非线性无约束优化问题

$$x^{k+1} = x^k + \alpha^k d^k$$

α^k：步长

d^k：搜索方向

精确线搜索：$\alpha^k = \arg\min_{\alpha>0} \phi(\alpha) := f(x^k + \alpha^k d^k)$

非精确线搜索：

Armijo 准则： $f(x^k + \alpha d^k) \le f(x^k) + c_1 \alpha \langle \nabla f(x^k), d^k \rangle$

梯度下降：$f(x^k - \alpha \nabla f(x^k)) \le f(x^k) - c_1 \alpha \|\nabla f(x^k)\|^2$

回退法：从大到小搜索步长

Goldstein 准则：
$$\begin{cases} f(x^k + \alpha d^k) \le f(x^k) + c_1 \alpha \langle \nabla f(x^k), d^k \rangle \\ f(x^k + \alpha d^k) \ge f(x^k) + (1-c_1) \alpha \langle \nabla f(x^k), d^k \rangle \end{cases}$$

Wolfe 准则：
$$\begin{cases} f(x^k + \alpha d^k) \le f(x^k) + c_1 \alpha \langle \nabla f(x^k), d^k \rangle \\ \langle \nabla f(x^k + \alpha d^k), d^k \rangle \ge c_2 \langle \nabla f(x^k), d^k \rangle \end{cases}$$

计算的复杂度

Flop 计数：

① flop：两个浮点数之间的一次四则运算

② 算法的复杂度：计算所需的 flop 数量.

a. 由问题规模的多项式函数表示，

b. 一般会简化表示，仅保留最高阶项　　$n^3 + n^2 \to O(n^3)$

③ 仅为复杂度的粗略估计，不能准确预测时间

E.g. ① 内积 x^Ty $\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \Rightarrow x^Ty = \sum_{i=1}^{n} x_i y_i$

n 个乘法，n-1 个加法 $\longrightarrow$ 2n-1 flops.

② 求和 $x+y$ n 个加法 $\longrightarrow$ n flops

③ 标量乘法 αx $\alpha \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{bmatrix}$ n 个乘法 $\longrightarrow$ n flops.

④ 矩阵乘法 $y = Ax$ $A \in \mathbb{R}^{m \times n}$ $\longrightarrow$ $m(2n-1)$ flops

$m=2$ $\begin{bmatrix} - A_1 - \\ - A_2 - \end{bmatrix} \begin{bmatrix} | \\ x \\ | \end{bmatrix} = \begin{bmatrix} A_1 x \\ A_2 x \end{bmatrix}$

$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix}$

(4.1) 稀疏矩阵：N 个非零元素，初始化 $y = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$

$y = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

a. N 个乘法：$A_{ij} \cdot x_j$ b. N 个加法：$A_{ij} x_j$ 加到对应 y 的位置

$\longrightarrow$ 2N flops

$A = \begin{pmatrix} 3 & 0 & 0 & 5 \\ 0 & 0 & -2 & 0 \\ 0 & 7 & 0 & 0 \end{pmatrix}$ $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ $y = \begin{pmatrix} 0 + 3x_1 + 5x_4 \\ 0 + -2x_3 \\ 0 + 7x_2 \end{pmatrix}$

(4.2) 低秩矩阵 $A = UV^T$ $U \in \mathbb{R}^{m \times p}$ $V \in \mathbb{R}^{n \times p}$ $m >> p$ $n >> p$

$y = Ax = UV^Tx = U(V^Tx)$

$\left.\begin{array}{l} z = V^Tx : \ p(2n-1) \approx 2pn \\ y = Uz : \ m(2p-1) \approx 2pm \end{array}\right\} \longrightarrow 2p(n+m)$ flops.

⑤ 矩阵乘法：$C = AB$ $A \in \mathbb{R}^{m \times n}$ $B \in \mathbb{R}^{n \times p}$

$AB = \underbrace{\begin{bmatrix} A \end{bmatrix}}_{A^{m \times n}} \underbrace{\begin{bmatrix} | & | & & | \\ B_1 & B_2 & \cdots & B_p \\ | & | & & | \end{bmatrix}}_{B^{n \times p}} = \begin{bmatrix} AB_1 & AB_2 & \cdots & AB_p \end{bmatrix} \longrightarrow p\,m(2n-1)$ flops $\approx 2mnp$

A: 1000×100

B: 100×2000

C: 2000×10

A B C

$A \cdot B : 2(1000 \times 100 \times 2000) \to \mathbb{R}^{1000 \times 2000}$

$(AB)C : 2 \cdot (1000 \times 10)$

$4 \times 10^8 + 2 \times 10^4$

$A(B \cdot C)$ $B \cdot C : 2 \times 100 \times 2000 \times 10 \to \mathbb{R}^{100 \times 10}$

$A \cdot (B \cdot C) : 2 \times 1000 \times 100 \times 10$

$4 \times 10^6 + 2 \times 10^6$

$$C = \begin{bmatrix} C_{11} & C_{12} & & \\ C_{21} & \ddots & & \\ & & \ddots & \\ & & & C_{mm} \end{bmatrix}$$

(5.1) 对称矩阵 $C = AA^T$ $A \in \mathbb{R}^{m\times n}$

一共有 $\frac{1}{2}m(m+1)$ 个元素需计算

$c_{ij} = A_i^T A_j$: A的第i行与第j行相乘: $2n-1$ flops

$\to \frac{1}{2}m(m+1)(2n-1)$ flops

$\approx m^2 n$ flops

求解 $Ax=b$ 的计算量

$\|Ax-b\|^2 = 0$

$\Leftrightarrow Ax=b$

$\min\limits_x \|Ax-b\|^2$

$f(x) = x^TA^TAx - 2b^TAx + b^Tb$

$\Rightarrow \nabla f(x) = 2A^TAx - 2b^TA = 0$

$\Rightarrow x = (A^TA)^{-1}A^Tb$

$A \in \mathbb{R}^{n\times n}$

① 最简单的方法: $x = A^{-1}b$ ① A是非奇异的 nonsingular

a. 求 A^{-1}: 高斯-约旦消元法

$[A|I] \Rightarrow [I|A^{-1}]$

第k步: 把 a_{kk} 之外的元素化为0 $\to 2n\cdot(n-1)$ 每个元素 共n-1个元素

$$[\underbrace{\overbrace{\cdots\ a_{kk}\ \cdots}^{n个}}_{n-1个}]$$

把 a_{kk} 化为1 n个flops

$$\begin{bmatrix} a_{kk} & 3 & \\ \cdots & 1 & \cdots \\ \cdots & 3 & \cdots \\ \cdots\ a_{kk}\text{-}3 & 0 & \cdots \end{bmatrix}$$

① αy n flops

② $x+y$ n flops

共进行n步: $n\left(2n(n-1)+n\right) = 2n^2(n-1)+n^2 = 2n^3-n^2$

$\approx 2n^3$ flops

b. $A^{-1}b$: $2n(n-1)$ flops

$\to 2n^3+n^2-2n$ flops

$\approx 2n^3$ flops

② 带结构的A:

(1) 对角矩阵 $x = A^{-1}b$ n个除法 + $n(2n-1)$ flops

$2n^2$ flops

$$A: \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} \frac{1}{a_{11}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{a_{nn}} \end{bmatrix}$$

(2) 下三角矩阵

$$\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$x_1 = b_1 / a_{11}$ 1

$x_2 = (b_2 - a_{21} x_1)/a_{22}$ 3

$x_3 = (b_3 - a_{31}x_1 - a_{32}x_2)/a_{33}$ 5

$\vdots$

$x_k = (b_k - a_{k1}x_1 - \cdots a_{kk-1}x_{k-1})/a_{kk}$ $2k-1$

一共：$\frac{1}{2}(1+2n-1)n = n^2$ flops

(3) 上三角矩阵：n^2 flops. 同上

(4) 正交阵：$AA^T = I$

$AA^T = I \Rightarrow A^{-1}AA^T = A^{-1}I = A^T = A^{-1}$

$Ax = b \Rightarrow x = A^T b$: $\rightarrow n(2n-1)$ flops

(5) 共轭梯度法：要求 A 为正定矩阵

共轭：若 u, v 是关于 A 共轭，那么 $\langle u, Av \rangle = 0$

算法：共轭梯度法

初始化：x^0：初始点

1. $r^0 = b - Ax^0$
2. $d^0 = r^0$
3. while $\|r^{k+1}\| > tol$
4. 步长：$\alpha_k = (r^k)^T r^k / (d^k)^T A d^k$
5. 迭代点：$x^{k+1} = x^k + \alpha_k d^k$
6. $r^{k+1} = r^k - \alpha_k A d^k$
7. $\beta_k = (r^{k+1})^T r^{k+1} / (r^k)^T r^k$
8. $d^{k+1} = r^{k+1} + \beta_k d^k$
9. End while

$O(\sqrt{\kappa(A)}\, n^2)$ flops

理论最大 $O(n^3)$ flops

$\min_x \frac{1}{2}x^TAx - bx$

$\nabla f(x) = Ax - b = 0$

$Ax = b$

共轭梯度法的收敛速率：Q-线性收敛

$$\|x_k - x^*\|_A \le 2\left(\frac{\sqrt{K}-1}{\sqrt{K}+1}\right)^k \|x_0 - x^*\|_A$$

k → 迭代数

$K = K(A)$ 是 A 的条件数

梯度下降法：$\|x_k - x^*\|_A \le 2\left(\frac{\sqrt{K}-1}{\sqrt{K}+1}\right) \|x_0 - x^*\|_A$

(6) LU 分解：非奇异矩阵

$A = PLU$

P：置换矩阵，L 下三角矩阵，U 上三角矩阵

$Ax = b \Rightarrow A_1A_2\cdots A_k x = b \qquad x = A_k^{-1}\cdots A_2^{-1}A_1^{-1}b$

P 置换矩阵

$P^TP = PP^T = I_n$

$P \in \{0,1\}^{n\times n}$

a. LU 分解：$\frac{2}{3}n^3$ flops.

b. 置换矩阵 $Pz_1 = b \Rightarrow 0$ flops

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \Rightarrow \begin{bmatrix} z_2 \\ z_1 \\ z_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

只需做重新排序.

c. 求解 z_2：$Lz_2 = z_1 \rightarrow n^2$ flops

d. 求解 x：$Ux = z_2 \rightarrow n^2$ flops

$\longrightarrow \frac{2}{3}n^3 + 2n^2$ flops

$\approx \frac{2}{3}n^3$ flops.