《数值分析》10

主要内容：

- 初等变分原理
- 最速下降法

I 方程组问题: $Ax = b$

II 极值问题: $\min\limits_{x \in R^n} \quad f(x) = \frac{1}{2} x^T A x - b^T x$

设$x, y \in R^n$,记 $(x, y) = x^T y$

- $(x, y) = (y, x)$;
- $(tx, y) = t(x, y)$;
- $(x+y, z) = (x, z) + (y, z)$;
- $(x, x) \geq 0$, 且$(x, x) = 0 \Leftrightarrow x = 0$;

设A是 n 阶对称正定阵

- $(Ax, y) = (x, Ay)$; **(对称)**
- $(Ax, x) \geq 0$, 且$(Ax, x) = 0 \Leftrightarrow x = 0$ **(正定)**

定理4.10 设$A=(a_{ij})_{n\times n}$为实**对称正定矩阵**，$x,b\in R^n$

则 x 使二次函数

$$f(x)=\frac{1}{2}(Ax,x)-(b,x)$$

取极小值 $\Leftrightarrow$ x 是线性方程组 $Ax=b$ 的解。

证明: (<=)设 u 是 $Ax=b$的解

$$\Rightarrow Au=b \Rightarrow f(u)=-\frac{1}{2}(Au,u)$$

对任意 $x\in R^n$，只须证明 $f(x)-f(u)\geqslant 0$

$$f(x)-f(u)=\frac{1}{2}(Ax,x)-(b,x)+\frac{1}{2}(Au,u)$$

$$=\frac{1}{2}(A(x-u),(x-u))\geq 0$$

初等变分原理

(=>) 设 u 使 $f(x)$ 取极小值. 取非零向量 $x\in R^n$, 对任意 $t\in R$, 有

$$f(u+tx)=\frac{1}{2}(A(u+tx),u+tx)-(b,u+tx)$$

$$=f(u)+t(Au-b,x)+\frac{t^2}{2}(Ax,x)$$

令 $g(t)=f(u+tx)$, 当 t=0 时, $g(0)=f(u)$达到极小值, 所以 $g'(0)=0$,即

$$(Au-b,x)=0 \quad \rightarrow \quad Au-b=0$$

所以, u 是方程组 $Ax=b$ 的解.

从初值点 $x^{(0)}$ 出发, 以负梯度方向 r 为搜索方向

选择步长 t_0, 使 $x^{(1)} = x^{(0)} + t_0 r$ 为 $f(x)$ 极小值点

在 x 处,梯度方向是 $f(x)$ 增长最快方向

负梯度方向是 $f(x)$ 下降最快方向

梯度: $\nabla f = \mathrm{grad} f(x) = [\, f_{x1},\ f_{x2},\ \cdots,\ f_{xn}\,]^T$

$$\boxed{\nabla f = Ax - b}$$

方向l：$l=[v_1, v_2, \cdots, v_n]^T$，$g(t)=f(x+tl)$

其中，$\|l\|=1$，$x=[x_1, x_2, \cdots, x_n]^T$

方向导数：$g'(0)=f_{x1}v_1+f_{x2}v_2+\cdots,+f_{xn}v_n$

$$\frac{\partial f}{\partial l}=\nabla f\cdot l=\|\nabla f\|\cos<\nabla f, l>$$

l 与∇f 方向一致时, 方向导数取得最大值

$\Rightarrow$ ∇f 是 $f(x)$ 增长最快方向

l 与 ∇f 方向相反时, 方向导数取得最小值

$\boxed{-\nabla f \text{ 是 } f(x) \text{ 下降最快方向}}$

最速下降法

$$g'(0)=f_{x1}v_1+f_{x2}v_2+\cdots,+f_{xn}v_n$$

$$g(t)=f(x+tl)=f(x)+t(Ax-b,l)+\frac{t^2}{2}(Al,l)$$

$$g'(0)=(Ax-b,l)$$

$$\nabla f=Ax-b$$

$$\Downarrow$$

最速下降方向: $r=-\nabla f=b-Ax$

最速下降法

取初值点 $x^{(0)}$,取负梯度方向 $r_0 = b - A\,x^{(0)}$

求点: $x^{(1)} = x^{(0)} + t_0 r_0$使得

$$f(x^{(0)} + t_0 r_0) = \min_{t \in R} f(x^{(0)} + t\, r_0)$$

记 $g(t) = f(x^{(0)}) + t(Ax^{(0)} - b, r_0) + t^2 (Ar_0, r_0)/2$

为选取最佳步长t_0，令

$$g'(t) = (Ax^{(0)} - b, r_0) + t(Ar_0, r_0) = 0$$

求解, 得 $t_0 = (r_0, r_0)/(Ar_0, r_0)$

解对称正定方程组$Ax=b$ 的**最速下降算法**:

第一步: 取初值 $x^{(0)}\in R^{(n)}$, ε>0,计算

$$r_0=b-Ax^{(0)} \quad , k \leftarrow 0;$$

第二步: 计算 $t_k=(r_k,r_k)/(Ar_k,r_k)$

$$x^{(k+1)}=x^{(k)}+t_k r_k ; \qquad r_{k+1}=b-Ax^{(k+1)} \quad ;$$

第三步: $k \leftarrow k+1$, 如果 $\|r_k\|\geq \varepsilon$,转第二步;

否则,输出: $x^{(k)}$, 结束.

例 3.6 通过梯度下降法求解如下优化问题的最优解

$$x^* = \arg\min_{x\in R^2} f(x)=x_1^2+x_2^2,$$ 其中最优解为$(0, 0)^T$.

解： 由于目标函数的梯度为$\nabla f=(2x_1,2x_2)^T$，梯度下降法的迭代格式为

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{pmatrix}=\begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \end{pmatrix}-\alpha\begin{pmatrix} 2x_1^{(k)} \\ 2x_2^{(k)} \end{pmatrix}=(1-2\alpha)\begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \end{pmatrix}.$$

选取初始向量$(x_1^{(0)},x_2^{(0)})^T=(1,3)^T$ 和步长$\alpha=0.1$，则迭代结果为

$$\begin{aligned} x^{(0)} &= (1.0000,3.0000)^T \\ x^{(1)} &= (0.8000,2.4000)^T \\ x^{(2)} &= (0.6400,1.9200)^T \\ x^{(3)} &= (0.5120,1.5360)^T \\ x^{(4)} &= (0.4096,1.2288)^T \\ &\vdots \\ x^{(10)} &= (0.1074,0.3221)^T \end{aligned}.$$

例 3.7 通过梯度下降法求解线性方程组 $Ax=b$，其中

$$A=\begin{pmatrix}2&2\\2&5\end{pmatrix},x=\begin{pmatrix}x_1\\x_2\end{pmatrix},b=\begin{pmatrix}6\\3\end{pmatrix}.$$

解： 方程组的精确解为 $(4,-1)^T$．由定理 3.10 知，线性方程组可以转化为如下二次优化问题求解

$$x^*=\underset{x\in R^2}{\arg\min} f(x)\triangleq\frac{1}{2}(Ax,x)-(b,x)=\frac{1}{2}\begin{pmatrix}x_1 & x_2\end{pmatrix}\begin{pmatrix}2&2\\2&5\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}-\begin{pmatrix}x_1 & x_2\end{pmatrix}\begin{pmatrix}6\\3\end{pmatrix}.$$

下面通过梯度下降法求解二次优化问题．由于目标函数的负梯度方向为

$$r\triangleq-\nabla f(x)=\begin{pmatrix}6\\3\end{pmatrix}-\begin{pmatrix}2&2\\2&5\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix},$$

则梯度下降法迭代格式为

$$\begin{pmatrix}x_1^{(k+1)}\\x_2^{(k+1)}\end{pmatrix}=\begin{pmatrix}x_1^{(k)}\\x_2^{(k)}\end{pmatrix}+\alpha\begin{pmatrix}r_1^{(k)}\\r_2^{(k)}\end{pmatrix}.$$

选取初始向量 $(x_1^{(0)},x_2^{(0)})^T=(0,0)^T$，则第一步迭代搜索方向 $r^{(0)}$ 为

$$r^{(0)}=-\nabla f(x^{(0)})=\begin{pmatrix}6\\3\end{pmatrix}-\begin{pmatrix}2&2\\2&5\end{pmatrix}\begin{pmatrix}0\\0\end{pmatrix}=\begin{pmatrix}6\\3\end{pmatrix},$$

而第一步步长 α_0 为

$$\alpha_0=\frac{(r^{(0)},r^{(0)})}{(Ar^{(0)},r^{(0)})}=\frac{5}{21}.$$

以此类推…

学到了什么?

初等变分原理

最速下降法