(1) LSE

Let $f(x) = (Ax-b)^2 + \lambda \|x\|^2$

Goal: $\underset{x}{\operatorname{argmin}} \left((Ax-b)^2 + \lambda \|x\|^2 \right)$

$$f(x) = (Ax-b)^2 + \lambda \|x\|^2$$

$$= (Ax-b)^T (Ax-b) + \lambda x^T x$$

$$= (x^T A^T - b^T)(Ax-b) + \lambda x^T x$$

$$= x^T A^T A x - x^T A^T b - b^T A x + b^T b + \lambda x^T x$$

$$= x^T A^T A x - 2 x^T A^T b + b^T b + x^T \lambda I x$$

$$= x^T (A^T A + \lambda I) x - 2 x^T A^T b + b^T b$$

$$\frac{\partial f(x)}{\partial x} = 2(A^T A + \lambda I) x - 2A^T b$$

To find the minimum, $\frac{\partial f(x)}{\partial x} = 0$

$$2(A^T A + \lambda I) x = 2A^T b$$

$$x = (A^T A + \lambda I)^{-1} A^T b$$

(2) Steepest descent method.

Goal: $\arg\min_x \left(\|Ax-b\|^2 + \lambda \|x\|_1 \right)$

LSE   L1 norm

gradient of LSE: $\frac{\partial E_{LSE}}{\partial x} = \frac{\partial}{\partial x}(Ax-b)^2$

$= \frac{\partial}{\partial x}(Ax-b)^T(Ax-b)$

$= \frac{\partial}{\partial x}(x^TA^T-b)(Ax-b)$

$= \frac{\partial}{\partial x}(x^TA^TAx - 2x^TA^Tb + b^Tb)$

$= 2A^TAx - 2A^Tb$

$= 2A^T(Ax-b)$

gradient for L1: $\frac{\partial E_{L1}}{\partial x} = \frac{\partial}{\partial x}(\lambda|x|) = \lambda \operatorname{sign}(x)$, where $\operatorname{sign}(x) = \begin{cases} -1, & \text{if } x<0 \\ 0, & \text{if } x=0 \\ 1, & \text{if } x>0 \end{cases}$

the gradient for total error: $\frac{\partial E}{\partial x} = 2A^T(Ax-b) + \lambda \operatorname{sign}(x)$

with initial $x = x^0$, learning rate α

update $x \leftarrow x - 2A^T(Ax-b) + \lambda \cdot \operatorname{sign} x$

(3) Newton's method.

From Taylor Expansion, we have

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2!}f''(x_0)(x-x_0)^2 + \frac{1}{3!}f^{3}(x_0)(x-x_0)^3 + \cdots = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

Let $g(x) = \sum_{n=0}^{2} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$

$$= f(x_0) + \nabla f(x_0)\cdot(x-x_0) + \frac{1}{2!}\nabla^2 f(x_0)\cdot(x-x_0)^2$$

Let $\Delta x = x - x_0$

$$\frac{\partial g(x)}{\partial \Delta x} = \nabla f(x_0) + 2\cdot\frac{1}{2}\nabla^2 f(x_0)\cdot\Delta x$$

to optimize, $\frac{\partial g(x)}{\partial \Delta x} = 0$

$$\nabla f(x_0) + \nabla^2 f(x_0)\cdot\Delta x = 0$$

$$\Delta x = \frac{-\nabla f(x_0)}{\nabla^2 f(x_0)}$$

$$x = x_0 + \Delta x = x_0 - \frac{\nabla f(x_0)}{\nabla^2 f(x_0)}$$

if $f(x) = LSE = \|Ax-b\|^2$

$$= (Ax-b)^T(Ax-b)$$

$$= (x^TA^T - b^T)(Ax-b)$$

$$= x^TA^TAx - 2x^TA^Tb + b^Tb$$

$$\nabla f(x) = 2A^TAx - 2A^Tb$$

$$H = \nabla^2 f(x) = 2A^TA$$

$$x_{n+1} = x_n - H^{-1}\nabla f(x_n)$$

Let $x_n = 0$

$$x_{n+1} = 0 - (2A^TA)^{-1}(2A^TA\cdot 0 - 2A^Tb)$$

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

To find $X = A^{-1}$ that $AX = I$

1. do LU decomposition that $A = LU \rightarrow AX = LUX = I$
2. let $UX = Y \rightarrow$ find Y by solve $LY = I$
3. find X by solve $UX = Y$

(1) $A = LU$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

① $R_2 \leftarrow R_2 - \frac{a_{21}}{a_{11}} R_1$, $R_3 \leftarrow R_3 - \frac{a_{31}}{a_{11}} R_1$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} - \frac{a_{21}}{a_{11}} \times a_{11} & a_{22} - \frac{a_{21}}{a_{11}} \times a_{12} & a_{23} - \frac{a_{21}}{a_{11}} \times a_{13} \\ a_{31} - \frac{a_{31}}{a_{11}} \times a_{11} & a_{32} - \frac{a_{31}}{a_{11}} \times a_{12} & a_{33} - \frac{a_{31}}{a_{11}} \times a_{13} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}' & a_{23}' \\ 0 & a_{32}' & a_{33}' \end{bmatrix}$$

② $R_3 \leftarrow R_3 - \frac{a_{32}'}{a_{22}'} R_2$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}' & a_{23}' \\ 0 & a_{32}' - \frac{a_{32}'}{a_{22}'} \cdot a_{22}' & a_{33}' - \frac{a_{32}'}{a_{22}'} \cdot a_{23}' \end{bmatrix} = \underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}' & a_{23}' \\ 0 & 0 & a_{33}'' \end{bmatrix}}_{= U} \qquad \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ \frac{a_{21}}{a_{11}} & 1 & 0 \\ \frac{a_{31}}{a_{11}} & \frac{a_{32}'}{a_{22}'} & 1 \end{bmatrix}}_{= L}$$

(2) $LY = I$, given L & I find Y, using forward substitution

$$\begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \cdot \begin{array}{c} \begin{matrix} y_1 & y_2 & y_3 \end{matrix} \\ \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{bmatrix} \end{array} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

① $L y_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

$$\begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} y_{11} \\ y_{21} \\ y_{31} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

#1. row 1

$l_{11} \cdot y_{11} = 1 \rightarrow y_{11} = \frac{1}{l_{11}}$

#2 row 2

$l_{21} \cdot y_{11} + l_{22} y_{21} = 0 \rightarrow l_{22} y_{21} = -l_{21} y_{11} \rightarrow y_{21} = \frac{-l_{21} y_{11}}{l_{22}}$

#3. row 3

$l_{31} \cdot y_{11} + l_{32} y_{12} + l_{33} y_{13} = 0 \rightarrow l_{33} y_{13} = -l_{31} y_{11} - l_{32} y_{12} \rightarrow y_{13} = \frac{-l_{31} y_{11} - l_{32} y_{12}}{l_{33}}$

② $L y_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

③ $L y_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

(3) $UX = Y$ given U & Y find X, using backward substitution

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} \end{array} = \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{bmatrix}$$

① $Ux_1 = y_1$

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} y_{11} \\ y_{21} \\ y_{31} \end{bmatrix}$$

#1 row 3

$$u_{33} \cdot x_{31} = y_{31} \rightarrow x_{31} = \frac{y_{31}}{u_{33}}$$

#2 row 2

$$u_{22} \cdot x_{21} + u_{23} \cdot x_{31} = y_{21} \rightarrow u_{22} x_{21} = y_{21} - u_{23} \cdot x_{31} \rightarrow x_{21} = \frac{y_{21} - u_{23} \cdot x_{31}}{u_{22}}$$

row 3

$$u_{11} \cdot x_{11} + u_{12} \cdot x_{21} + u_{13} \cdot x_{31} = y_{11} \rightarrow u_{11} x_{11} = y_{11} - u_{12} x_{21} - u_{13} x_{31}$$

$$x_{11} = \frac{y_{11} - u_{12} x_{21} - u_{13} x_{31}}{u_{11}}$$