二、最速下降法

最速下降法基于梯度，我们都知道梯度方向代表函数增长最快的方向，那么负梯度方向就是下降最快的方向，也就是最陡峭的方向，我们沿该方向更新参数，效果一定是最好的。因此，对我们的无约束最优化公式 $min\phi(x) = \frac{1}{2}(Ax, x) - (b, x)$ 对x求导的得到 $\frac{d(\frac{1}{2}x^TAx - b^Tx)}{dx} = Ax - b$，那么负梯度方向就是 $r = -\nabla\varphi(x) = b - Ax$ 。

步长：寻求步长就是希望一步跨到最优解，那么 $\varphi(x + \alpha x) = min$ ，带入min中求导，令其导数等于0，即可算出 α。

$$\min \varphi(x) := \frac{1}{2}(Ax, x) - (b, x)$$

下降方向： 负梯度方向

$$r = -\nabla\varphi(x) = -(Ax - b) = b - Ax$$

步长： 寻求步长α，使得 $\varphi(x+\alpha r) = \min$

$$\varphi(x+\alpha r) = \frac{1}{2}\left(A(x+\alpha r),(x+\alpha r)\right) - \left(b,(x+\alpha r)\right)$$

$$\frac{d\varphi(x+\alpha r)}{d\alpha} = \alpha(Ar, r) - (r, r) = 0$$

$$\alpha = \frac{(r,r)}{(Ar,r)}$$

最速下降法

算法流程: 对于给的初始向量x_0

$$r_k = b - Ax_k$$

$$\alpha_k = (r_k, r_k)/(Ar_k, r_k) \qquad k = 0,1,2,\ldots$$

$$x_{k+1} = x_k + \alpha_k r_k$$

收敛性定理:

$$(r_{k+1}, r_k) = 0, \qquad k = 0,1,2,\ldots$$

$$\|x_k - x^*\|_A \le \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^k \|x_0 - x^*\|_A$$

其中 λ_1, λ_n 分别为矩阵A的最大最小特征值，$\|x\|_A = \sqrt{(Ax,x)}$

三、共轭梯度法（CG法）

共轭梯度法同时考量上一步和当前方向

$$\min \varphi(x) := \frac{1}{2}(Ax, x) - (b, x)$$

下降方向： **上步下降方向与当前残量的线性组合**

$$p_k = r_k + \beta_{k-1} p_{k-1}$$

步长： **寻求步长**α**，使得** $\varphi(x + \alpha p) = \min$

$$\alpha = \frac{(r, p)}{(Ap, p)}$$

$$\min \varphi(x) := \frac{1}{2}(Ax, x) - (b, x)$$

对于 $p_k = r_k + \beta_{k-1} p_{k-1}$**，选择** β_{k-1}**，使得**

$$\varphi\left(x_k + \alpha_k\left(r_k + \beta_{k-1} p_{k-1}\right)\right) = \min$$

$$\frac{d\varphi\left(x_k + \alpha_k\left(r_k + \beta p_{k-1}\right)\right)}{d\beta} = \left(\alpha_k^2 p_{k-1}^T A p_{k-1}\right)\beta - \alpha_k\left(p_{k-1}^T r_k - \alpha_k p_{k-1}^T A r_k\right) = 0$$

由于 $p_{k-1}^T r_k = 0$**，**

$$\beta_{k-1} = \frac{p_{k-1}^T A r_k}{p_{k-1}^T A p_{k-1}}$$

知乎 @乔胤博

$$(p_i, r_k) = 0, \quad i = 0, 1, \ldots, k-1$$

$$(r_i, r_k) = 0, \quad i = 0, 1, \ldots, k-1$$

$$(Ap_i, p_k) = 0, \quad i = 0, 1, \ldots, k-1$$

共轭梯度法最多迭代n步即可求得问题的精确解

$$\alpha_k = \frac{(r_k, p_k)}{(Ap_k, p_k)} = \frac{(r_k, r_k + \beta_{k-1} p_{k-1})}{(Ap_k, p_k)} = \frac{(r_k, r_k)}{(Ap_k, p_k)}$$

$$\beta_{k-1} = \frac{p_{k-1}^T A r_k}{p_{k-1}^T A p_{k-1}} = \frac{r_k^T (r_{k-1} - r_k)}{p_{k-1}^T A p_{k-1}} \frac{1}{\alpha_{k-1}} = \frac{(r_k, r_k)}{(r_{k-1}, r_{k-1})}$$

这样我们就得到了共轭梯度法的基本流程，下面给出算法流程。

对于给的初始向量 x_0 ， $r_0 = p_0 = b - Ax_0$

$$\begin{aligned} &\alpha_k = (r_k, r_k)/(Ap_k, p_k) \\ &x_{k+1} = x_k + \alpha_k p_k \\ &r_{k+1} = r_k - \alpha_k A p_k \qquad\qquad k = 0,1,2,\ldots \\ &\beta_k = (r_{k+1}, r_{k+1})/(r_k, r_k) \\ &p_{k+1} = r_{k+1} + \beta_k p_k \end{aligned}$$