\mathbf{1}_{\{expr\}}$	1 if *expr* is true, otherwise 0
$P \ll Q$	the measure P is absolutely continuous w.r.t. the measure Q.

abbreviations:

ADI	Alternating Direction Implicit
ATS	Affine Term Structure
BGM	Brace, Gatarek, and Musiela
CDF	Cumulative Distribution Function
CIR	Cox, Ingersoll, and Ross
ELS	Equity Linked Security
EWMA	Exponentially Weighted Moving Average
FDM	Finite Difference Method
FEM	Finite Elements Method

GARCH	Generalized AutoRegressive Conditional Heteroscedacity
GBM	Geometric Brownian Motion
iff	if and only if
HJM	Heath, Jarrow, and Morton
LIBOR	London InterBank Offer Rate
MC	Monte Carlo
PCA	Principal Component Analysis
PDE	Partial Differential Equation
PDF, pdf	Probability Density Function
PSOR	Projected Successive Over Relaxation
QMC	Quasi Monte Carlo
SDE	Stochastic Differential Equation
SOR	Successive Over Relaxation
w.r.t.	with respect to
YTM	Yield To Maturity

Bibliography

[1] Arnold, L. (1974). *Stochastic Differential Equations: Theory and Applications*. John Wiley & Sons.

[2] Babsiri, M. E. and G. Noel (1998). Simulating path-dependent options: A new approach. *Journal of Derivatives 6*(no. 2), 65–83.

[3] Baxter, M. and A. Rennie (1996). *Financial Calculus; An Introduction to Derivative Pricing*. Cambridge University Press.

[4] Björk, T. (1998). *Arbitrage Theory in Continuous Time*. Oxford University Press.

[5] Black, F. and M. Scholes (1973, May-June). The pricing of options and corporate liabilities. *Journal of Political Economy 81*, 637–59.

[6] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*, 307–27.

[7] Brace, A., D. Gatarek, and M. Musiela (1997). The market model of interest dynamics. *Mathematical Finance 7*, 127–155.

[8] Brennan, M. and E. Schwartz (1978). Finite difference method and jump processes arising in the pricing of contingent claims. *Journal of Financial and Quantitative Analysis 13(September)*, 461–474.

[9] Broadie, M. and P. Glasserman (1996). Estimating security price derivatives using simulation. *Management Science 42*, 269–285.

[10] Brockhaus, O., A. Ferraris, C. Gallus, D. Long, R. Martin, and M. Overhaus (1999). *Modelling And Hedging Equity Derivatives*. Risk Publications.

[11] Burden, R. and J. Faires (1997). *Numerical Analysis* (6th ed.). Brooks/Cole Publishing Company.

[12] Clewlow, L. and C. Strickland (1998). *Implementing Derivatives Models*. John Wiley & Sons.

[13] Cox, J., S. Ross, and M. Rubinstein (1979). Option pricing: a simplified approach. *Journal of Financial Economics vol. 7(October)*, 229–264.

[14] Craig, I. and A. Sneyd (1988). An alternating direction implicit scheme for parabolic equations with mixed derivatives. *Computers & Mathematics with Applications vol. 16*(no. 4), 341–350.

[15] Derman, E., D. Ergener, and I. Kani (1995, Summer). Static options replication. *Journal of Derivatives 2*(s), 78–95.

[16] Derman, E., D. Ergener, I. Kani, and I. Bardhan (1995, May). Enhanced numerical methods for options with barriers. *Goldman Sachs: Quantitative Strategies Research Notes*.

[17] Duffy, D. J. (2004). *Financial Instrument Pricing Using C++*. John Wiley & Sons,Ltd.

[18] Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of united kindom inflation. *Econometrica 50*, 987–1007.

[19] Fourniè, E., J. Lasry, J. Lebuchoux, P. Lions, and N. Touzi (1999). Applications of malliavin calculus to monte carlo methods in fiance. *Finance and Stochastics 3*, 391–412.

[20] Friz, P. K. (2003). Malliavin calculus in finance. *Courant Institute, New York University*.

[21] Geske, R. and K. Shastri (1985). Valuation by approximation: A comparison of option valuation techniques. *Journal of Financial and Quantitative Analysis 20*, 45–71.

[22] Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering; Sotochastic Modelling and Applied Prabability*. Springer.

[23] Glasserman, P. and X. Zhao (2000). Arbitrage-free discretization of lognormal forward libor and swap rate models. *Finance and Stochastics 4*, 35–68.

[24] Glynn, P. (1989). Optimization of stochastic systems via simulation. *In Proceedings of the 1989 Winter Simulation Conference. San Diego, Society for Computer Simulation*, 90–105.

[25] Gobet, E. and A. Higa (2003). Computation of greeks for barrier and lookback options using malliavin calculus. *Electronic Computations in Probability 8*, 51–62.

[26] Heston, S. (1993). A closed-form solution for options with stochastic volatility with application to bond and currency options. *Review of Financial Studies 6*, 327–343.

[27] Hogg, R. V. and A. T. Craig (1995). *Introduction To Mathematical Statistics* (Fifth ed.). Prentice Hall.

[28] Hui, C. (1996). One-touch double barrier binary option values. *Applied Financial Economics 6*, 343–346.

[29] Hull, J. C. (2000). *Options, Futures, & Other Derivatives* (4th ed.). Prentice Hall.

[30] Hull, J. C. and A. White (1990). Valuing derivative securities using the explicit finite difference method. *Journal of Finance and Quantitative Analysis 25(March 1990a)*, 87–100.

[31] Jäckel, P. (2002). *Monte Carlo Methods in Fiance*. John Wiley & Sons, LTD.

[32] Kemna, A. and A. Vorst (1990, March). A pricing method for options based on average asset values. *Journal of Banking and Finance 14*, 113–29.

[33] Lamberton, D. and B. Lapeyre (1996). *Introduction to Stochastic Calculus Applied to Finance*. Chapman & Hall.

[34] Levy, E. (1992). Pricing european average rate currency options. *Journal of International Money and Finance 11*, 474–491.

[35] Lin, X. (1998). Double barrier hitting time distributions with applications to exotic options. *Insurance: Mathematics and Economics 23*, 45–58.

[36] Margrabe, W. (1978, March). The value of an option to exchance one asset for another. *Journal of Finance 33*, 177–86.

[37] Nelson, D. (1990). Arch medels as diffusion approximations. *Journal of Econometrics 45*, 7–38.

[38] **O**ksendal, B. (2003). *Stochastic Differential Equations: An Introduction with Applications* (6th ed.). Springer Verlag.

[39] Press, W., S. Teukolsky, W. Vetterling, and B. Flannery (2002). *Numerical Recipes in C++(The Art of scientific Computing)* (2nd ed.). CAMBRIDGE.

[40] Schwartz (1977). The valuation of warrants: implmenting a new approach. *Journal of Financial Economics 4(January 1977)*, 79–93.

[41] Seydel, R. (2002). *Tools for Compuational Finance*. Springer.

[42] Tavella, D. and C. Randall (2000). *Pricing Financial Instruments: The Finite Difference Method*. John Wiley & Sonns, Inc.

[43] Thomas, J. (1997). *Numerical Partial Differential Equations: Finite Difference Method*. Springer-Verlag.

[44] Topper, J. (2005). *Financial Engineering with Finite Elements*. Wiley.

[45] Turnbull, S. and L. Wakeman (1991, September). A quick algorithm for pricing european average options. *Journal of Financial and Quantitative Analysis 26*, 377–89.

[46] Wilmott, P. (1999). *Derivatives-The Theory and Practice of Financial Engineering*. John Wiley & Sons.

[47] Wilmott, P. (2000). *Paul Wilmott on Quantitative Finance*. John Wiley & Sons.

[48] Wilmott, P., J. Dewynne, and S. Howison (1993). *Option Pricing-Mathematical Models and Computation*. Oxford Financial Press.

[49] Yanenko, N. (1971). *The Method of Fractional Steps*. Springer, Berlin.

Index

A

absorbed process . . . 142
  density of . . . 143
acceptance rejection method . . . 418
ADI . . . 389
  D'Yakonov scheme . . . 390
  Douglas-Rachford Scheme . . . 391
  mixed derivative . . . 392
  Peaceman-Rachford Scheme . . . 389
affine term structure . . . 283
  CIR . . . 290
  Ho-Lee . . . 294
  Hull-White . . . 297
  Vasicek . . . 286
American option . . . 8, 102
  analytic . . . 110
  discrete . . . 100
  early exercise . . . 107
  mathematical formulation . . . 96
  optimal stopping . . . 98
  perpetual . . . 93
antithetic variates . . . 411
Asian option . . . 130
  arithmetic average . . . 138
  geometric average . . . 132
    discrete . . . 136
  pricing equation . . . 131
  similarity reduction . . . 134
asset or nothing option . . . 230

B

Barone-Adesi-Whaley . . . 110
barrier option . . . 178, 187, 190
  asset . . . 204
  binary asset-or-nothing-call . . . 210
  binary asset-or-nothing-put . . . 212
  binary cash-or-nothing-call . . . 206
  binary cash-or-nothing-put . . . 208
  bond . . . 195
  call option . . . 197
  double barrier . . . 214
    distribution . . . 158
    hitting time . . . 148, 153
  put option . . . 201
  up-out-and-down-in call . . . 218
Bermudan option . . . 109
best of two assets . . . *see* exotic option
BGM
  forward measure . . . 342
  spot measure . . . 340
BGM model . . . 331
  calibration . . . 338, 350
  multi-factor . . . 339
  simulation . . . 347
  single-factor . . . 334
binary option . . . 230
binomial tree . . . 10, 15
Black formula . . . 256, 337, 345
Black's model . . . 255
Black-Scholes
  double asset PDE . . . 67, 77
  formula . . . 25, 49, 318
  generalization of formula . . . 32
  model . . . 21
  single asset PDE . . . 65
bond option . . . 320
  coupon bearing bond . . . 326
Box-Muller method . . . 419
Brownian bridge . . . 431
Brownian motion . . . 494
  conditional density . . . 184
  geometric . . . 15
  joint density function . . . 180
  joint distribution . . . 182, 184
  maximum distribution . . . 144

C

callable bond . . . 55
cap . . . 328
cap floor parity . . . 329
caplet . . . 328, 337, 345
Cholesky decomposition . . . 415

chooser option ... 229
CIR model ... 290
    bond option ... 322
CMS ... *see* constant maturity swap
CMT ... *see* constant maturity treasury swap
collar ... 329
complementarity problem ... 103
    American put ... 106
compound option ... 224
constant maturity swap ... 56
constant maturity treasury swap ... 56
convexity ... 248, 249
convexity adjustment ... 258

D

digital option ... 230
    two assets ... *see* exotic option
dividend & GBM ... 19
double barrier
    American binary knock out ... 220
    call ... 214
    knock out binary ... 218
    put ... 216
Dupire equation ... 117
duration ... 247
    modified ... 248
dynamic hedging ... 23, 407

E

exchange option ... *see* exotic option
exotic option ... 224
    asset or nothing ... 230
    binary ... 230
    call on best ... 241
    call on worst ... 243
    chooser ... 229
    cliquet ... 55
    compound ... 55, 224
    corridor ... 55
    digital ... 230
    digital call on best ... 239
    digital call on worst ... 240
    digital put on best ... 240
    digital put on worst ... 241
    exchange option ... 50
    forward start ... 55, 230
    one touch ... 231
    Parisian ... 55
    put on best ... 243
    put on worst ... 244
    relative digital option ... 237
    relative outperformance option ... 238
    shout ... 231
exponential martingale ... 63, 496

F

Faure sequence ... 424
FDM ... 363
    boundary condition ... 382, 401
    Crank-Nicolson ... 368, 378
    discrete cashflow ... 398
    explicit method ... 364, 374
    implicit method ... 377
    jump ... 399
    theta method ... 372, 382
FEM ... 434
    Galerkin method ... 437
    hat function ... 436
Feynman-Kac formula ... 90, 505
finite difference method ... *see* FDM
finite element method ... *see* FEM
first variation process ... 452
floating rate note ... 277
floor ... 328
floorlet ... 329
Fokker-Planck equation ... 116, 506
forward contract ... 260, 276
forward LIBOR ... 334, 339
forward measure ... 315, 336, 342
    affine term structure ... 284
    Hull-White ... 301
forward price ... 261, 279
forward rate ... 295
forward start option ... 230
forward swap ... 278
forward swap rate ... 279
FRN ... *see* floating rate note
fundamental transform ... 127
futures price ... 261
futures rate ... 295

G

Galerkin method ... *see* FEM
GARCH model ... 123, 128
GBM ... 15, 18
    conditional density ... 185
    discrete ... 20
    maximum ... 163, 164
    SDE solution ... 58
generating distribution ... 417

acceptance rejection method........418
Box-Muller method........419
central limit theorem........419
discrete........417
inverse transform........417
normal distribution........419
polar rejection........420
Girsanov........28, 48, 496
Greeks........34, 409
2-dim........81
delta........34
gamma........35
rho........38
theta........36
vega........37

H

Halton sequence........422
heat equation........69
kernel........69
two dimension........77
Heston model........126
hitting time
Brownian motion........179
double barrier........153
lower barrier........146
upper barrier........146
with drift........146
HJM model........304
bond option........320
bond price dynamics........308
forward measure........315
Ho-Lee........310
Hull-White........313
multi-factor........305
short rate........307
single-factor........304
Ho-Lee model........294
bond option........322
Hull-White model
contingent claim on $r(T)$........327
forward measure........301
HJM model........313
short rate........297

I

incomplete market........7
infinitesimal operator........504
instantaneous short rate........269
interest rate tree........357
inverse Gaussian........147, 180
Ito formula........58
Ito isometry........63
Ito lemma........58

K

Kolmogorov........90
backward equation........506
forward equation........506

L

Levy's approximation........141
linear SDE........61
LMM........*see* BGM
lookback option........165
call........169
forward call........172
forward put........175
put........165
low discrepancy........422
LU decomposition........508

M

Malliavin calculus........442
barrier options........483
double asset........453
double asset delta........463
double asset gamma........463
double asset Greeks........445
double asset rho........467
double asset vega........466
European options........460
likelihood ratio method........442, 448
multi-asset........454, 467
multi-asset delta........467
multi-asset gamma........467
multi-asset rho........472
multi-asset vega........470
path dependent options........473
Malliavin Derivatives........451
market price of risk........42, 43, 47
multi assets........44
Markov process........90
martingale measure........21
min-max distribution........146
Monte Carlo method........403
antithetic variates........411
control variates........412
Greeks........409
interest futures........408

moments maching . . . 413
multivariate . . . 414
quasi . . . 422
variance reduction . . . 410
Musiela parametrization . . . 332

N

normal distribution . . . 491
bivariate . . . 491
conditional density . . . 492
multivariate . . . 491
table . . . 526
Novikov condition . . . 496
numeraire . . . 46, 47, 315, 327
multi asset . . . 43, 44

O

obstacle problems . . . 103
one touch option . . . 156, 231
asset . . . 236
deferred payment . . . 234, 236
operator splitting method . . . 394
Black-Scholes PDE . . . 396
optional sampling theorem . . . 150
Ornstein-Uhlenbeck process 60, 280, 288, 292, 323, 357

P

par yield . . . 262
polar rejection . . . 420
portfolio insurance . . . 6
primitive polynomial . . . 426
principal components analysis . . 253, 352, 510
PSOR . . . 108, 521
put call parity . . . 5
for bond option . . . 325
puttable bond . . . 55

Q

quasi Monte Carlo . . *see* Monte Carlo method

R

range forward . . . 55
reflection principle . . . 179
relative digital option . . . *see* exotic option
relative outperformance option . . . *see* exotic option
reverse convertible . . . 56
Riccati equation . . . 283
risk reversal . . . 55

S

SDE . . . 58
self financing . . . 22, 65
series solution . . . 52
double barrier . . . 218, 220
short rate . . . 280
shout option . . . 231
Skorohod integral . . . 450
Snell Envelope . . . 98
Sobol sequence . . . 426
direction number . . . 429
Gray code . . . 428
spot measure . . . 340
spot rate . . . 260
static hedging . . . 54
swap . . . 263, 278
forward . . . 278
swap rate . . . 264, 346
forward . . . 279
swaption . . . 55, 258, 330, 346

T

tree . . . 10, 15
interest rate . . . 357
tridiagonal matrix . . . 507
trinomial tree . . . 13
Turnbull-Wakeman . . . 138

V

variance reduction . . . 410
variational inequality . . . 104
Vasicek model . . . 286
bond option . . . 321
volatility . . . 113
estimation . . . 128
implied . . . 114
surface . . . 116
time dependent . . . 113

W

Wald distribution . . . 180
Wiener-Ito chaos expansion . . . 450
worst of two assets . . . *see* exotic option

Y

yield to maturity . . . 262
YTM . . . *see* yield to maturity

Z

zero rate . . . 260

