

# 1 Probability and Stochastic Process

## 1.1 Probability

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

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

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \text{Cov}(X, Y).$$

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

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

矩母函数 Moment generating function: For rv  $X$ , MGF is  $M_X(t) = \mathbb{E}[e^{tX}]$ . For vec-valued rv  $\mathbf{X}$ ,  $M_X(\mathbf{t}) = \mathbb{E}[e^{\langle \mathbf{t}, \mathbf{X} \rangle}]$ . 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  $\mathbf{X}$ ,  $\varphi_X(\mathbf{t}) = \mathbb{E}[e^{i\langle \mathbf{t}, \mathbf{X} \rangle}]$ .

偏度 Skewness  $\gamma_1 = \mathbb{E}[(X - \mathbb{E}X)^3] / \text{Var}(X)^{3/2}$ . 峰度 Kurtosis  $\gamma_2 = \mathbb{E}[(X - \mathbb{E}X)^4] / \text{Var}(X)^2$ .

一些分布

- Binomial( $p, n$ ):  $P_X(k) = C_n^k p^k (1-p)^{n-k}$ .  $n$  次尝试的成功次数. 期望  $np$ ,  $\text{Var } np(1-p)$ .
- Geometric( $p$ ):  $P_X(k) = (1-p)^{k-1}p$ . 第一次成功需要的尝试次数. 期望  $\frac{1}{p}$ ,  $\text{Var } \frac{1-p}{p^2}$ .
- Poisson( $\lambda$ ):  $P_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}$ . 期望  $\lambda$ ,  $\text{Var } \lambda$ . 当  $\lambda > 20$ , 分布  $\approx \mathcal{N}(\lambda, \lambda)$ .
- 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  $\frac{1}{\lambda}$ ,  $\text{Var } \frac{1}{\lambda^2}$ .
- $\Gamma(\alpha, \beta)$ :  $f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ . 均值  $\frac{\alpha}{\beta}$ , 方差  $\frac{\alpha}{\beta^2}$ . MGF:  $(\frac{\beta}{\beta-t})^\alpha$ . 独立的  $X_i \sim \Gamma(\alpha_i, \beta)$  有  $\sum X_i \sim \Gamma(\sum \alpha_i, \beta)$ .
- $\chi_n^2$  自由度  $n$  的卡方分布为  $\sum_{i=1}^n X_i^2$  其中  $X_i$  是独立标准正态. 是特殊的Gamma分布  $\Gamma(\frac{n}{2}, \frac{1}{2})$ . 均值等于自由度  $n$ , 方差为  $2n$ .
- Fisher  $F$  分布  $F_{k,m} = \frac{X/k}{Y/m}$  其中  $X \sim \chi_k^2, Y \sim \chi_m^2$  为独立卡方分布. 期望为  $\frac{m}{m-2}$  对  $m > 2$ , 期望发散对  $m \leq 2$ .
- Student  $t$  分布  $T = \frac{X}{\sqrt{\frac{1}{n} \sum_{i=1}^n Y_i^2}}$  其中  $X, Y_i$  为独立标准正态分布. 对称分布, 形状类似正态, 但tail更肥. pdf有  $f_T(t) = t f_{1,n}(t^2)$  其中  $f_{1,n}$  为  $F_{1,n}$  的pdf.

**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.

$n$ 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(独立正态分布的和一定是正态分布). 反例:  $X \sim \mathcal{N}(0, 1), Y = \varepsilon X$  where  $\varepsilon = \pm 1$  各一半概率且与 $X$ 独立.  $X+Y$ 不是正态.

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  $f(\mathbf{x}) = (2\pi)^{-d/2} \det(\Sigma)^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu)\right)$ ,

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

条件概率条件期望

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

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 $f_Y(y) = f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$ .
- $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2), X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$  独立, 则 $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ ,
- $X \sim \mathcal{N}(\mu, \sigma^2)$ , 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.
- $X \sim \mathcal{N}(\mu_X, \sigma_X^2), Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ , corr  $\rho$ , 则  $X|Y \sim \mathcal{N}\left(\mu_X + \rho \frac{\sigma_X}{\sigma_Y}(Y - \mu_Y), (1 - \rho^2)\sigma_X^2\right)$ .
- 对于 std norm  $X, Y$  corr  $\rho$ , 有 quadrant prob formula  $P(X > 0, Y > 0) = \frac{1}{4} + \frac{1}{2\pi} \arcsin(\rho)$ .
- 在 $[0, 1]$ 的均匀分布下取  $n$  个值, 它们的和小于1的概率是  $\frac{1}{n!}$ .
- 在  $[0, 1]$  的均匀分布下取  $n$  个值, 最大值的期望是  $\frac{n}{n+1}$ , 最大值的密度  $f_X(x) = n(1-x)^{n-1}$ . 第  $i$  大的值期望是  $\frac{i}{n+1}$ .

- 当  $X \sim f, Y \sim g$  独立,  $X + Y$  有  $\mathbb{P}(X + Y \leq z) = \int \mathbb{P}(x \leq z - y)g(y)dy$ , pdf 为  $h(z) = \int f(z - y)g(y)dy$ .
- 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.
- 从  $\mathcal{N}(\mu, \sigma^2)$  抽样  $Y_1, \dots, Y_n$ . 样本均值有  $\bar{Y}_n \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$ . 样本方差  $s^2 = \frac{1}{n-1} \sum (Y_i - \bar{Y}_n)^2$  有  $(n-1) \frac{s^2}{\sigma^2} \sim \chi_{n-1}^2$ .

## 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:  $\frac{1}{a} B(a^2 t)$  是 BM. 可得 first passage time  $T_a = a^2 T_1$  in distr.
- Time inversion:  $X(t) = t B(\frac{1}{t}) (t > 0), X(0) = 0$  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.
- 考虑停时  $\tau = \min\{t > 0 : B(t) = -a \text{ or } B(t) = b\}$ . 有  $\mathbb{E}[\tau] = ab, \text{Var}(\tau) = \frac{ab(a^2 + b^2)}{3}$ .
- 指数鞅  $M_t(\lambda) = \exp(\lambda B(t) - \frac{1}{2} \lambda^2 t)$  对  $\lambda \in \mathbb{R}$ . 可以由指数鞅生成鞅  $H_n(B_t, t) = \frac{\partial^n}{\partial \lambda^n} M_t(\lambda) |_{\lambda=0}$ , 如  $H_1 = B_t, H_2 = B_t^2 - t, H_3 = B_t^3 - 3tB_t, H_4 = B_t^4 - 6tB_t^2 + 3t^2$ .
- 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: 由  $dX(t) = \Delta(X(t), t)dW(t) + \Theta(X(t), t)dt$  形式SDE决定的随机过程  $X(t)$  为Ito过程.

特别地, 若  $\Delta$  只关于  $t$ , 则有  $[X, X](t) = \int_0^t \Delta(u)^2 du$ .

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

### 1.3 Risk-Neutral Pricing

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  $Z(t) = \exp \left( - \int_0^t \Theta(u) dW(u) - \frac{1}{2} \int_0^t \Theta(u)^2 du \right)$ ,  $\widetilde{W}(t) = W(t) + \int_0^t \Theta(u) du$ , 并假设  $\Theta(u)Z(u)$  平方可积. 令  $Z = Z(T)$ , 则  $\mathbb{E}Z = 1$ , 且  $\widetilde{W}(t)$  是  $\widetilde{\mathbb{P}}$  下的BM, 其中  $\widetilde{\mathbb{P}}$  是  $Z$  作为 Radon-Nikodym 导数得到的测度, 即  $\widetilde{\mathbb{P}}(A) = \int_A Z(\omega) d\mathbb{P}(\omega)$ .

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

Martingale Representation Thm.  $M(t)$  是平方可积鞅, 则存在适应的平方可积过程  $\theta(t)$  使得  $M(t)$  a.s. 有积分表示  $M(t) = \mathbb{E}[M_0] + \int_0^t \theta(s) dW(s)$ .

贴现后的 deriv value process  $V^*(t)$  是鞅, 由鞅表示定理有  $\theta(t)$  使  $V^*(T) = \widetilde{\mathbb{E}}V_0 + \int_0^T \theta(t) d\widetilde{W}(t) = \widetilde{\mathbb{E}}V_0 + \int_0^T \frac{\theta(t)}{\sigma(t)S^*(t)} dS^*(t)$ , 即为复制策略.

Ornstein-Uhlenbeck (OU): 连续时间均值回归随机过程. SDE定义:  $dX_t = \kappa(\theta - X_t)dt + \sigma dW_t$ . 显示定义:  $X_t = \theta + (X_0 - \theta)e^{-\kappa t} + \sigma \int_0^t e^{-\kappa(t-s)} dW_s$ . (因此不是鞅) 分布定义:  $X_t | X_0 \sim \mathcal{N}(\theta + (X_0 - \theta)e^{-\kappa t}, \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t}))$ . 应用: Vasicek, Heston Var.

## 1.4 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 Stats

### 2.1 Regression

假设: (1)  $y = X\beta + \varepsilon$  (2) 随机抽样 (3) 无完全共线性 (4) 严格外生性  $\mathbb{E}[\varepsilon|X] = 0$  (5) 球形扰动/Homoskedasticity  $\text{Var}(\varepsilon|X) = \sigma^2 I_n$  (6) 正态性  $\varepsilon|X \sim \mathcal{N}(0, \sigma^2 I_n)$ .

估计  $\hat{\beta} = (X^T X)^{-1} X^T y$ . (1)-(4) 时  $\hat{\beta}$  是无偏的, (1)-(5) 时是 BLUE 最佳线性无偏估计, (1)-(6) 时  $\hat{\beta}$  是最优估计量.

如果异方差(对角线不相等)/自相关(非对角不为0), 违反(5), 估计不再是最佳的(用广义最小二乘GLS变量转换), 并且假设检验失效(用稳健标准误).

估计的分布是正态,  $\mathbb{E}[\hat{\beta}] = \beta, \text{Var}(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$ .

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

Notation: 解释变量的个数为  $k$  不包括截距, 回归参数总数  $p$  包含截距 =  $k + 1$  不包含截距 =  $k$ .

记  $H = X(X^T X)^{-1} X^T$ , 有  $\hat{y} = Hy$ .  $H$  对称幂等,  $HX = X$ , rank 和 trace 均为  $p$ .

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

$ESS = SSR = \sum (\hat{y}_i - \bar{y})^2, df = p - 1. RSS = SSE = \sum (\hat{y}_i - y_i)^2. SST = \sum (y_i - \bar{y})^2$ . 有  $SST = SSE + SSR$ . 定义  $R^2 = SSR/SST$ . (ESS: explained, SSE: error, SSR: regression, RSS: residual).

$\text{Var}(\hat{\beta}_j) = \frac{\sigma^2}{SST_j(1 - R_j^2)}$ , 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 $\hat{\sigma}^2 = \frac{RSS}{n-p}$ (或写为 $se^2$ ). 残差自由度 $df = n - p$ .

$\hat{\beta}_j$ 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 $Cov(\hat{\beta})$ 对应角元的开根.

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 $F = \frac{R^2/k}{(1 - R^2)/(n - k - 1)}$ .

多重共线性会导致 $X^T X$ 某些特征值非常小, 求逆后方差膨胀(放大样本噪声), 系数不稳定, t检验失效. 衡量多重共线性 $VIF_j = \frac{1}{1 - R_j^2}$ , 大于10认为严重. 系数有 $SE(\beta_j) = \frac{\sigma^2 VIF_j}{(n - 1)Var(X_j)}$ .

解决方法: 删除列, 做PCA, 使用Ridge/Lasso.

Ridge回归目标函数 $\min_{\beta} \|y - X\beta\|^2 + \lambda\|\beta\|_2^2$ , 通过惩罚项约束系数的大小. 有闭式解 $\hat{\beta} = (X^T X + \lambda I)^{-1} X^T y$ . 惩罚项避免了特征值接近0时求逆的不稳定. Lasso惩罚项是 $L^1$ 的, 有变量选择的效果.  $L1$ 约束的菱形角点容易与最小二乘等高线相交于坐标轴上, 因此Lasso倾向于让部分系数精确为零.

$y$ 对 $x$ 做一元回归, 回归系数(无截距) $\hat{\beta}_{yx} = \frac{\sum x_i y_i}{\sum x_i^2}$ , (有截距) $\hat{\beta}_{yx} = \frac{Cov(x, y)}{Var(x)}$ .

有截距的一元回归 $y \sim (1, x)$ 有 $R^2 = r^2$ . 如果 $y \sim (1, x_1)$ 得到 $r_1^2$ ,  $y \sim (1, x_2)$ 得到 $r_2^2$ , 且 $x_1, x_2$ 的相关系数是 $\rho$ , 则 $y \sim (1, x_1, x_2)$ 的 $R^2 = \frac{r_1^2 + r_2^2 - 2r_1 r_2 \rho}{1 - \rho^2}$ .

对于 $k$ 个解释变量, 包含截距的回归 $R^2$ 的分布 $\sim Beta(\frac{k}{2}, \frac{n - k - 1}{2})$ , 其中 $Beta(\alpha, \beta)$ 的期望为 $\frac{1}{1 + \beta/\alpha}$ . 不含截距时没有Beta分布性质, 并且 $R^2$ 可能为负.

样本翻一倍:  $\hat{\beta}$ 以及 $R^2$ 不变, 在 $\sigma^2$ 固定的情况下,  $Var(\hat{\beta})$ 变为原来的 $\frac{1}{2}$ , 相应 $t$ 变为 $\sqrt{2}$ 倍,  $F$ 变为2倍. 但实际 $\hat{\sigma}_{new}^2 = (1 - \frac{p}{2n - p})\hat{\sigma}^2$ ,  $t$ 变为 $\sqrt{2 + \frac{p}{n - p}}$ 倍, 其中 $n$ 为样本量,  $p$ 为解释变量个数.

分块OLS: 将很大的数据 $X$ 分成 $X_1, X_2, \dots, X_M$ 块, 对第 $m$ 份计算 $S_{xx}^{(m)} = X_m^T X_m$ 和 $S_{xy}^{(m)} = X_m^T y_m$ . 合并 $S_{xx} = \sum_m S_{xx}^{(m)}$ ,  $S_{xy} = \sum_m S_{xy}^{(m)}$ . 最后计算 $\hat{\beta} = (S_{xx})^{-1} S_{xy}$ . 这与直接对 $X$ 计算OLS结果相同.

## 2.2 Hypothesis Testing

Confidence Interval: 95% 置信区间指在同样模型假设下做重复抽样无数次, 有95%的概率包含真实参数值的区间. 即 $\mathbb{P}(\hat{\theta}_{lower}(X) < \theta_{true} < \hat{\theta}_{upper}(X)) = 0.95$ , 这里 $X$ 是随机的, 而不是 $\theta_{true}$ .

检查正态分布的均值: t-test,  $H_0: \mu = \mu_0$ .

统计量 $T = \sqrt{\frac{n-1}{\hat{\sigma}^2}}(\hat{\mu} - \mu_0)$ ,  $H_0$ 下有 $T \sim t_{n-1}$ . 如果 $|T| > t_{\alpha/2, n-1}$ , 则拒绝 $H_0$ .

检查正态分布的方差:  $H_0: \sigma^2 = \sigma_0^2$ .

统计量 $Q = \frac{n\hat{\sigma}^2}{\sigma_0^2}$ ,  $H_0$ 下有 $Q \sim \chi_{n-1}^2$ . 如果 $Q > \chi_{\alpha, n-1}^2$ , 则拒绝 $H_0$ .

对两个正态分布 $X, Y$ 的均值做假设检验:  $H_0: \mu_X = \mu_Y$ .

(1) 数据量相同, 直接检验 $Z_i = X_i - Y_i$ .

(2)  $X$  有  $n$  个,  $Y$  有  $m$  个. 在  $\sigma_X^2 = \sigma_Y^2$  假设下, 统计量  $T = \frac{(\hat{\mu}_X - \hat{\mu}_Y)(\sqrt{m+n-2})}{\sqrt{\frac{1}{n} + \frac{1}{m}} \sqrt{n\hat{\sigma}_X^2 + m\hat{\sigma}_Y^2}}$ .

$H_0$  下有  $T \sim t_{n+m-2}$ . 如果  $|T| > t_{\alpha/2, n+m-2}$ , 则拒绝  $H_0$ .

对两个正态分布  $X, Y$  的方差做假设检验, F-test,  $H_0 : \sigma_X^2 = \sigma_Y^2$ .

统计量  $F = \frac{(m-1)n\hat{\sigma}_X^2}{(n-1)m\hat{\sigma}_Y^2}$ ,  $H_0$  下有  $F \sim \frac{\sigma_X^2}{\sigma_Y^2} F_{n-1, m-1}$ .

OLS 单个 regressor 系数的假设检验:  $H_0 : \beta_l = \beta_l^0$  对某个  $l$ .

记估计结果为  $b$ , 有  $b_l - \beta_l \sim \mathcal{N}(0, \sigma^2((X^T X)^{-1})_{ll})$ . 统计量  $t_l = \frac{b_l - \beta_l^0}{\hat{s}((X^T X)^{-1})_{ll}^{1/2}}$  其中  $s^2 = \frac{e^T e}{n-k}$ . 在  $H_0$  下有  $t_l \sim t_{n-k}$ .

### 3 Finance

#### 3.1 Pricing

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

Option value wrt strike price is convex (consider the payoff of fly). Option value wrt vol is usually (but not always) convex (see Volga expression).

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

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: 关于 Spot 一阶导. Gamma: 关于 Spot 二阶导. Vega: 关于 Vol 一阶导. Theta: 关于  $t$  一阶导. Rho: 关于利率一阶导.

##### 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.

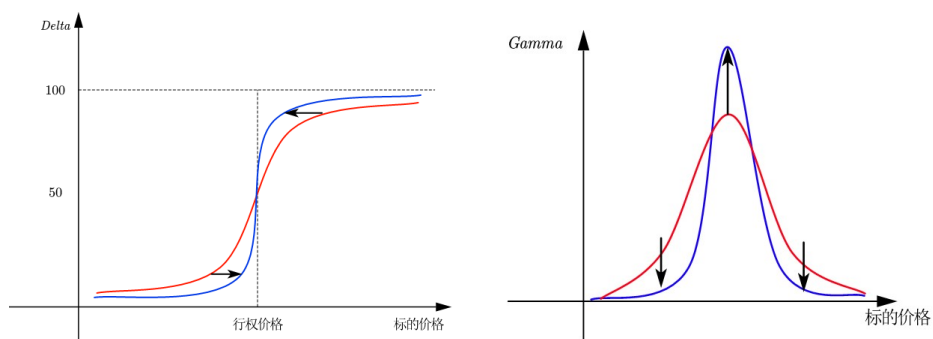


图 1: Fig1

### 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.

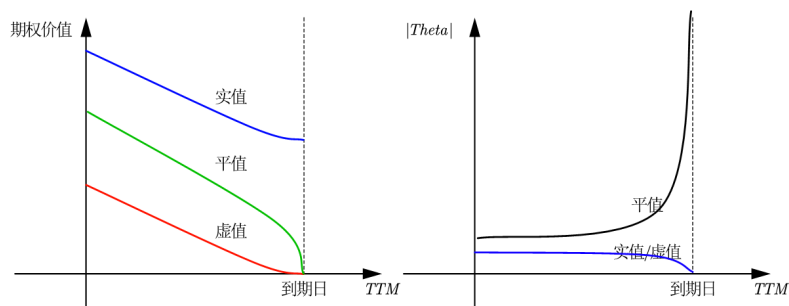


图 2: Fig2

- 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.



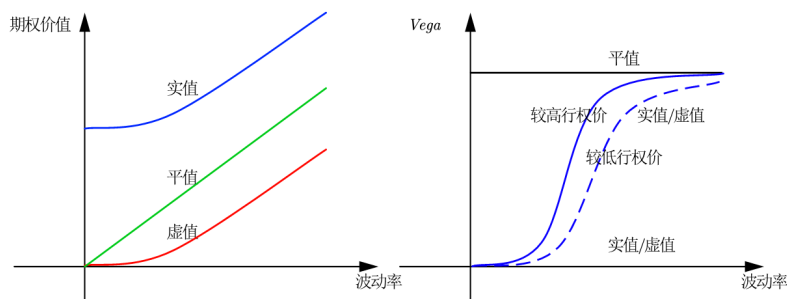


图 3: Fig3

- Vega值随到期日临近的变化见Fig4.

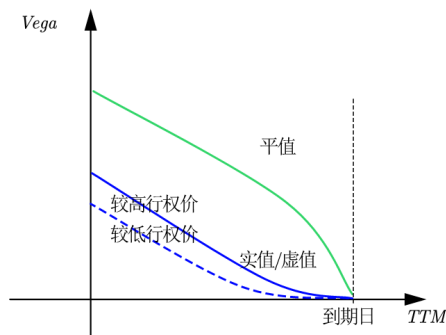


图 4: Fig4

### 3.3 Structure

#### 对称策略

跨式期权 Straddle: Long 100 C + Long 100 P ( $\backslash /$ ). Long Gamma, Long Vega, Short Theta.

宽跨式期权 Strangle: Long 105 C + Long 95 P ( $\backslash - /$ ). Long Gamma, Long Vega, Short Theta.

蝶式期权 Fly: Long 95 C + Long 105 C + Short 2 \* 100 C ( $- / \backslash -$ ). Short Gamma, Short Vega, Long Theta.

鹰式期权 Iron Condor: Long 95 C + Long 105 C + Short 97.5 C + Short 102.5 C ( $- / - \backslash -$ ). Short Gamma, Short Vega, Long Theta.

#### 时间价差

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

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

Long Vega, 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).

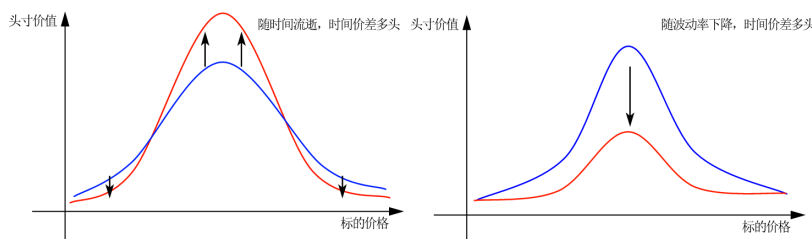


图 5: Fig5

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.

Call Spread: Buy 105 C + Sell 110 C. Put Spread: Buy 95 P + Sell 90 P.

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.

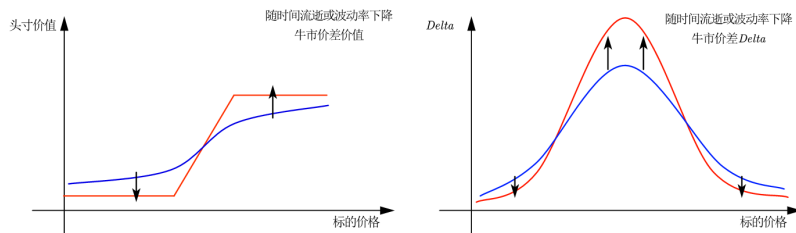


图 6: Fig5

## 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.

严格说明在 no dividend 情况下 Am Call 不会 early exercise: 考虑欧式期权Put-Call Parity, 有  $C_{eu} = P_{eu} + S - Ke^{-r(T-t)} > P + S - K > S - K$ , 从而  $C_{am} \geq C_{eu} > S - K$ .

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.

美式期权没有Put Call Parity, 没有显式定价公式(Perpetual American Option有).

### Exotic

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.

Autocall Hedging: The Vega of autocall peaks somewhere between KI and spot, with vega at KI only about 20% peak vega. When peak vega is hit, need to hedge skew exposure (sell vol around peak vega and buy vol around KI). Spot near KO, hedge with buying tight call spreads (digital payoff). Spot near KI, buy back all hedges.

### Bond and Rates

$n$ 年零息利率 spot rate / zero rate: 此时投入本金, 所有利息与本金都在 $n$ 年后支付的收益率.

平价收益率(par yield): 债券每期支付券息 $c$ , 所有现金流(每期的 $c$ 与到期的本金加 $c$ )用零息利率贴现等于债券面值(100).

若 $d$ 为到期时\$1贴现值,  $A$ 为年金(annuity, 每个付息日付\$1)现金流贴现值,  $m$ 为每年付息次数, 则 $100 = A \frac{c}{m} + 100d$ .

债券收益率 $y$ : 所有现金流用 $y$ 进行贴现, 现值等于债券市场价格.

久期 $D = -\frac{1}{B} \frac{\partial B}{\partial y}$ , 其中 $B$ 为债券价格,  $y$ 为收益率. 刻画了债券价格关于收益率的敏感度:  $\Delta B \simeq -DB\Delta y$ .

Zero Coupon Bond的久期等于它的maturity.

### VaR

$N$  day  $C\%$  VaR has  $VaR(N, C) \simeq \sigma_v z_C \sqrt{\frac{N}{252}} V(0)$ , where  $C$  is the confidence level,  $\sigma_v$  is the annualized std of return,  $z_C$  is the z-score i.e,  $\mathbb{P}(Z \leq z_C) = C$  for std norm  $Z$ .  $V(0)$  is the current value of portfolio.

## 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 $r(A) = m$ , 即 $A$ 行满秩, 则必存在秩等于 $m$ 的  $n \times m$ 矩阵 $C$ 使 $AC = I_m$ .

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

## 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.
- Cayley-Hamilton Thm.  $f(x)$ 为矩阵 $A$ 的特征多项式, 则 $f(A) = O$ .
- 若 $A$ 特征值为 $\lambda_1, \dots, \lambda_n$ , 则 $f(A)$ 的特征值为 $f(\lambda_1), \dots, f(\lambda_n)$ . 若 $A$ 适合多项式 $g(x)$ ,则 $A$ 的任意特征值适合 $g(x)$ .
- 特征值降阶公式.  $A, B$ 是 $m \times n$ 与 $n \times m$ 矩阵,  $m \geq n$ .则  $|\lambda I_m - AB| = \lambda^{m-n} |\lambda I_n - BA|$ .
- 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; Jordan块都是一阶矩阵; 有完全的特征向量系 (即 $m$ 重特征值有 $m$ 个线性无关特征向量).
- Jordan标准型求法  
 $\lambda I - A$ 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.
- $J = J_r(\lambda)$ 为Jordan块.则 $f(J)$ 主对角线上均为 $f(\lambda)$ , 上次对角线上均为 $f'(\lambda)$ , 距离主对角线 $j$ 位的对角线上均为 $\frac{1}{j!} f^{(j)}(\lambda)$ . 对 $\lambda \neq 0$ ,  $J_r(\lambda)^m$ 的Jordan标准型为 $J_r(\lambda^m)$ .
- $A^k$ 求法: 先求出Jordan标准型 $J$ , 根据 $AP = PJ$ 解方程得 $P$ , 求 $J^k$ 再乘过渡矩阵即可.

## 4.3 二次型

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

余子式法求逆阵:  $A$  是非异阵, 有  $A^{-1} = \frac{1}{\det A} C^T$ . 其中  $C_{ij}$  是  $a_{ij}$  的代数余子式 (有  $\pm 1$ ).

初等变换法求逆阵:  $A$  是非异阵, 作  $n \times 2n$  矩阵  $(A, I_n)$ , 实施初等行变换, 将左侧化为  $I_n$ , 此时右侧矩阵即为  $A^{-1}$ .

$A$  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

等价量代换 在  $x \rightarrow 0$  时,

$$(1+ax)^b \sim 1+abx, \cos x \sim 1 - \frac{1}{2}x^2, x - \sin x \sim \frac{1}{6}x^3, x - \tan x \sim -\frac{1}{3}x^3, \tan x - \sin x \sim \frac{1}{2}x^3,$$

$$\log_a(1+x) \sim \frac{x}{\ln a}, a^x - 1 \sim x \ln a.$$

常用导数

$$(\tan x)' = \sec^2 x \quad (\cot x)' = -\csc^2 x \quad (\sec x)' = \tan x \sec x \quad (\csc x)' = -\cot x \csc x$$

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}} \quad (\arccos x)' = -\frac{1}{\sqrt{1-x^2}}, x \in (-1, 1)$$

$$(\arctan x)' = \frac{1}{1+x^2} \quad (\operatorname{arccot} x)' = -\frac{1}{1+x^2}$$

$$(\log_a x)' = \frac{1}{x \ln a} \quad (a^x)' = a^x \ln a$$

常用高阶导数

$$(x^k)^{(n)} = k(k-1)\cdots(k-n+1)x^{k-n} \quad (n \leq k),$$

$$(a^x)^{(n)} = a^x (\ln a)^n, (\ln x)^{(n)} = (-1)^{n-1} (n-1)! x^{-n},$$

$$(\sin x)^{(n)} = \sin\left(x + \frac{n\pi}{2}\right), (\cos x)^{(n)} = \cos\left(x + \frac{n\pi}{2}\right)$$

$$\text{分部积分法 } \int u(x)v'(x)dx = u(x)v(x) - \int v(x)du(x) = uv - \int v(x)u'(x)dx$$

按照  $u$  在左,  $v$  在右的次序: 反对幂指三

常用变量代换

$$\text{含有 } \sqrt[n]{\frac{ax+b}{cx+d}} \text{ 的积分, 令 } t = \sqrt[n]{\frac{ax+b}{cx+d}};$$

$$\text{含有 } \sqrt{x^2+a^2} \text{ 的积分, 令 } x = a \tan t \text{ 或 } x = a \cot t;$$

$$\text{含有 } \sqrt{x^2-a^2} \text{ 的积分, 令 } x = a \sec t \text{ 或 } x = a \csc t;$$

$$\text{含有 } \sqrt{a^2-x^2} \text{ 的积分, 令 } x = a \sin t \text{ 或 } x = a \cos t;$$

$$\text{三角有理函数积分 } \int R(\sin x, \cos x)dx$$

$$R(-\sin x, \cos x) = -R(\sin x, \cos x), \text{ 令 } t = \cos x$$

$$R(\sin x, -\cos x) = -R(\sin x, \cos x), \text{ 令 } t = \sin x$$

$$R(-\sin x, -\cos x) = R(\sin x, \cos x), \text{ 令 } t = \tan x$$

## 常用积分

$$\int \tan x \, dx = -\ln |\cos x| + C, \int \cot x \, dx = \ln |\sin x| + C$$

$$\int \csc x \, dx = \ln |\csc x - \cot x| + C, \int \sec x \, dx = \ln |\sec x + \tan x| + C$$

$$\int a^x \, dx = \frac{1}{\ln a} a^x + C, \int \ln x \, dx = x \ln x - x + C$$

$$\int \frac{1}{\sqrt{a^2 - x^2}} \, dx = \arcsin \frac{x}{a} + C, \int \frac{1}{\sqrt{x^2 \pm a^2}} \, dx = \ln \left| x + \sqrt{x^2 \pm a^2} \right| + C$$

$$\int \frac{1}{a^2 + x^2} \, dx = \frac{1}{a} \arctan \frac{x}{a} + C, \int \frac{1}{a^2 - x^2} \, dx = \frac{1}{2a} \ln \left| \frac{a+x}{a-x} \right| + C$$

$$\int \sin^2 x \, dx = \frac{1}{2}(x - \sin x \cos x) + C, \int \cos^2 x \, dx = \frac{1}{2}(x + \sin x \cos x) + C$$

$$\int \tan^2(x) \, dx = \tan x - x + C, \int \cot^2(x) \, dx = -\cot x - x + C$$

有理函数不定积分: 化为简单分式  $\frac{A}{(x-a)^k}, \frac{Ax+B}{((x-a)^2+b^2)^k}$  的和, 然后  $\int \frac{Ax+B}{((x-a)^2+b^2)^k} dx =$

$$\int \frac{A(x-a)}{((x-a)^2+b^2)^k} dx + \int \frac{B+aA}{((x-a)^2+b^2)^k} dx, \int \frac{1}{((x-a)^2+b^2)^k} dx \stackrel{t=\frac{x-a}{b}}{=} \frac{1}{b^{2k-1}} \int \frac{1}{(t^2+1)^k} dt,$$

$$\int \frac{1}{(x^2+1)^2} dx = \frac{\arctan(x)}{2} + \frac{x}{2+2x^2} + C$$

## 级数

$$S(x) = \sum_{k=0}^{\infty} a_k (x-a)^k \text{ 收敛域 } R = 1/\limsup_{k \rightarrow \infty} |a_k|^{1/k}.$$

$$f(x) \text{ 任意阶可导, Taylor展开 } T(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k \text{ 收敛域 } R = \frac{1}{e} \liminf_{k \rightarrow \infty} \frac{k}{|f^{(k)}(a)|^{1/k}}.$$

$$\text{Stirling公式. } n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \quad (n \rightarrow \infty)$$

$$\text{Viète公式. } \prod_{n=1}^{\infty} \cos\left(\frac{x}{2^n}\right) = \frac{\sin x}{x}$$

$$\text{Wallis公式. } \prod_{n=1}^{\infty} \left(1 - \frac{1}{(2n)^2}\right) = \frac{2}{\pi}$$

## 5.2 多元

## 多元微分

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

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

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

Hessian 矩阵正定的极值点是极小值点, 负定是极大值点.

多元函数  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  二阶 Taylor 展开:  $f(x) = f(a) + Df(a)(x-a) + \frac{1}{2}(x-a)^T D^2 f(a)(x-a)$ , 其中  $D$  为求梯度,  $D^2$  为 Hessian 矩阵.

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  $\begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = -\frac{\partial(F, G)}{\partial(u, v)}^{-1} \frac{\partial(F, G)}{\partial(x, y)}$ .

Lagrange 乘数法: 目标函数  $f(x_1, \dots, x_n)$ , 约束条件  $g_i(x_1, \dots, x_n), i = 1 : m$ .  $f, g_i$  有连续偏导且  $g$  Jacobi 矩阵满秩. 构造 Lagrange 函数  $L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) = f(x_1, \dots, x_n) - \sum_{i=1}^m \lambda_i g_i(x_1, \dots, x_n)$  则条件极值点一定为  $\frac{\partial L}{\partial x_k} = 0, g_l = 0 (k = 1 : n, l = 1 : m)$  的解对应的  $x$  中. 若  $(\frac{\partial^2 L}{\partial x_k \partial x_l})_{n \times n}$  正定, 为条件极小值点. 若负定, 为条件极大值点.

### 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  $\frac{\partial(x_1, x_2, \dots, x_n)}{\partial(r, \varphi_1, \dots, \varphi_{n-1})} = r^{n-1} \sin^{n-2} \varphi_1 \sin^{n-3} \varphi_2 \cdots \sin \varphi_{n-2}$ .

### 曲线与曲面积分

Green公式.  $P, Q$  在  $D$  上有连续偏导, 则  $\int_{\partial D} P dx + Q dy = \iint_D (Q_x - P_y) dx dy$ . 其中沿着边界方向时区域始终在边界左侧.

Gauss公式.  $P, Q, R$  在  $\Omega$  上有连续偏导, 则  $\iint_{\partial \Omega} P dy dz + Q dz dx + R dx dy = \iiint_{\Omega} (P_x + Q_y + R_z) dx dy dz$ . 其中  $\partial \Omega$  定向外侧.

Stokes公式.  $P, Q, R$  在  $\Sigma$  与  $\partial \Sigma$  上有连续偏导, 则  $\int_{\partial \Sigma} P dx + Q dy + R dz = \iint_{\Sigma} (R_y - Q_z) dy dz + (P_z - R_x) dz dx + (Q_x - P_y) dx dy$ . 其中  $\partial \Sigma$  方向与曲面选定一侧的法向量呈右手关系.

注. (1) 需要连续偏导, 不可导点要单独处理. 需要闭合, 不闭合可以补成闭合. 需要连通, 环状区域可写成两个连通区域的差. (2) 由Green公式, 若  $Q_x = P_y$ , 则  $\int_L P dx + Q dy$  与路径无关.

### Euler积分

Beta 函数  $B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx$

在定义域  $(0, +\infty) \times (0, +\infty)$  连续.  $B(p, q) = B(q, p)$ . 递推公式  $B(p, q) = \frac{q-1}{p+q-1} B(p, q-1)$ .



其他表示:  $B(p, q) = 2 \int_0^{\pi/2} \cos^{2p-1} \varphi \sin^{2q-1} \varphi d\varphi = \int_0^1 \frac{t^{p-1} + t^{q-1}}{(1+t)^{p+q}} dt.$

Gamma函数  $\Gamma(s) = \int_0^{+\infty} x^{s-1} e^{-x} dx$

在定义域  $s > 0$  连续且  $n$  阶可导,  $\Gamma^{(n)}(s) = \int_0^{+\infty} x^{s-1} e^{-x} (\ln x)^n dx$ . 递推公式  $\Gamma(s+1) = s\Gamma(s)$ .

其他表示: 令  $x = t^2$ , 则  $\Gamma(s) = 2 \int_0^{+\infty} t^{2s-1} e^{-t^2} dt.$

相关公式

(1)  $B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$  (2) Legendre公式  $\Gamma(s)\Gamma(s+\frac{1}{2}) = \frac{\sqrt{\pi}}{2^{2s-1}} \Gamma(2s)$

(3) 余元公式  $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin(\pi s)}$  (4) 一些取值  $B(\frac{1}{2}, \frac{1}{2}) = \pi, \Gamma(n+1) = n!, \Gamma(\frac{1}{2}) = \sqrt{\pi}.$

## 6 常微

### 1. 分离变量

(1) 齐次方程  $\frac{dx}{dt} = f(\frac{x}{t})$ . 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  $x = e^{\int P(t)dt} (\int e^{-\int P(t)dt} Q(t) dt + C).$

### 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  $x^{(n)} + a_1x^{(n-1)} + \cdots + a_{n-1}x^{(1)} + a_nx = 0$

特征方程  $\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n = 0$ .  $\lambda_1, \dots, \lambda_s$  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  $x^{(n)} + a_1x^{(n-1)} + \cdots + a_{n-1}x^{(1)} + a_nx = f(t)$

解为  $x(t) = x_1(t) + x^*(t)$ , 其中  $x_1(t)$  为  $x^{(n)} + a_1x^{(n-1)} + \cdots + a_{n-1}x^{(1)} + a_nx = 0$  的通解,  $x^*(t)$  为  $x^{(n)} + a_1x^{(n-1)} + \cdots + a_{n-1}x^{(1)} + a_nx = f(t)$  的特解.

$x^*(t) = \int_0^t K(t-s)f(s)ds$ , 其中  $K(t)$  为  $x^{(n)} + a_1x^{(n-1)} + \cdots + a_{n-1}x^{(1)} + a_nx = 0$  在初值条件  $K(0) = K'(0) = \cdots = K^{(n-1)}(0) = 0, K^{(n)}(0) = 1$  下的解.

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

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

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  $\vec{x}(t) = Pe^{Jt}\vec{c}$ , 其中对  $\lambda_i$  的 Jordan 块  $J_i, e^{J_i t}$  第一行为  $e^{\lambda_i t}, te^{\lambda_i t}, \dots, \frac{t^{n_i-1}}{(n_i-1)!}e^{\lambda_i t}$ . 记  $\vec{p}_i$  为  $P$  的第  $i$  个列向量, 则有

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

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

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

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

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

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

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

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

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

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

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

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

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

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

### 7.3 采样

均匀采正态: Box-Muller变换.  $U_1, U_2$ 是独立的 $[0, 1]$ 均匀分布, 则 $Z_0 = \sqrt{-2 \ln U_1} \cos(2\pi U_2)$ ,  $Z_1 = \sqrt{-2 \ln U_1} \sin(2\pi U_2)$  是标准正态分布随机变量.

逆变换采样:  $[0, 1]$ 均匀分布取 $u$ , 目标分布CDF为 $F$ ,  $F^{-1}(u)$ 可采样出目标分布.

Acceptance-Rejection Sampling. 由给定pdf  $f(x)$ , 产生服从该分布的样本. 需要建议分布 $G$  (pdf  $g(y)$  已知, 通常用均匀/正态, 选哪个取决于定义域)产生候选样本. 计算 $c$ 为满足 $cg(x) \geq f(x)$ 的最小值. 从 $G$ 抽样出样本 $Y$ , 从均匀分布 $U(0, 1)$ 抽样出样本 $U$ . 如果 $U \leq \frac{f(Y)}{cg(Y)}$ 则接受 $Y$ , 否则拒绝.