

# 1 Probability and Stochastic Process

## 1.1 Probability

$$\text{Var } X = \mathbb{E}[(X - \mathbb{E}X)^2] = \mathbb{E}[X^2] - (\mathbb{E}X)^2. \quad \text{Var}(aX + b) = a^2 \text{Var } X.$$

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]. \quad \text{Cov}(aX + b, Y) = a \text{Cov}(X, Y).$$

$$\text{Cov Matrix for } X = (X_1, \dots, X_n)^T: \text{Cov}(X) = (\text{Cov}(X_i, X_j))_{i,j}.$$

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var } X \text{Var } Y}}$$

矩母函数 Moment generating function: For rv  $X$ , MGF is  $M_X(t) = \mathbb{E}[e^{tX}]$ . For vec-valued rv  $\mathbf{X}$ ,  $M_X(\mathbf{t}) = \mathbb{E}[e^{i\langle \mathbf{t}, \mathbf{X} \rangle}]$ . 可以生成  $n$  阶矩  $\mathbb{E}[X^n] = \frac{d^n}{dt^n} M_X \Big|_{t=0}$ . 矩母函数相同说明分布相同.

特征函数 Ch.f.: For rv  $X$ ,  $\varphi_X(t) = \mathbb{E}[\exp(itX)]$ . For vec-valued rv  $\mathbf{X}$ ,  $\varphi_X(\mathbf{t}) = \mathbb{E}[e^{i\langle \mathbf{t}, \mathbf{X} \rangle}]$ .

一些分布

- Binomial( $p, n$ ):  $P_X(k) = C_n^k p^k (1-p)^{n-k}$ . 期望  $np$ , Var  $np(1-p)$ .
- Poisson( $\lambda$ ):  $P_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}$ . 期望  $\lambda$ , Var  $\lambda$ .
- Normal( $\mu, \sigma^2$ ):  $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{(x-\mu)^2}{2\sigma^2})$ . 偏度为0, 峰度为3. MGF:  $\exp(\frac{1}{2}u^2t)$ .
- Exponential( $\lambda$ ):  $f_X(x) = \lambda e^{-\lambda x} (x \geq 0)$ . 期望  $\lambda$ , Var  $\lambda^2$ .

### Multivariate Gaussian $\mathcal{N}(\mu, \Sigma)$

定义: Random vec  $X$  是  $d$  维 Gaussian Vec 若对任意  $t \in \mathbb{R}^d$ ,  $\langle t, X \rangle$  都是 Gaussian.

Note that  $n$  个正态分布拼起来不一定是 Gaussian vec, 但  $n$  个独立正态分布拼起来一定是.

给定 vec  $\mu$  和半正定对称阵  $\Sigma$ , 存在 Gaussian Vector  $X$  with mean  $\mu$  and Cov  $\Sigma$ . 如果  $\Sigma$  是满秩的, 则存在 density  $f(\mathbf{x}) = (2\pi)^{-d/2} \det(\Sigma)^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu)\right)$ ,

$$\text{MGF } \exp\left(\mu^T t + \frac{1}{2} t^T \Sigma t\right). \quad \text{ChF } \exp\left(i\mu^T t - \frac{1}{2} t^T \Sigma t\right).$$

$$\text{条件概率: } \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \quad \text{贝叶斯公式: } \mathbb{P}(A|B) = \mathbb{P}(B|A) \frac{\mathbb{P}(A)}{\mathbb{P}(B)}.$$

条件期望:  $\mathbb{E}[X|\mathcal{F}]$  是一个随机变量使得 (1) 是  $\mathcal{F}$  可测的; (2)  $\mathbb{E}[\mathbb{E}[X|\mathcal{F}]1_A] = \mathbb{E}[X1_A], \forall A \in \mathcal{F}$ .

对两个 rv  $X, Y$ , 有 joint density  $f_{X,Y}(x, y)$ , 则 Marginal density  $f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x, v) dv$ . 条

件密度  $f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$ , 条件期望  $\mathbb{E}[X|y] = \int_{\mathbb{R}} x f_{X|Y}(x|y) dx$ .  $\mathbb{E}[g(X)|Y] = h(Y)$  where

$$h(y) = \frac{\int g(x) f_{X,Y}(x, y) dx}{\int f_{X,Y}(x, y) dx}.$$

独立性引理. 若  $X, Y$  独立, 则  $\mathbb{E}[\varphi(X, Y)|Y] = h(Y)$  where  $h(y) = \mathbb{E}[\varphi(X, y)]$ . 更一般地,  $X_1, \dots, X_k$  是  $\mathcal{G}$  可测的,  $Y_1, \dots, Y_l$  独立于  $\mathcal{G}$ , 则  $\mathbb{E}[f(X_1, \dots, X_k, Y_1, \dots, Y_l)|\mathcal{G}] = g(x_1, \dots, x_k)$  where  $g(x_1, \dots, x_k) = \mathbb{E}[f(x_1, \dots, x_k, Y_1, \dots, Y_l)]$ .

一些命题

- $F$ 递增+右连续+[0, 1], 即可定义一个分布函数. 如果存在非负  $f$  使  $F(t) = \int_{-\infty}^t f(x)dx$ , 则分布有 density.
- $X$  有 pdf  $f_X$ ,  $Y = g(X)$  其中  $g$  可微且严格单调, 则  $f_Y(y) = f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$ .
- $X \sim \mathcal{N}(\mu, \sigma^2)$ , 则  $\mathbb{E}[|X|] = \mu \left[ 2\Phi\left(\frac{\mu}{\sigma}\right) - 1 \right] + \frac{2\sigma}{\sqrt{2\pi}} \exp\left\{-\frac{\mu^2}{2\sigma^2}\right\}$ . 其中  $\Phi$  为 std norm CDF.
- 在  $[0, 1]$  的均匀分布下取  $n$  个值, 最大值的期望是  $\frac{n}{n+1}$ .
- Distribution of  $X + Y$  when  $X \sim f, Y \sim g$  are independent:  $\mathbb{P}(X + Y \leq z) = \int \mathbb{P}(x \leq z - y)g(y)f(y)dy$ , density of  $X + Y$  is  $h(z) = \int f(z - y)g(y)dy$ .
- 顺序统计量  $k$ th order statistic: the  $k$ th smallest value in a sample of  $n$  from a random distribution. There is  $\mathbb{P}(X_{(k)} \leq x) = \sum_{j=k}^n F(x)^j (1 - F(x))^{n-j}$ .

大数定律:  $X_1, \dots, X_n$  是 iid rv 序列, 期望  $\mu$  有限. 令  $S_n = X_1 + \dots + X_n$ , 则 (弱)  $\frac{S_n}{n} \rightarrow \mu$  in probability i.e.  $\lim_{n \rightarrow \infty} \mathbb{P}(|\frac{S_n}{n} - \mu| > \varepsilon) = 0$ . (强) Almost surely  $\lim_{n \rightarrow \infty} \frac{S_n}{n} = \mu$ .

中心极限定理 Central Limit Theorem:  $X_1, \dots, X_n$  是 iid rv 序列, 期望  $\mu$  和方差  $\sigma^2$  有限, 令  $S_n = X_1 + \dots + X_n$ , 则  $Y_n = \frac{S_n - \mu n}{\sigma \sqrt{n}}$  converges in distribution to std norm.

## 1.2 Stochastic Process

### 鞅和停时

$(X_n)$  or  $(X_t)$  是鞅,  $T$  是停时, 则  $(X_{n \wedge T})$  or  $(X_{t \wedge T})$  是鞅.

[Optional Stopping Thm]  $(X_n)$  u.i. sub-mtg,  $S \leq T$  是停时, 则  $\mathbb{E}[X_T | \mathcal{F}_S] \geq X_S$ .  $(X_t)$  u.i. mtg,  $S \leq T$  是停时, 则  $\mathbb{E}[X(T) | \mathcal{F}(S)] = X(S)$ .

Rk. 一致可积 u.i. 即  $\forall \varepsilon, \exists K$  s.t.  $\mathbb{E}[|X_n| 1_{|X_n| \geq K}] < \varepsilon (\forall n)$ . 对离散 sub-mtg, 一致可积  $\Leftrightarrow$  convergence a.s. and in  $L^1 \Leftrightarrow$  convergence in  $L^1$ . 对连续 mtg, 一致可积  $\Leftrightarrow$  convergence in  $L^1 \Leftrightarrow \exists Y \in L^1, X(t) = \mathbb{E}[Y | \mathcal{F}(t)]$ .

### Brownian Motion 性质

- Scaling invariance:  $\frac{1}{a}B(a^2t)$  是 BM. 可得 first passage time  $T_a = a^2T_1$  in distr.
- Time inversion:  $X(t) = tB(\frac{1}{t})(t > 0), X(0) = 0$  是 BM. 可得  $\lim_{t \rightarrow \infty} \frac{B(t)}{t} = 0$  a.s.
- Reflection Principle:  $T$  是停时,  $B_t 1_{t \leq T} + (2B_T - B_t) 1_{t > T}$  是 BM.
- $B(t), B(t)^2 - t$  是鞅. 指数鞅  $X(t) = \exp(\lambda B(t) - \frac{1}{2}\lambda^2 t)$  对  $\lambda \in \mathbb{R}$ .
- Arcsine law:  $L = \sup\{t \in [0, 1] : B(t) = 0\}$  is distributed as  $\mathbb{P}(L \leq x) = \frac{2}{\pi} \arcsin \sqrt{x}$ .

- Wald's Lemma: 停时 $T$ 有 $\mathbb{E}T < \infty$ , 则 $\mathbb{E}[B(T)] = 0$ ,  $\mathbb{E}[B(T)^2] = \mathbb{E}T$ .
- Law of the Iterated Logarithm:  $\limsup_{t \rightarrow \infty} \frac{B(t)}{\sqrt{2t \log \log(t)}} = 1$  a.s.

### Stochastic Calculus

二次变差  $[f, f](T) = \lim_{|\Pi| \rightarrow 0} \sum_{j=0}^{n-1} (f(t_{j+1}) - f(t_j))^2$ . 有  $[W, W](t) = t$  a.s.

Ito Isometry:  $\mathbb{E} \left[ \left( \int_0^T X_t dW_t \right) \left( \int_0^T Y_t dW_t \right) \right] = \mathbb{E} \left[ \int_0^T X_t Y_t dt \right]$ .

$\mathbb{E}[W_t^2] = t$ ,  $\mathbb{E}[W_t^4] = 3t^2$ .  $\int_0^t W_s ds \sim \mathcal{N}(0, \frac{1}{3}t^2)$ ,  $\int_0^t f(s) dW_s \sim \mathcal{N} \left( 0, \int_0^t f(s)^2 ds \right)$ .

$\int_0^T W_t dW_t = \frac{1}{2}(W_T^2 - T)$  (apply Ito's lemma to  $W_t^2$ ).

Ito过程:  $X(t)$ 满足 $dX(t) = \Delta(t)dW(t) + \Theta(t)dt$ . 有 $[X, X](t) = \int_0^t \Delta(u)^2 du$ .

Ito Lemma:  $df(W(t)) = f'(W_t)dW_t + \frac{1}{2}f''(W_t)dt$ . 对Ito过程 $X_t$ , 有 $df(t, X_t) = f_t(t, X_t)dt + f_x(t, X_t)dX_t + \frac{1}{2}f_{xx}(t, X_t)dX_t dX_t$ .

Stock Price  $S_t$  following log-normal Brownian motion:  $dS_t = \alpha S_t dt + \sigma S_t dW_t$ ,  $d \log S_t = (\alpha - \frac{1}{2}\sigma^2)dt + \sigma dW_t$ .

Girsanov Thm. 定义 $Z(t) = \exp \left( - \int_0^t \Theta(u) dW(u) - \frac{1}{2} \int_0^t \Theta(u)^2 du \right)$  作为Radon-Nikodym导数得到 $\tilde{\mathbb{P}}(A) = \int_A Z(\omega) d\mathbb{P}(\omega)$ . 则 $\tilde{W}(t) = W(t) + \int_0^t \Theta(u) du$ 在 $\tilde{\mathbb{P}}$ 下是BM.

在 $\Theta(t) = \frac{\alpha(t) - R(t)}{\sigma(t)}$ 定义的 $\tilde{\mathbb{P}}$ ,  $\tilde{W}$ 下,  $dS(t) = R(t)S(t) + \sigma(t)S(t)d\tilde{W}(t)$ 贴现后是鞅.

### 1.3 Stochastic Differential Equation

线性SDE求解 考虑  $\begin{cases} dX(t) = (c(t) + d(t)X(t))dt + (e(t) + f(t)X(t))dW(t) \\ X(0) = X_0 \end{cases}$  其中  $c, d, e, f$

非随机. 结论: 解为  $X_1(t) \cdot \left( X_0 + \int_0^t (c(s) - f(s)e(s)) (X_1(s))^{-1} ds + \int_0^t e(s) (X_1(s))^{-1} dW(s) \right)$ ,

其中  $X_1(t) = \exp \left( \int_0^t f(s) dW(s) - \int_0^t (d(s) - \frac{1}{2}f(s)^2) ds \right)$ .

解法: Duhamel原理.

先得到  $\begin{cases} dX_1(t) = d(t)X_1(t)dt + f(t)X_1(t)dW(t) \\ X_1(0) = 1 \end{cases}$  的解 $X_1(t)$  (对 $d(\ln X_1(t))$ 用Ito即可).

再分离变量, 设  $X(t) = X_1(t) \cdot X_2(t)$  其中 $X_2$ 满足  $\begin{cases} dX_2(t) = A(t)dt + B(t)dW(t) \\ X_2(0) = X_0 \end{cases}$  是原方

程的解,  $A(t), B(t)$ 待定. 写开 $dX(t)$ 代回原方程, 有  $\begin{cases} A(t) = (c(t) - f(t)e(t)) (X_1(t))^{-1} \\ B(t) = e(t) (X_1(t))^{-1} \end{cases}$ , 积分即

得 $X_2$ . 从而有原方程解 $X_1 \cdot X_2$ .

**Feynman Kac Theorem** 考虑方程  $dX(u) = \beta(u, X(u))du + \gamma(u, X(u))dW(u)$ , 记  $g(t, x) = \mathbb{E}^{t,x}[h(X(T))]$ , 其中  $X$  是初值  $X(t) = x$  下的解. 则  $g_t + \beta g_x + \frac{1}{2}\gamma^2 g_{xx} = 0$ .

## 2 OLS

$y = X\beta + \varepsilon$ , G-M假设  $\text{Cov}(e) = \sigma^2 I_n$ .

估计  $\hat{\beta} = (X^T X)^{-1} X^T y$ . 满足  $\mathbb{E}[\hat{\beta}] = \beta$ ,  $\text{Var}(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$ . 是BLUE最佳线性无偏估计.

$X$  已知非随机,  $y$  已知随机,  $\beta$  未知非随机,  $\hat{\beta}$  随机,  $e$  未知随机 ( $\sigma^2$  未知). 每个  $y_i$  都有一个对应的总体  $N(\beta^T x_i, \sigma^2)$

记  $H = X(X^T X)^{-1} X^T$ , 有  $\hat{y} = Hy$ .  $H$  对称幂等,  $HX = X$ , 迹为  $k+1$ .

$\mathbb{E}[\hat{y}] = y$ ,  $\text{Var}(\hat{y}) = \sigma^2 H$ . 残差  $e = y - \hat{y}$  (注意区别于误差向量  $\varepsilon$ ) 有  $\mathbb{E}[e] = 0$ ,  $\text{Var}(e) = \sigma^2 (I - H)$ .

$SSR = \sum \varepsilon_i^2$ ,  $SSE = \sum (\hat{y}_i - \bar{y})^2$ ,  $SST = \sum (y_i - \bar{y})^2$ . 有  $SST = SSE + SSR$ . 定义  $R^2 = 1 - SSR/SST$ .

$\text{Var}(\hat{\beta}_j) = \frac{\sigma^2}{SST_j(1 - R_j^2)}$ , 其中  $SST_j = \sum_i (x_{ij} - \bar{x}_j)^2$ ,  $R_j^2$  为  $x_j$  关于其他自变量含截距回归的  $R^2$ .

MLE:  $\hat{\sigma}^2 = \frac{RSS}{n}$  是有偏的, 无偏估计为  $\hat{\sigma}^2 = \frac{RSS}{n-k-1}$  (或写为  $se^2$ ).

$\hat{\beta}_j$  标准误 (标准化残差)  $se(\hat{\beta}_j) = \frac{\hat{\sigma}}{(SST_j(1 - R_j^2))^{1/2}}$ .  $SE(\hat{\beta}_j)$  是  $\text{Cov}(\hat{\beta})$  对应对角元的开根.

T化残差为原始残差除以估计的标准差.

t检验统计量  $t_{\hat{\beta}_j} = \hat{\beta}_j / se(\hat{\beta}_j)$ . 整体回归F统计量  $F = \frac{R^2/k}{(1 - R^2)/(n - k - 1)}$ .

衡量多重共线性  $VIF_j = \frac{1}{1 - R_j^2}$ , 大于10认为严重. 有  $SE(\beta_j) = \frac{\sigma^2}{(n-1)\widehat{\text{Var}}(X_j)} VIF_j$ .

$y$  对  $x$  做一元回归, 回归系数  $\hat{\beta}_{yx} = \frac{\text{Cov}(x, y)}{\text{Var}(x)}$ . 回归  $R^2$  的分布  $\sim \text{Beta}(\frac{k-1}{2}, \frac{n-k}{2})$ .  $\text{Beta}(\alpha, \beta)$  的期望为  $\frac{1}{1 + \beta/\alpha}$ .

## 3 Options

### 3.1 Pricing

Binomial Model. 风险中性概率  $\tilde{p} = \frac{1+r-d}{u-d}$ ,  $\tilde{q} = \frac{u-1-r}{u-d}$ .  $V_n(\omega) = \frac{1}{1+r} [\tilde{p}V_{n+1}(\omega H) + \tilde{q}V_{n+1}(\omega T)]$ .  $\Delta_n(\omega) = \frac{V_{n+1}(\omega H) - V_{n+1}(\omega T)}{S_{n+1}(\omega H) - S_{n+1}(\omega T)}$ .

Put option value wrt Strike price 应该是凸的.

当total variance  $\sigma^2 T$  较小, ATM put 价值有近似  $P_{ATM} \simeq 0.4\sigma S_0 \sqrt{T}$ . ATM Call 和 Put 都有  $\Delta \simeq 0.5$ .

Put-Call Parity  $C(t) - P(t) = S(t) - Ke^{-r(T-t)}$ .

BS equation: 假设股价符合几何布朗运动  $dS = \mu S dt + \sigma S dW$ , 即  $S_t = S_0 \exp((r - \frac{1}{2}\sigma^2)t + \sigma W(t))$ .  $\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0$  (推导: 对  $e^{-rt}V(t, S(t))$  用Ito).

BSM formula:  $c = S_0 N(d_1) - Ke^{-rT} N(d_2)$ ,  $p = Ke^{-rT} N(-d_2) - S_0 N(-d_1)$ , 其中  $d_1 = \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}$ ,  $d_2 = \frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$ .

### 3.2 Greeks

Delta: 标的资产价格变化引起的期权价格变化. Gamma: 标的资产价格变化引起的Delta值变化. Vega: 隐含波动率变化引起的期权价格变化. Theta: 期权的时间价值随时间流逝损耗的速度.

#### Delta & Gamma

- Call的Delta是正的, Put的Delta是负的. Call和Put的Gamma都是正的.
- 对Call, 随着Vol下降, ITM Delta上升(范围从50到100), OTM Delta下降; 随着到期日临近, ITM Delta上升, OTM Delta下降. 波动率下降/到期日临近下, Delta-标的价格变化图像为Fig1左.
- 对Put, 随着波动率上升, OTM Delta下降(范围从0到-50), ITM Delta上升; 随着到期日临近, ITM Delta下降, OTM Delta上升.
- Gamma在ATM最大. ATM的Gamma与Strike负相关.
- 随波动率下降/到期日临近下, 期权的Gamma值变化图像为Fig1右.

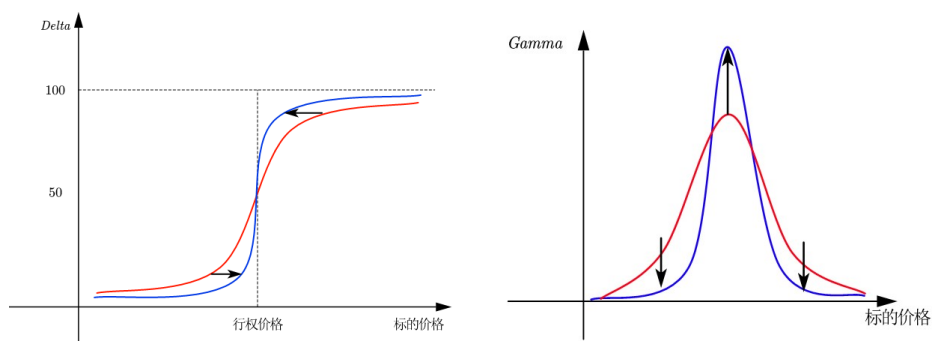


图 1: Fig1

#### Theta

- Theta通常是负的.
- Theta值在平值时绝对值最大, 深度实值或深度虚值时绝对值较小.
- 期权价值与—Theta—在临近到期时的变化见Fig2.

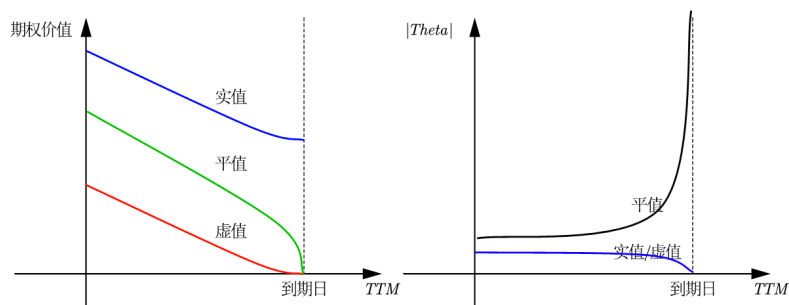


图 2: Fig2

- 平值期权的Theta与波动率成比例变化.

### Vega

- Vega在平值时最大, 深度实值或深度虚值时较小.
- 期权价值与Vega在不同波动率下的情况见Fig3.

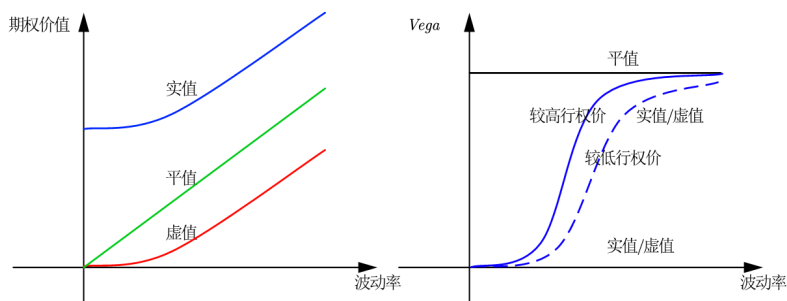


图 3: Fig3

- Vega值随到期日临近的变化见Fig4.

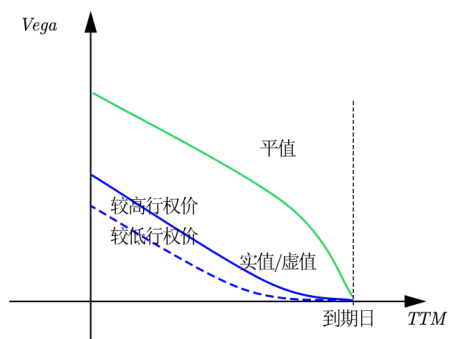


图 4: Fig4

### 3.3 Structure

#### 对称策略

以下默认期权多头. 价指Strike Price.

跨式期权 Straddle: Long 同价的Call+Put.  $\backslash /$  形. Neutral Delta, Long Gamma, Long Vega, Short Theta.

宽跨式期权 Strangle: Long 高价Call + 低价Put.  $\backslash - /$  形. Neutral Delta, Long Gamma, Long Vega, Short Theta.

蝶式期权 Fly: Long 低价Call + 高价Call, Short 2 中间价Call.  $- / \backslash -$  形. Neutral Delta, Short Gamma, Short Vega, Long Theta.

鹰式期权 Iron Condor: Long 低价Call + 高价Call, Short 中低价Call+中高价Call.  $- / - \backslash -$  形. Neutral Delta, Short Gamma, Short Vega, Long Theta.

#### 时间价差

时间价差多头: 买入长期期权, 卖出短期期权.(因为长期期权通常更贵) 使用平值期权1:1构成Delta中性.

时间价差多头是Short Gamma的, 因为若标的价格不变, 随时间流逝, 短期期权价值减少是多于长期期权的.

时间价差多头是Long Vega的, 因为波动率变化(上升)对长期期权有更大影响(增加更多时间价值).

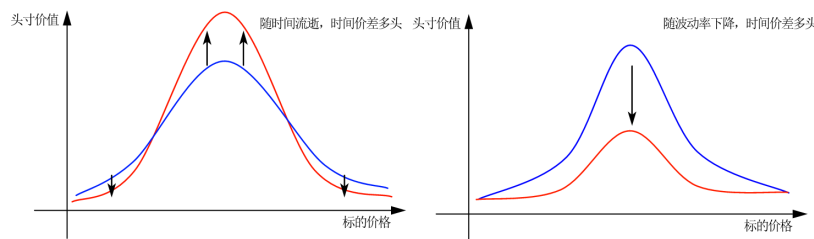


图 5: Fig5

策略选择: 如果隐含波动率低, 要 Long Vega. 当隐含波动率很低但被认为会升高时, 时间价差多头可能获利.

#### 垂直价差

买入与卖出相同类型、同时到期、行权价格不同的期权. 牛市价差 Call Spread: 买低, 卖高. (Call: 牛市看涨, Put: 牛市看跌). 熊市价差 Put Spread: 卖低, 买高. (Call: 熊市看涨, Put: 熊市看跌).

买低卖高, 无论是Call还是Put都具有牛市特征, 即具有正的Delta值. 两个行权价之间差距越大, 牛市特征越明显(价差的Delta越大).

Call Spread Collar: CS + underwrite. Put Spread Collar: PS + overwrite.

Risk reversal: Buy downside put, partially funded by selling an upside call, or vice versa.

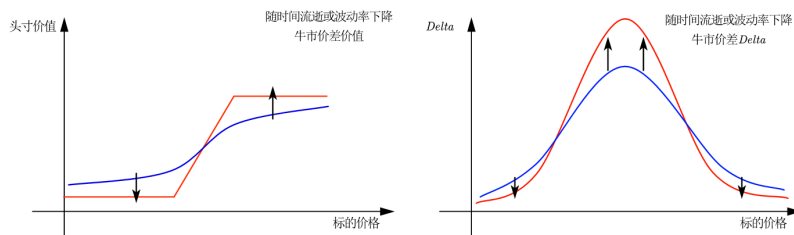


图 6: Fig5

### 3.4 其他

#### 美式期权

看涨期权价值=内在+波动率+利率-股利. 因此美式看涨若提前行权, 只可能在股利支付前一天. 看跌期权价值=内在+波动率-利率+股利. 因此美式看跌若提前行权, 要距离除息日足够远, 且波动率价值较小. 无论看涨看跌, 越是实值, 美式与欧式价格差越大; 波动率越小, 美式与欧式价格差越大.

#### 结构化

Down and out + Down and in = Vanilla Call.

Knock-in: Structure is activated after being knocked in (before that coupon is protected).

Knock-out: Immediate termination of structure.

雪球是什么: FCN-KI autocallable (fixed coupon note kick-in variant)

## 4 Linear Algebra

### 4.1 矩阵性质

- $|AB| = |A||B|$ ,  $|kA| = k^n |A|$
- $(kA)^{-1} = k^{-1}A^{-1}$ ,  $|A^{-1}| = |A|^{-1}$
- $\text{tr}(AB) = \text{tr}(BA)$ ,  $\text{tr}(AA') \geq \text{tr}(A^2)$ .
- 若  $A$  是  $n$  阶实矩阵, 则  $\text{tr}(AA') \geq 0$ , 等号成立  $\Leftrightarrow A = O$
- $r(AB) \leq \min\{r(A), r(B)\}$ ;
- $r(A+B) \leq r(A) + r(B)$ ,  $r(A-B) \leq r(A) + r(B)$ ;
- Sylvester不等式. 设  $A$  是  $m \times n$  矩阵,  $B$  是  $n \times t$  矩阵, 则  $r(AB) \geq r(A) + r(B) - n$ .
- 设  $A$  是  $m \times n$  阶实矩阵, 则  $r(A'A) = r(AA') = r(A)$
- 设  $A$  是  $m \times n$  矩阵, 则:
  - (1) 若  $r(A) = n$ , 即  $A$  列满秩, 则必存在秩等于  $n$  的  $n \times m$  矩阵  $B$  使  $BA = I_n$ ;



(2)若 $r(\mathbf{A}) = m$ ,即 $\mathbf{A}$ 行满秩,则必存在秩等于 $m$ 的 $n \times m$ 矩阵 $\mathbf{C}$ 使 $\mathbf{AC} = \mathbf{I}_m$ .

- 满秩分解. 设 $\mathbf{A}$ 是秩为 $r$ 的 $m \times n$ 矩阵,则有满秩分解 $\mathbf{A} = \mathbf{BC}$ , 其中 $\mathbf{B}$ 是秩为 $r$ 的 $m \times r$ 矩阵, $\mathbf{C}$ 是秩为 $r$ 的 $r \times n$ 矩阵.
- 设 $\mathbf{A}$ 是 $n$ 阶幂等矩阵,则存在 $n$ 阶非异阵 $\mathbf{P}$ ,使得 $\mathbf{P}^{-1}\mathbf{AP} = \begin{pmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$ , 其中 $r = r(\mathbf{A})$

## 4.2 特征值

- 任一复方阵必相似于一上三角阵,且主对角线上元素为特征值.
- 若 $\mathbf{A}$ 特征值为 $\lambda_1, \dots, \lambda_n$ ,则 $f(\mathbf{A})$ 的特征值为 $f(\lambda_1), \dots, f(\lambda_n)$ . 若 $\mathbf{A}$ 适合多项式 $g(x)$ ,则 $\mathbf{A}$ 的任意特征值适合 $g(x)$ .
- 特征值降阶公式. 设 $\mathbf{A}$ 是 $m \times n$ 矩阵, $\mathbf{B}$ 是 $n \times m$ 矩阵,且 $m \geq n$ .则 $|\lambda \mathbf{I}_m - \mathbf{AB}| = \lambda^{m-n} |\lambda \mathbf{I}_n - \mathbf{BA}|$
- 可对角化判定方法  
充分条件: 有 $n$ 个不同的特征值, 相似于实对称阵  
充要条件: 有 $n$ 个线性无关的特征向量, 极小多项式无重根, 初等因子都是一次多项式或Jordan块都是一阶矩阵
- Jordan标准型求法  
 $\lambda \mathbf{I} - \mathbf{A}$ 相抵于 $\text{diag}\{1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda)\}$ , 对角元 $d_i | d_{i+1}$ .  $d_1, \dots, d_k$ 的准素因子全体(初等因子)对应Jordan块拼起来即为Jordan标准型.
- $\mathbf{J} = \mathbf{J}_r(\lambda)$ 为Jordan块.则 $f(\mathbf{J})$ 主对角线上元素均为 $f(\lambda)$ ,上次对角线上元素均为 $f'(\lambda)$ . 距离主对角线 $j$ 位的对角平行线上的元素均为 $\frac{1}{j!} f^{(j)}(\lambda)$ . 对 $\lambda \neq 0$ ,则 $\mathbf{J}_r(\lambda)^m$ 的Jordan标准型为 $\mathbf{J}_r(\lambda^m)$
- $\mathbf{A}^k$ 求法: 先求出Jordan标准型 $\mathbf{J}$ , 根据 $\mathbf{AP} = \mathbf{PJ}$ 解方程得 $\mathbf{P}$ , 求 $\mathbf{J}^k$ 再乘过渡矩阵即可.

## 4.3 二次型

- 设 $\mathbf{A}$ 为 $n$ 阶实对称阵,则以下结论等价:  
(1) $\mathbf{A}$ 正定 (2) $\mathbf{A}$ 合同于 $\mathbf{I}_n$  (3)存在非异实矩阵 $\mathbf{C}$ ,  $\mathbf{A} = \mathbf{C}'\mathbf{C}$  (4) $\mathbf{A}$ 的主子式都大于0  
(5) $\mathbf{A}$ 的顺序主子式都大于0 (6) $\mathbf{A}$ 的特征值都大于0.
- $\mathbf{A} = (a_{ij})$ 是 $n$ 阶半正定实对称矩阵,则 $|\mathbf{A}| \leq a_{11}a_{22} \cdots a_{nn}$ , 等号成立当且仅当 $\mathbf{A}$ 是对角矩阵或某个 $a_{ii} = 0$ .
- 设 $\mathbf{A}$ 为 $n$ 阶实对称矩阵,则 $\mathbf{A}$ 半正定或半负定 $\Leftrightarrow$  对任意 $\alpha$ 满足 $\alpha' \mathbf{A} \alpha = 0$ , 均有 $\mathbf{A} \alpha = \mathbf{0}$ .
- 若实矩阵 $\mathbf{A}$ 满足 $\mathbf{A}' = \mathbf{A}^{-1}$ ,则称 $\mathbf{A}$ 为正交阵.  $\mathbf{A}$ 为正交阵/酉阵的充要条件为 $\mathbf{A}$ 的 $n$ 个行向量构成一组标准正交基.

- 实对称阵与Hermite阵特征值全为实数,反对称阵特征值全为0或纯虚数.
- 实对称阵与Hermite阵合同于 $\text{diag}\{I_p, -I_q, 0\}$ . 实对称阵(Hermite阵)正交相似(酉相似)于实对角阵.
- 复正规阵酉相似于复对角阵,对角元模长都为1.实反对称阵酉相似于复对角阵.
- Gram-Schmidt正交化方法.  $\{u_1, \dots, u_m\}$ 为一组线性无关的向量, 令

$$v_{k+1} = u_{k+1} - \sum_{i=1}^k \frac{(u_{k+1}, v_i)}{\|v_i\|^2} v_i$$

则 $\{v_1, \dots, v_m\}$ 为一组正交向量.

#### 4.4 一些分解

- QR分解

设 $A$ 为 $n$ 阶实/复矩阵,则 $A = QR$ ,其中 $Q$ 为正交/酉阵, $R$ 为主对角元均 $\geq 0$ 的上三角阵.若 $A$ 非异,则分解唯一.

- Jordan-Chevalley分解

设 $A$ 为 $n$ 阶复方阵,则有分解 $A = B + C$ ,满足:

(1) $B$ 可对角化 (2) $C$ 为幂零阵 (3) $BC = CB$  (4) $B, C$ 可表示为 $A$ 的多项式

并且满足前三条的分解唯一.

- Cholesky分解

设 $A$ 为 $n$ 阶实对称阵/Hermite阵,若 $A$ 正定,则存在对角元均为正实数的下三角(或上三角)实矩阵 $L$ 使得 $A = LL' / A = L\overline{L}'$ , 并且分解唯一.若 $A$ 为半正定阵,则对角元均为非负数,且分解不一定唯一.

- 极分解

$\varphi \in \mathcal{L}(V)$ ,则 $\varphi = \omega\psi$ .其中 $\omega$ 保积, $\psi$ 半正定自伴随.若 $\varphi$ 可逆,则分解唯一.

$A \in M_n(R)$ ,则 $A = OS$ ,其中 $O$ 为正交阵, $S$ 为半正定实对称阵.

$A \in M_n(C)$ ,则 $A = UH$ ,其中 $U$ 为酉阵, $H$ 为半正定Hermite阵.

- 奇异值分解

$A \in M_{m \times n}(R)$ ,则 $A = P \text{diag}\{S, 0\} Q'$ . 其中 $P, Q$ 为 $m, n$ 阶正交阵, $S = \text{diag}\{\sigma_1, \dots, \sigma_r\}$ ,  $\sigma_1 \geq \dots \geq \sigma_r > 0$ 为 $A$ 的所有奇异值,即 $\sigma_1^2, \dots, \sigma_r^2$ 为 $A'A$ 的非零特征值全体.

## 4.5 计算

$A$ 是实对称阵, 求非异矩阵 $C$ 使得 $C'AC$ 是对角阵: 作 $n \times 2n$ 矩阵 $(A, I_n)$ , 实施初等行变换, 并对左侧矩阵实施对应初等列变换, 将左侧化为对角阵. 此时右侧矩阵的即为 $C'$ .

$A$ 是实对称阵, 求正交矩阵 $P$ 使得 $P'AP$ 是对角阵: 求特征值, 对每个特征值 $\lambda$ , 求解 $(\lambda I - A)x = 0$ 得到基础解系. 若基础解系有多个向量, 进行正交化. 所有列向量单位化后拼接得到 $P$ .

# 5 Calculus

## 5.1 一元微积分

Please see 数分2 notes attached.

## 5.2 多元微分

Hessian matrix:  $f$ 为 $\mathbb{R}^n \rightarrow \mathbb{R}$ 函数,  $H(f) = \left( \frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{n \times n}$

Jacobi matrix:  $f$ 为 $\mathbb{R}^n \rightarrow \mathbb{R}^m$ 向量值函数,  $J(f) = \left( \frac{\partial f_i}{\partial x_j} \right)_{m \times n}$

Prop.  $H(f(x)) = J(\nabla f(x))^T$ .

Hessian矩阵正定的极值点是极小值点, 负定是极大值点.

隐函数存在定理: 对多元函数 $F(x_1, \dots, x_n, y) = 0$ , 在周围闭矩形连续且有连续偏导,  $F_y \neq 0$ , 有 $\frac{\partial y}{\partial x_i} = -\frac{F_{x_i}}{F_y}$ . 多元向量值函数 $F(x_0, y_0, u_0, v_0) = 0, G(x_0, y_0, u_0, v_0) = 0$ , Jacobi行列式 $\frac{\partial(F, G)}{\partial(u, v)} \neq 0$ , 则有 $\begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = -\frac{\partial(F, G)}{\partial(u, v)}^{-1} \frac{\partial(F, G)}{\partial(x, y)}$ .

几何应用

1. 曲线 $\Gamma: x = x(t), y = y(t), z = z(t)$ , 在 $P_0(x_0, y_0, z_0)$ 处切向量 $\vec{\tau} = (x', y', z')|_{P_0}$ , 法平面为 $x'(x - x_0) + y'(y - y_0) + z'(z - z_0) = 0$ .

2. 曲线 $\Gamma: \begin{cases} F(x, y, z) = 0 \\ G(x, y, z) = 0 \end{cases}$ , 在 $P_0(x_0, y_0, z_0)$ 处切向量 $\vec{\tau} = \left( \frac{\partial(F, G)}{\partial(y, z)}, \frac{\partial(F, G)}{\partial(z, x)}, \frac{\partial(F, G)}{\partial(x, y)} \right)|_{P_0}$ , 法平面为 $\text{grad } F(P_0), \text{grad } G(P_0)$ 张成的平面.

3. 曲面 $F(x, y, z) = 0$ 在 $P_0(x_0, y_0, z_0)$ 处法向量 $\vec{n} = (F_x, F_y, F_z)|_{P_0}$ , 切平面为 $\frac{x - x_0}{F_x(P_0)} = \frac{y - y_0}{F_y(P_0)} = \frac{z - z_0}{F_z(P_0)}$ .

4. 曲面 $x = x(u, v), y = y(u, v), z = z(u, v)$ , 法向量 $\vec{n} = \left( \frac{\partial(y, z)}{\partial(u, v)}, \frac{\partial(z, x)}{\partial(u, v)}, \frac{\partial(x, y)}{\partial(u, v)} \right)$ .

### 5.3 重积分

重积分变量代换: 映射  $T: \begin{cases} x = x(u, v) \\ y = y(u, v) \end{cases}, D \rightarrow T(D)$  有连续偏导且  $\frac{\partial(x, y)}{\partial(u, v)} \neq 0$ .  $f(x, y)$  在  $T(D)$  连续, 则  $\iint_{T(D)} f(x, y) dx dy = \iint_D f(x(u, v), y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv$

常用变量代换:

极坐标变换  $x = r \cos \theta, y = r \sin \theta$  下  $dx dy = r dr d\theta$ .

柱坐标变换  $x = r \cos \theta, y = r \sin \theta, z = z$  下  $dx dy dz = r dr d\theta dz$ .

球坐标变换  $x = r \sin \varphi \cos \theta, y = r \sin \varphi \sin \theta, z = r \cos \varphi, r \in [0, +\infty), \varphi \in [0, \pi], \theta \in [0, 2\pi]$  下  $dx dy dz = r^2 \sin \varphi dr d\varphi d\theta$ .

广义球坐标变换  $x = ar \sin \varphi \cos \theta, y = br \sin \varphi \sin \theta, z = cr \cos \varphi$  下  $dx dy dz = abcr^2 \sin \varphi dr d\varphi d\theta$ .

$n$  重球坐标变换  $x_1 = r \cos \varphi_1, x_2 = r \sin \varphi_1 \cos \varphi_2, \dots, x_{n-1} = r \sin \varphi_1 \cdots \sin \varphi_{n-2} \cos \varphi_{n-1}, x_n = r \sin \varphi_1 \cdots \sin \varphi_{n-1}, 0 \leq \varphi_1, \dots, \varphi_{n-2} \leq \pi, 0 \leq \varphi_{n-1} \leq 2\pi$  下  $\frac{\partial(x_1, x_2, \dots, x_n)}{\partial(r, \varphi_1, \dots, \varphi_{n-1})} = r^{n-1} \sin^{n-2} \varphi_1 \sin^{n-3} \varphi_2 \cdots \sin \varphi_{n-2}$ .

### 5.4 曲线与曲面积分

Please see 数分3 notes attached.

## 6 常微

#### 1. 分离变量

(1) 齐次方程  $\frac{dx}{dt} = f\left(\frac{x}{t}\right)$ . 令  $\frac{x}{t} = u$ , 则  $\frac{du}{dt} t = f(u) - u$ .

(2)  $\frac{dx}{dt} = f(at + bx + c)$ . 令  $y = at + bx + c$ , 则  $\frac{dy}{dt} = a + bf(y)$ .

#### 2. 非齐次线性方程 $\frac{dx}{dt} = P(t)x + Q(t)$

常数变易, 两边乘  $e^{-\int P(t)dt}$ . 常数变易公式  $x = e^{\int P(t)dt} \left( \int e^{-\int P(t)dt} Q(t) dt + C \right)$ .

#### 3. 全微分方程 $M(t, x)dt + N(t, x)dx = 0$

(1)  $\frac{\partial M(t, x)}{\partial x} = \frac{\partial N(t, x)}{\partial t}$ , 则存在  $u(t, x)$  使得  $du(t, x) = M(t, x)dt + N(t, x)dx$

设  $u(t, x) = \int M(t, x)dt + \varphi(x)$ , 由  $N(t, x) = \frac{\partial}{\partial x} \int M(t, x)dt + \varphi'(x)$  可求出  $\varphi(x)$ .

(2)  $\frac{\partial M(t, x)}{\partial x} \neq \frac{\partial N(t, x)}{\partial t}$ , 要找积分因子  $\mu$  使得  $\frac{\partial \mu M}{\partial x} = \frac{\partial \mu N}{\partial t}$

可寻找特殊形式的  $\mu(t, x)$ , 如取  $\mu(x), \mu(t), \mu(xt), \dots$ , 代入  $\frac{\partial \mu M}{\partial x} = \frac{\partial \mu N}{\partial t}$  求解.

#### 4. 导数未解出的一阶方程

$x = f(t, \frac{dx}{dt})$  或  $t = f(x, \frac{dx}{dt})$ , 引入  $p = \frac{dx}{dt}$  求解.

### 5. 高阶方程的降阶

(1) 不含  $x, F(t, x^{(k)}, \dots, x^{(n)}) = 0$ . 令  $x^{(k)} = y$ , 则化为  $F(t, y, \dots, y^{(n-k)}) = 0$ .

(2) 不含  $t, F(x, \frac{dx}{dt}, \dots, \frac{d^n x}{dt^n}) = 0$ . 引入  $y = \frac{dx}{dt}$ , 以  $x$  为新自变量. 则  $x^{(k)}$  可用  $y, \frac{dy}{dx}, \dots, \frac{d^{k-1}y}{dx^{k-1}}$  表示.

### 6. 齐次常系数线性方程 $x^{(n)} + a_1 x^{(n-1)} + \dots + a_{n-1} x^{(1)} + a_n x = 0$

特征方程  $\lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n = 0$ .  $\lambda_1, \dots, \lambda_s$  为不同实根, 重数  $n_1, \dots, n_s$ ,  $\alpha_1 \pm i\beta_1, \dots, \alpha_p \pm i\beta_p$  为共轭虚根, 重数为  $m_1, \dots, m_p$ . 则实值通解

$$x(t) = \sum_{i=1}^s P_i(t) e^{\lambda_i t} + \sum_{i=1}^p e^{\alpha_i t} (W_i(t) \cos \beta_i t + V_i(t) \sin \beta_i t)$$

其中  $P_i(t)$  为  $n_i - 1$  次实系数多项式,  $W_i(t), V_i(t)$  为  $m_i - 1$  次实系数多项式.

### 7. 非齐次常系数线性方程 $x^{(n)} + a_1 x^{(n-1)} + \dots + a_{n-1} x^{(1)} + a_n x = f(t)$

解为  $x(t) = x_1(t) + x^*(t)$ , 其中  $x_1(t)$  为  $x^{(n)} + a_1 x^{(n-1)} + \dots + a_{n-1} x^{(1)} + a_n x = 0$  的通解,  $x^*(t)$  为  $x^{(n)} + a_1 x^{(n-1)} + \dots + a_{n-1} x^{(1)} + a_n x = f(t)$  的特解.

$x^*(t) = \int_0^t K(t-s) f(s) ds$ , 其中  $K(t)$  为  $x^{(n)} + a_1 x^{(n-1)} + \dots + a_{n-1} x^{(1)} + a_n x = 0$  在初值条件  $K(0) = K'(0) = \dots = K^{(n-1)}(0) = 0, K^{(n)}(0) = 1$  下的解.

### 8. Euler方程 $t^n x^{(n)} + a_1 t^{n-1} x^{(n-1)} + \dots + a_{n-1} t x^{(1)} + a_n x = 0$

引入  $s = \ln |t|$ , 记  $D_s = \frac{d}{ds}$ , 则  $t^n x^{(n)} = D_s(D_s - 1) \dots (D_s - n + 1)x$ .

### 9. 齐次常系数线性微分方程组 $\frac{d\vec{x}}{dt} = A\vec{x}$

若  $A$  可对角化, 则  $\vec{x}(t) = \sum_{i=1}^n c_i e^{\lambda_i t} \vec{p}_i$ ,  $\vec{p}_i$  为  $\lambda_i$  的特征向量.

若  $A$  有复数特征值  $\lambda = \alpha + i\beta$  和  $\bar{\lambda} = \alpha - i\beta$ , 则方程组实值解形如  $\vec{x} = e^{\alpha t} \{ \vec{p}(t) \cos \beta t + \vec{q}(t) \sin \beta t \}$ , 其中  $\vec{p}(t)$  和  $\vec{q}(t)$  是  $t$  的次数小于  $\lambda$  重数的实向量多项式, 其向量系数由微分方程组确定.

对一般矩阵  $A$ , Jordan 标准型为  $J, P^{-1}AP = J$ , 则  $\vec{x}(t) = P e^{Jt} \vec{c}$ , 其中对  $\lambda_i$  的 Jordan 块  $J_i, e^{J_i t}$  第一行为  $e^{\lambda_i t}, t e^{\lambda_i t}, \dots, \frac{t^{n_i-1}}{(n_i-1)!} e^{\lambda_i t}$ . 记  $\vec{p}_i$  为  $P$  的第  $i$  个列向量, 则有

$$\vec{x}(t) = \sum_{i=1}^s \left\{ c_{i_1} e^{\lambda_i t} \vec{p}_{i_1} + c_{i_2} [t e^{\lambda_i t} \vec{p}_{i_1} + e^{\lambda_i t} \vec{p}_{i_2}] + \dots + c_{i_{n_i}} \left[ \frac{t^{n_i-1}}{(n_i-1)!} e^{\lambda_i t} \vec{p}_{i_1} + \dots + e^{\lambda_i t} \vec{p}_{i_{n_i}} \right] \right\}.$$

$\forall \operatorname{Re}\{\lambda_i\} < 0 \Leftrightarrow \|\vec{x}(t)\| \leq M.$

### 10. 非齐次常系数线性微分方程组 $\frac{d\vec{x}}{dt} = A\vec{x} + \vec{f}(t)$

方程的解  $\vec{x}(t) = Pe^{Jt}\vec{c}(t) = Pe^{Jt}\vec{c} + \int_0^t e^{A(t-s)}\vec{f}(s)ds$

对初值问题  $\vec{x}(t_0) = \vec{x}_0$ , 有  $\vec{x}(t) = e^{A(t-t_0)}\vec{x}_0 + \int_{t_0}^t e^{A(t-s)}\vec{f}(s)ds$ .

## 7 Others

### 7.1 矩阵范数

Frobenius 范数  $\|A\| = \sqrt{\sum_{i,j} a_{ij}^2}$

由向量范数  $\|\vec{x}\|_p = (\sum_i |x_i|^p)^{1/p}$  诱导的矩阵范数  $\|A\|_p = \sup_{\vec{x} \neq \vec{0}} \frac{\|A\vec{x}\|_p}{\|\vec{x}\|_p}$ .

$\|A\|_1 = \max_{1 \leq j \leq n} \left\{ \sum_{i=1}^n |a_{ij}| \right\}$ ,  $\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$ ,  $\|A\|_\infty = \max_{1 \leq i \leq n} \left\{ \sum_{j=1}^n |a_{ij}| \right\}$ .

由向量范数诱导的矩阵范数有性质 (1)  $\|A\vec{x}\| \leq \|A\| \cdot \|\vec{x}\|$  (2)  $\|A\| = \sup_{\|\vec{x}\|=1} \|A\vec{x}\|$  (3)  $\|AB\| \leq$

$$\|A\| \cdot \|B\| \quad (4) \left\| \int_\alpha^\beta A(s)ds \right\| \leq \left| \int_\alpha^\beta \|A(s)\|ds \right|$$

矩阵指数函数:  $(e^{At})^{-1} = e^{-At}$ ,  $e^{A(t+s)} = e^{At} \cdot e^{As}$ .

### 7.2 矩阵求导

关于  $t$  求导与积分: 对矩阵每个元素求导与积分.

$$\frac{d}{dt}[A(t)B(t)] = A'(t)B(t) + A(t)B'(t).$$

将矩阵  $X_{m \times n}$  按列堆栈向量化,  $vec(X) = [x_{11}, \dots, x_{m1}, \dots, \dots, x_{mn}]^T$ . 将矩阵函数  $F: m \times n \rightarrow p \times q$  向量化,  $vec(F(X)) = [f_{11}(X), \dots, f_{p1}(X), \dots, \dots, f_{pq}(X)]^T$ .

$$D_X F(X) = \frac{\partial vec_{pq \times 1}(F(X))}{\partial vec_{mn \times 1}^T X} \in \mathbb{R}^{pq \times mn}.$$

向量变元标量函数:  $\frac{\partial a^T x}{\partial x} = a$ ,  $\frac{\partial x^T A x}{\partial x} = Ax + A^T x$ ,  $\frac{\partial a^T x x^T b}{\partial x} = ab^T x + ba^T x$ .

矩阵变元标量函数  $\frac{\partial a^T X b}{\partial X} = ab^T$ ,  $\frac{\partial a^T X X^T b}{\partial X} = ab^T X + ba^T X$ .