conjugate method

$$Ax=b \iff \min \frac{1}{2}\|Ax-b\|_2^2 \iff \min \frac{1}{2}x^TA^TAx - b^TAx + \frac{1}{2}b^Tb$$

$$\iff \min \frac{1}{2}x^TA^TAx - b^TAx \quad (QP)$$

① steepest descent : too slow

② Newton's method : too expensive

CG 的 n 个独立最小.

$$\min \frac{1}{2}x^TQx - b^Tx, \quad Q \in S_+^n, \ x \in R^n$$

def: Q-conjugate, $x, y \neq 0$, if x, y 为 Q-conjugate then $x^TQy = 0$

find $d_1, d_2 \ldots d_n$ 相互共轭

Thero 1.3. lemma $Q \in S_+^n$ $\{d_i\}_{i=1}^{k}$ are Q-conjugate, then they are linearly independent. if $d_k = \sum_{i=1}^{k-1} \alpha_i d_i$

$$0 < d_k^TQd_k = d_k^TQ\left(\sum_{i=1}^{k-1}\alpha_i d_i\right) = \sum_{i=1}^{k-1}\alpha_i d_k^TQd_i \equiv 0$$ 矛盾

1.4. $x \in R^n$, $\exists \{\alpha_i\}_{i=0}^{n-1}$, $x = \sum_{i=0} \alpha_i d_i$ $\{d_i\}_{i=2}^{n-1}$

$$\min \frac{1}{2}x^TQx - b^Tx$$

$$\Rightarrow \frac{1}{2}\left(\sum_i \alpha_i d_i^T\right)Q\left(\sum_j \alpha_j d_j\right) - b^T\left(\sum_i \alpha_i d_i\right)$$

$$= \sum_i \frac{1}{2}\alpha_i^2 d_i^TQd_i - \alpha_i b^Td_i$$

Gram-schimit 正交化

$$\frac{\partial L}{\partial \alpha_i} = \alpha_i d_i^TQd_i - b^Td_i = 0 \Rightarrow \alpha_i = \frac{b^Td_i}{d_i^TQd_i}$$

$$\alpha_k = \frac{-g_k^T d_k}{d_k^T Q d_k^T} = \frac{-(Q_k - b)^T d_k}{d_k^T \alpha d_k}$$

$$g_k = Q x_k - b$$

1.5 conjugate direction theorem.

Then: $d_0, \cdots d_{n-1}$ ... Q-conjugate ; $x_0 \in R^n$, then $X_{k+1} = X_k + \alpha_k d_k$

after n steps, $X_n = x^*$, $(Qx^* = b)$

pf: $\{d_i\}$ independent , $x^* - x_0 = \sum_{i=0}^{n-1} \alpha_i d_i$ for some $\{\alpha_i\}$

$$X_1 = X_0 + \alpha_0 d_0$$

$$x_k = x_0 + \alpha_0 d_0 + \alpha_1 d_1 + \cdots + \alpha_{k-1} d_{k-1}$$

$$X_n = x_0 + \alpha_0 d_0 + \cdots, + \alpha_{n-1} d_{n-1}$$

之要 $\alpha_k = \dfrac{-g_k \, d_k}{d_k^T Q d_k}$ $\Rightarrow$ $x^* - x_0 = \sum_{i=0}^{k-1} \alpha_i d_i$

$$x^k - x_0 = \sum_{i=0}^{k-1} \alpha_i d_i$$

$$\Rightarrow d_k^T Q (x^* - x_0) = \alpha_k d_k^T Q d_k$$

$\alpha_k = \dfrac{d_k^T Q (x^* - x_0)}{d_k^T Q d_k}$ $\Rightarrow$ $\alpha_k = \dfrac{-d_k^T g_k}{d_k^T Q d_k}$

$$d_k^T Q (x^* - x_0) = d_k^T Q (x^* - x_k + x_k - x_0) = d_k^T Q (x^* - x_k) = -d_k^T (Q x_k - b)$$

$$= -d_k^T g_k$$

例子: $d_0 = -g_0 = b - Qx_0$

repeat: $g_k = Qx_k - b$

$$\alpha_k = \frac{-g_k^T d_k}{d_k^T Q d_k}$$

$$x_{k+1} = x_k + \alpha d_k$$

$$d_{k+1} = -g_{k+1} + \beta_k d_k$$

$$\beta_k = \frac{g_{k+1}^T Q d_k}{d_k^T Q d_k}$$