向量微积分 $\frac{df}{d(s,t)}$ 维度

多元函数：可微求可导

全微分

多元函数链式法则 $\frac{df}{d\vec{x}}\cdot\frac{d\vec{x}}{dt}$

梯度：$\nabla_x f = \frac{df}{d\vec{x}} = \left[\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n}\right]$

向量值函数：$f: R^n \to R^m$ $f(x) \in R^m$

$$\nabla_{\vec{x}} f = \begin{bmatrix} \frac{\partial f_1(x)}{x_1} & \cdots & \frac{\partial f_1(x)}{x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m(x)}{x_1} & \cdots & \frac{\partial f_m(x)}{x_n} \end{bmatrix}$$ 雅可比矩阵

$\nabla_{B_{p\times q}} A_{m\times n} = [J_{ijkl} = \partial A_{ij}/\partial B_{kl}]_{m\times n\times p\times q}$

有用的恒等式 $\frac{\partial [f(X)]^T}{\partial X} = \left(\frac{\partial f(X)}{\partial X}\right)^T$ $\frac{\partial a^T X b}{\partial X} = ab^T$

$\frac{\partial[tr(f(X))]}{\partial X} = tr\left(\frac{\partial f(X)}{\partial X}\right)$ $\frac{\partial x^T B x}{\partial x} = x^T(B+B^T)$

$\frac{\partial x^T a}{\partial x} = a^T$ $\frac{\partial}{\partial s}(\vec{x}-As)^T W(\vec{x}-As)$

$\frac{\partial a^T x}{\partial x} = a^T$

泰勒级数 $f(x) \approx f(a) + \nabla f(a)^T(x-a) + \frac{1}{2}(x-a)^T H_f(a)(x-a) + \dots$

Optimization $\max\{f_1(x), f_2(x)\}$ 凸函数非负加权和

凸集，保凸运算：交集运算，仿射函数。

下水平集：$C_\alpha = \{x \in \mathrm{dom} f \mid f(x) \le \alpha\}$

上镜图 $\{(x,t) \mid x \in \mathrm{dom} f, f(x) \le t\}$

凸函数充要条件 | 图像在切线上方
| 海森矩阵半正定

原问题 → 对偶问题 $g(\lambda,\nu) = \inf_x L(x,\lambda,\nu)$

可行域 对偶可行域

弱对偶性：对偶问题得到原问题最优值 p^* 的最大下界

强对偶性：$d^* = p^*$

slater：$\exists x \in \mathrm{relint} D$ s.t. $f_i(x) < 0$

即存在约束严格成立（仿射约束除外）

KKT：① 原问题约束
② $\nabla_x L(x, \lambda, \nu) = 0$
③ $\lambda g(x) = 0$ 互补松弛
④ $\lambda \ge 0$ 对偶问题约束

Newton $x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$ 求根

$x_n = x_{n-1} - \frac{f'(x_{n-1})}{f''(x_{n-1})}$ 求极值

算法：O Ω Θ

贪心

分治：$n\log n$

①将问题分解成几部分

②递归地解决每个部分

③将子问题的解决方案合并

多元正态分布：$p(x_1, x_2, \dots x_n) = p(x) = \frac{1}{(2\pi)^{\frac{n}{2}}|B|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-a)^T B^{-1}(x-a)}$

$ax+by \sim N(a u_x + b u_y, a^2\Sigma_x + b^2\Sigma_y)$

$Ax \sim N(Au_x, A\Sigma_x A^T)$

伯努利共轭：$x \sim Ber(u)$

$p(x|u) = u^x(1-u)^{1-x}$

$u \sim Beta(\alpha, \beta)$

$p(u|\hat{x}) \sim p(\hat{x}|u)\cdot p(u) = Beta(\alpha+\hat{x}, \beta+1-\hat{x})$

$a\int_a^{+\infty} p(x)dx \le \int_a^{+\infty} p(x)x dx \le \int_{-\infty}^{+\infty} p(x)x\,dx = E(x).$

$P\{x \ge a\} \le \frac{E(x)}{a}$

$P_r((x-E(x))^2 \ge a^2) \le \frac{D(x)}{a^2} \Rightarrow P_r(|x-E(x)| \ge a) \le \frac{D(x)}{a^2}$

随机投影定理：

采样.

基本采样算法（逆采样算法）

拒绝采样法：

重要性采样

正态分布生成.

① $u_1[0,1]$ $u_2[0,1]$ BOX-Muller

$\sqrt{-2\ln u_1}\cos(2\pi u_2)$ $\sqrt{-2\ln u_1}\sin(2\pi u_2)$

② $x \in [-1,1]$ $y \in [-1,1]$ $\Rightarrow x^2+y^2 \le 1$ Marsaglia polar

$s = x^2+y^2 = r^2 \sim [0,1]$

$\sqrt{\frac{-2\ln s}{s}}\cdot x$ $\sqrt{\frac{-2\ln s}{s}}\cdot y$

高维球面均匀分布：

由正态分布归一化生成.