

Minimization of Quadratic Function with Linear Constraints

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad Ax = b$$

Quadratic function

Application $\min_{x \in \mathbb{R}^n} \|x\|^2 \quad \text{s.t.} \quad Ax = b$

Find the solution with min norm

Suppose this has lots (infinitely many) solns (e.g. 3×5 , $m \times n$ $m < n$)
undetermined system

Lagrange Multiplier λ (a vector)

$$\begin{aligned} L(x, \lambda) &= \|x\|^2 + \lambda^T (Ax - b) \\ &= x^T x + \lambda^T A x - \lambda^T b \end{aligned}$$

$A^{m \times n}$ $\mathbb{R}^n$ $\mathbb{R}^m$
 $\mathbb{R}^m$ (row) $\mathbb{R}^m$ (column)

$$\nabla_X L = 2X + A^T \lambda = 0$$

$$\nabla_\lambda L = AX - b = 0$$

(to be explained)

$$X = -\frac{1}{2} A^T \lambda$$

$$AX = b$$

X is known
once λ is found

$$\Rightarrow AX = -\frac{1}{2} (AA^T) \lambda$$

$$b = -\frac{1}{2} (AA^T) \lambda$$

solvable for λ
($\because AX = b$
is solvable for X)

Note: AA^T is a square matrix $m \times m$
 $m \times n$ $n \times m$

Suppose $(AA^T)^{-1}$ exists

($\Leftrightarrow$ rows of A (or columns of A^T) are lin ind.)

Then

$$\lambda = - (AA^T)^{-1} 2b$$

$$X = A^T (AA^T)^{-1} b$$

Not to be confused with Least Square Soln:

$$AX = b$$

(inconsistent)

$$A^T A X = A^T b$$

(Normal equation)

$$X = (A^T A)^{-1} A^T b$$

(Suppose $(A^T A)^{-1}$ exists
 $\iff$ cols of A are
lin ind.)

Ex $\min \quad x^2 + y^2 + z^2$

s.t. $x + 2y - z = 1$
 $x + y + z = 0$

$$\begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

A1 (Apply formula)

$$X = A(A^T A)^{-1} b$$

$$= \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 1 \end{bmatrix} \frac{\begin{bmatrix} 3 & -2 \\ -2 & 6 \end{bmatrix}}{14} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}}{ad-bc}$$

$$= \frac{1}{14} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \frac{1}{14} \begin{bmatrix} 1 \\ 4 \\ -5 \end{bmatrix}$$

A2 (Start from scratch)

$$L(\vec{x}, \vec{\lambda}) = x^2 + y^2 + z^2 + \lambda_1 (x + 2y - z - 1) + \lambda_2 (x + y + z - 0)$$

$\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ $\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$

$$\partial_x = 0 \Rightarrow 2x + \lambda_1 + \lambda_2 = 0$$

$$\partial_y = 0 \Rightarrow 2y + 2\lambda_1 + \lambda_2 = 0$$

$$\partial_z = 0 \Rightarrow 2z - \lambda_1 + \lambda_2 = 0$$

$$\partial_{\lambda_1} = 0 \Rightarrow x + 2y - z = 1$$

$$\partial_{\lambda_2} = 0 \Rightarrow x + y + z = 0$$

$$x = -\frac{1}{2} (\lambda_1 + \lambda_2)$$

$$y = -\frac{1}{2} (2\lambda_1 + \lambda_2)$$

$$z = -\frac{1}{2} (-\lambda_1 + \lambda_2)$$

$$\Rightarrow \begin{aligned} 6\lambda_1 + 2\lambda_2 &= -2 \\ 2\lambda_1 + 3\lambda_2 &= 0 \end{aligned}$$

$$\lambda_1 = -\frac{3}{7}, \quad \lambda_2 = \frac{2}{7},$$

$$\Rightarrow x = \frac{1}{14}, \quad y = \frac{4}{14}, \quad z = -\frac{5}{14}$$

A3 (Solve $AX=b$ and minimize)

$$x + 2y - z = 1$$

$$x + y + z = 0$$

set z to be free
(express x, y in terms of z)

$$x + 2y = 1 + z$$

$$x + y = -z$$

$$\Rightarrow \underline{y = 1 + 2z}, \quad \begin{aligned} x &= -y - z \\ &= -1 - 2z - z \\ &= \underline{-1 - 3z} \end{aligned}$$

$$x^2 + y^2 + z^2 = \underbrace{(-1-3z)^2 + (1+2z)^2 + z^2}_{\text{min over } z}$$

$$\frac{d}{dz} \quad \downarrow \quad 2 \left[(-1-3z)(-3) + (1+2z)2 + z \right] = 0$$

$$3 + 9z + 2 + 4z + z = 0$$

$$z = -\frac{3}{14}$$

$$x = -1 - 3z = -1 + \frac{15}{14} = \frac{1}{14}$$

$$y = 1 + 2z = 1 - \frac{10}{14} = \frac{4}{14}$$

$$f(x) = \underline{C^T X} = c_1 x_1 + c_2 x_2 + \dots + c_n x_n \\ = \underline{x^T C}$$

$$\nabla f(x) = \begin{pmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \\ \vdots \\ \partial f / \partial x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \underline{C} \quad (\underline{\text{col. vec.}})$$

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

$$\nabla_x L = \nabla_x [x^T x + \lambda^T A x - \lambda^T b] \\ = x + x + (\lambda^T A)^T = 2x + A^T \lambda$$

$$\nabla_\lambda L = \nabla_\lambda [\lambda^T A x - \lambda^T b] \\ = A x - b$$

Some textbooks use $\underline{\partial f(x)} = (\underline{\nabla f(x)})^T$
row vec col vec

More general Quadratic functions

$$f(X) = \underbrace{\frac{1}{2} X^T Q X}_{\text{quadratic fct}} + \underbrace{p^T X}_{\text{linear fct}}$$

(1) $X \in \mathbb{R}^n$

(2) $Q^{n \times n}$ (often assumed to be symmetric) $Q^T = Q$

(3) $p \in \mathbb{R}^n$

$$\min f(X) = \frac{1}{2} X^T Q X + p^T X$$

s.t. $AX = b$ ↪ assumed solvable

$$L(X, \lambda) = \frac{1}{2} X^T Q X + p^T X + \lambda^T (AX - b)$$

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

$$\nabla_X L \Rightarrow \frac{QX + Q^T X}{2} + p + A^T \lambda = 0$$

$$\left(\frac{Q + Q^T}{2} \right) X = -A^T \lambda - p$$

$Q^T = Q \Rightarrow$

$Q X = -A^T \lambda - p$

$$\nabla_\lambda L \Rightarrow$$

$$AX = b$$

↖

Assumed Q^{-1} exists

$$\Rightarrow X = Q^{-1}[-A^T \lambda - p] \\ = -Q^{-1}A^T \lambda - Q^{-1}p$$

$$A(-Q^{-1}A^T \lambda - Q^{-1}p) = b$$

$$(AQ^{-1}A^T)\lambda + AQ^{-1}p = -b$$

Assume $(AQ^{-1}A^T)^{-1}$ exists

$$\lambda = -(AQ^{-1}A^T)^{-1}[AQ^{-1}p + b]$$

$$X = Q^{-1}A^T(AQ^{-1}A^T)^{-1}[AQ^{-1}p + b] - Q^{-1}p$$