

R - Py 컴퓨팅: 기말보고서

Due: 2019-12-23, 블랙보드 제출, 파일 형식은 pdf로, 코드는 부록으로 첨부

Part 1: 시뮬레이션을 통한 옵션 가격 결정

옵션(option)이라는 금융상품에는 여러 종류가 있으나, 그 가운데 가장 대표적인 것으로 콜옵션(call option)과 풋옵션(put option)을 들 수 있다. 우선 콜옵션은 기준자산(underlying asset)을 만기일(maturity date/expiration date)에 행사가격(strike price/exercise price)을 주고 살 수 있는 권리라고 정의할 수 있다. 예를 들어 (기준자산, 만기일, 행사가격)이 (A 주식, 7월 12일, 50,000원)으로 정해진 콜옵션을 매입한 사람은 7월 12일에 A 주식을, 그 날의 실제 주가가 얼마이든지 50,000원에 살 수 있는 권리를 가진 것이다. 예를 들어 7월 12일 A 주식의 가격이 65,000원이 된다면, 이 콜옵션을 보유한 투자자는 65,000원의 가치를 갖는 주식을 50,000원에 살 수 있으므로 65,000원과 50,000원의 차액인 15,000원의 수입으로 얻게 된다. 한편, 7월 12일에 A 주식의 가격이 45,000원이 된다면, 45,000원의 가치를 갖는 주식을 50,000원에 살 수 있는 권리는 아무런 가치도 갖지 않게 된다. 왜냐하면 내가 가진 옵션을 행사해서 50,000에 사는 것보다 주식시장에서 사는 것이 훨씬 더 저렴하기 때문이다.

다시 말해서 콜옵션 매입자는 만기 기준자산의 가격이 행사가격보다 크면 그 차액을 수입으로 얻지만, 행사가격보다 작으면 아무런 수입도 얻지 못하게 된다. 콜 매입자가 만기에 기준자산을 행사가격에 사겠다고 요구하는 것을 콜옵션을 행사한다(exercise)고 표현한다. 콜의 매입자는 만기의 기준자산 가격이 행사가격을 초과할 경우만 그 옵션을 행사할 것이다.

옵션은 하나의 권리가기 때문에, 상응하는 의무를 지게 되는 거래의 상대방이 존재해야 한다. 콜옵션을 소유한 사람이 요구할 경우, 만기일에 기준자산을 행사가격에 매입하는 의무를 지겠다고 약속하는 것을 콜옵션을 발행한다고 표현한다. 콜옵션을 발행하려는 사람이, 그러한 의무를 지겠다는 약속에 대해서, 콜옵션을 매입하려는 사람에게 요구하는 대가를 옵션 프리미엄(option premium)이라고 하며 일종의 옵션의 가격으로 해석이 된다. 옵션을 매입하는 사람은 이러한 가격

을 지급하고 만기에 기준자산의 가격에 따라서 플러스 또는 0의 수입을 얻는 것이다.¹

수식으로 표현하여 행사가격을 K , 만기일의 기준자산 가격을 S_T 라고 표시할 때, 콜의 소유자는 $\max(S_T - K, 0)$ 의 수입을 얻게 된다. $\max(X, Y)$ 란 X 와 Y 의 두 값 가운데 큰 값을 의미한다. 유사하게 $\min(X, Y)$ 는 둘 중 작은 값을 의미한다.

정리하자면, 오늘 콜옵션의 가격을 C 로 표현할 때 콜의 소유자와 발행자 각각의 거래일의 현금흐름과 만기에서의 수입은 다음과 같다. 즉

콜옵션 소유자

$$\text{오늘의 현금흐름}^2 = -C$$

$$\text{만기의 수입} = \max(S_T - K, 0)$$

콜옵션 발행자

$$\text{오늘의 현금흐름} = +C$$

$$\text{만기의 수입} = -\max(S_T - K, 0) \text{ 또는 } \min(K - S_T, 0)$$

다음은 R 코드로 작성하고 실행 결과를 보고하라.

Q1 만기에서의 콜옵션 소유자의 수입 Y 를 기준자산, 여기서는 주식 S_T 의 가격의 그래프로 나타내라. S_T 의 범위는 0부터 200까지이며 $K=100$ 으로 고정한다.

¹ 용어를 정리하면, 옵션의 매입자(option buyer/holder)는 옵션에 대해서 사자(long) 포지션을, 옵션의 발행자(option writer/seller)는 팔자(short) 포지션을 취하였다고 표현한다.

² 오늘 옵션 거래가 일어났다고 생각한다.

Q2. 노벨상을 수상한 블랙-숄즈-머튼은 이와 같은 콜옵션의 프리미엄 C 가 특정가정들을 만족한 경우 다음과 같은 식에 의해서 결정되어야 한다고 증명하였다.

● B-S 옵션가격결정모형

$$C = S \cdot N(d_1) - Ke^{-r_f T} \cdot N(d_2)$$

C : 콜옵션의 가격 S : 주식의 가격

K : 행사가격 r_f : 무위험이자율

T : 옵션의 만기 σ : 주식수익률의 연간 표준편차

$N(d)$: 표준정규분포의 $z=d$ 까지의 누적확률

$$d_1 = \frac{\ln(S/K) + \left(r_f + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}} \quad d_2 = d_1 - \sigma\sqrt{T}$$

여기서 S 는 오늘의 주식 가격을 의미하고 e 는 자연대수를 의미한다. 따라서 오늘의 콜옵션가격은 오늘의 주가(S), 만기(T), 행사가격(K), 주식의 변동성 σ , 무위험 이자율 r_f 에 의해 결정된다. 이 다섯가지 요소를 함수의 인자로 받고, 콜옵션 가격을 계산해주는 BSM 함수를 R로 코딩한 후, $S=100$, $K=100$, $T=0.5$, $\sigma=0.1$, $r_f=0.04$ 에 대해 결과값을 계산하라.

Q3. 재무 이론의 발전은 블랙숄즈머튼의 방법론과 동일한 결과를 다음과 같이 옵션 만기 T 시점 까지 주가의 경로를 N 번 시뮬레이션을 함으로써 얻을 수 있음을 확인하였다. 여기서 i 는 i 번째 주가의 path를 의미하고 t 는 0부터 만기 T 까지의 시간 경로를 의미한다. 주가의 경로를 $0, 0+\Delta t, 0+2\Delta t, \dots, 0+M\Delta t$, $\Delta t=T/M$ 시점에 대해서 각각 산출한다는 의미이다. 이와 같은 주가 S_t^i 에서 $S_{t+\Delta t}^i$ 로의 경로식은 아래와 같이 주어진다.

$$S_{t+\Delta t}^i = S_t^i \exp \left[(r - 0.5\sigma^2)\Delta t + \sigma\sqrt{\Delta t}\epsilon_t^i \right]$$

$$\epsilon_t^i \sim N(0,1)$$

이러한 i 번째 주가 경로에 대해서 만기시점의 옵션의 가치 V_i 는

$$V^i = \max(S_T^i - K, 0)$$

과 같이 정리될수 있다. 최종적으로 옵션의 가치 C는

$$C = \exp(-r_f T) \frac{1}{N} \sum_{i=1}^N V^i$$

로 계산된다. $\Delta t = 1/1000$, $N = 10000$ 번으로 하여 Q2와 동일한 상황에서 옵션 가격을 계산한다. 처음 다섯개의 주가 경로에 대해 오늘부터 만기까지($0 \rightarrow T$) 꺾은선 형태의 그래프로 표기한다.

Part 2: 교수출제 기말과제 (개별과제 대체)

주식시장에 상장된 종목들은 두 가지의 이름을 가진다. 흔히 알고 있는 상장사들의 이름과 '종목 코드'라고 불리는 식별코드가 있다. 예를들어 상장 종목 가운데 대장주인 삼성전자는 '005930'이라는 종목코드를 갖고 있다. 2018년 한국거래소에 따르면 이날 상장기업수는 총 2194사(코스피 777사, 코스닥 1269사, 코넥스 148사)다. 주식시장에 상장된 종목들은 각기 다른 종목코드를 가지고 있다. 그렇다면 상장사의 또다른 이름인 종목코드는 무엇일까?

주식 종목코드는 상장사의 식별코드다. 기업들은 사람으로 치면 주민등록번호를 가지고 있는 셈이다. 코스피 시장의 종목 코드는 총 12자리다. 앞 두자리는 국가코드이며, 증권의 속성(주건, 국채, 지방채 등)을 표시하는 코드번호 1자리, 해당종목의 코드번호 5자리, 증권의 종류(보통주, 우선주, 신주 등)를 나타내는 3자리, 검사숫자 1자리 등으로 구성돼 있다. 예를 들면 삼성전자의 종목코드는 「KR 7 00593 000 3」이다. KR은 대한민국을 의미하며, 7은 증권의 속성을 구분 짓는데 활용되는데, 삼성전자의 경우 증권이기에 때문에 7을 쓴다.

주식시장에서는 사용자의 사용편리를 위해 표준코드 12자리를 6자리로 단축시킨 「단축코드」와 업종별 시장동향, 업종별 주가지수 산정 및 업종별 통계 등에 이용하기 위한 별도의 「업종코드」가 함께 표시된다. 삼성전자의 경우 단축코드는 005930이며 이와 같은 여섯자리 식별 코드를 이용해 주식 거래 정보를 획득한다.

STEP 0. 준비 스텝

파이썬에서 finance-datareader라는 라이브러리를 인스톨하면 한국 증권시장에서 거래되는주 주식 종목 코드를 증권거래소 별로 얻을 수 있다. 한국의 증권거래소는 KOSPI, KOSDAQ, KONEX 세 곳이며 KOSPI와 KONEX가 주된 시장이다. 반드시 아래 링크의 설명을 확인하도록 한다.

<https://financedata.github.io/posts/finance-data-reader-users-guide.html>

Q4) KOSPI에 상장된 종목을 df_KOSPI 코스닥에 상장된 종목을 df_KOSDAQ라는 데이터프레임으로 각각 만든다. StockListing메소드를 사용하면 된다. 위 두 데이터 프레임은 ['Symbol', 'Name', 'Sector', 'Industry']을 칼럼으로 가지고 있으며 Symbol 변수가 위에서 지칭한 여섯자리 코드를 의미한다.

- KOSPI와 KOSDAQ 그룹에 대해서 각각 몇 개의 상장사가 있는지 확인한다.
- KOSPI라는 list로 df_KOSPI의 Symbol값들을 저장하고, KOSDAQ이라는 이름으로 df_KOSPI의 Symbol값들을 저장한다.

STEP1 주가 내려 받기:DataReader 메소드를 사용

다음 명령어는 주식코드 215600 신라젠 주식에 대해서 2018-01-01일부터 2018-03-30까지 정보를 제공한다.

```
df = fdr.DataReader('215600', '2018-01-01', '2018-03-30')
```

```
df.head()
```

	Open	High	Low	Close	Volume	Change
Date						
2018-01-02	95900	104000	93300	102500	6824281	0.096257
2018-01-03	102600	104900	99500	103000	4760150	0.004878
2018-01-04	102600	104000	92200	92200	6538665	-0.104854
2018-01-05	85800	101200	85700	100000	8318372	0.084599
2018-01-08	98000	98400	92500	93800	6352776	-0.062000

여기서 open은 신라젠 주식가격의 장시작가, High는 당일 최고가, Low는 당일 최저가, Close는 당일 종가를 의미한다. Change는 일별 수익률을 의미한다.

Q5) 삼성전자에 대한 주식 종목 코드를 df_KOSPI에서 검색하거나 간단히 검색엔진을 사용하여 검색한후 fdr.DataReader 메소드를 사용하여 자료를 받은후에 각 날짜에 대한 일별 주가 수익률을 꺾은선 그래프 형태로 그려본다.

Q6) KOSDAQ과 KOSPI에서 거래되는 전체 기업들에 대해서 각각 일별 수익률의 연간 평균과 표준편차를 구해본다. 전체기업들에 대해 평균수익률-표준편차의 산포도를 그려본다. 이러한 산포도를 KOSPI에 상장된 기업과 KOSDAQ에 상장된 기업에 대해 구분해서 각각 산포도를 그려본다. 이 두 변수간의 양의 관계가 있는가? 평균수익률은 일종의 사람들이 기대하는 수익률과 밀접한 관련이 있으며, 주가수익률의 표준편차가 클수록 일반적으로 수익률의 리스크(위험도)가 크다고 생각한다.

STEP2

자본자산 가격결정 모형(Capital Asset Pricing Model, CAPM)은 자본시장의 균형하에서 위험이 존재하는 자산의 균형수익률을 도출해내는 모형이다. 마코위츠의 포트폴리오 이론을 바탕으로 하여, 샤프 등에 의해 무위험자산의 가정을 포함하여 발전되었다. 넓은 의미로는 자본시장선과 증권시장선을 포함하는 개념이나, 보통 CAPM이라 하면 증권시장선을 의미하는 경우가 많다.

증권시장선(SML; security market line)은 개별자산 또는 포트폴리오의 균형수익률을 도출해내는 모형으로, 체계적 위험의 지표인 베타에 비례하는 위험프리미엄을 측정하여 균형수익률을 이끌어 낸다. SML은 위험프리미엄, 즉 기대수익률과 무위험수익률의 차이의 보상기준이 되는 위험이 총 위험이 아닌 체계적위험이며, 따라서 효율적포트폴리오뿐만이 아닌 개별주식과 비효율적포트폴리오의 균형수익률도 측정가능하다. SML 식은 다음과 같다.

$$(E(R_i) - R_f) = \beta(E(R_M) - R_f)$$

여기서 R_i 는 개별자산의 수익률, R_M 은 KOSPI200과 같은 시장 인덱스의 수익률, R_f 는 무위험 수익률로 정의가 된다. 앞에 E는 기대값을 나타내는 기호이다. β 는 리스크를 의미하며, β 가 클수록 평균 수익률이 크다는 것을 알 수 있다.

이와 같은 모형을 개별 주식에 대해 실제로 분석하기 위해서는 다음과 같은 회귀 분석 모형을 사용한다.

$$R_i - R_f = \alpha + \beta(R_M - R_f) + \varepsilon$$

$$R_i^e = \alpha + \beta R_M^e$$

여기서 R_i^e 는 $R_i - R_f$, R_M^e 는 $R_M - R_f$ 이다. α 와 β 는 우리가 모형을 통해 Fitting 해야 할 계수이다.

Q7) 여러분은 이 회귀분석식을 한국시장에서 거래되는 모든 주식에 각각 α 와 β 를 Fitting 한다.

1. 우리의 target 변수는 개별 주식의 일별수익률(change) - 무위험 수익률(0.03/365)이며
2. feature는 KOSPI200의 일별수익률(change)- 무위험 수익률(0.03/365)이다.
3. 모든 주식에 대해서 리스크의 측도 β 를 각각 구해보고, 평균 수익률과의 산포도를 그려 보아라. 양의 관계를 발견할수 있는가?.
4. 3의 분석을 KOSPI 시장과 KOSDAQ 시장을 분리해서 시행한다.

STEP 3

Q8) Jegadeesh 와 Titman이 1993년에 작성한 논문은 미국 주식시장에서 흥미로운 사실을 발견한다. 즉 과거 수익률이 우수한 종목을 매수하고, 수익률이 부진한 종목을 매도할 경우, 3에서 12개월 정도 보유 기간에서 유의미한 수익률을 기록함을 확인하였다.

이를 한국 시장에 있어서도 잘 적용되는지 간단한 형태로 확인을 해보고자 한다.

1. 전체 주식에 대해서 1월 2일과 6월 29일까지 누적 수익률을 구한다 이 누적수익률은 (6월 29일 종가-1월 2일 종가)/ 1월 2일 종가로 계산한다.
2. 이 전체 수익률을 바탕으로 수익률 상위 - 하위까지 10개 그룹으로 구분한다.
3. 각 기업에 대해 7월 2일부터 9월 28일 (3개월 누적 수익률) 및 7월 2일부터 12월 28일까지의 누적수익률 (6개월 누적 수익률)을 1번과 동일한 방법으로 구한다.
4. 1-6월부터 누적 수익률 10개 기업 그룹별로 3개월 그룹수익률의 평균 및 6개월 누적 수익률의 평균값을 계산하여 막대그래프 형식으로 표시한다.
5. KOSPI와 KOSDAQ으로 상장 기업그룹을 나누어 1-4번까지 작업을 반복한다.

Part 2: 개별 기말과제

각 개인은 교수출제 기말과제 대신 본인이 자유주제로 기말보고서를 작성할 수 있다. 이를 위해서는 다음과 같은 부분을 반드시 포함해야 한다.

1. 연구 목적에 대한 1~2페이지 정도의 기술: 연구 문제에 대한 구체적인 기술이 필요
2. 데이터: 사용할수 있는 데이터는 XML, JSON 형식 혹은 웹크롤링을 통한 데이터 확보: 구체적인 데이터에 대한 1~2페이지 정도의 기술
3. 모형: 분석 모형 설정과 그 해석에 대한 1 ~2페이지 정도 기술
4. 모형 분석의 결과를 제시하고 연구목적의 관점에서 해석: 반드시 하나이상의 그래프를 그리고 그 분석을 포함해야 함: 약 4-5 페이지
5. 프로그래밍 언어: 파이썬/ R 코드로 작성하며 부록으로 반드시 포함