Constrained LS (CLS)

$$\min_{x \in \mathbb{R}^n} \|Ax-b\|_2^2 \qquad A \in \mathbb{R}^{m\times n}, \quad b \in \mathbb{R}^m$$

$$\text{s.t.} \quad Cx = d \qquad C \in \mathbb{R}^{p\times n}, \quad d \in \mathbb{R}^p$$

$$C = \begin{bmatrix} c_1^T \\ c_2^T \\ \vdots \\ c_p^T \end{bmatrix} \quad \text{where} \quad c_i^T \in \mathbb{R}^{1\times n}$$

Construct the Lagrangian function:

$$L(x, z) = \|Ax-b\|_2^2 + z_1(c_1^T x - d_1) + z_2(c_2^T x - d_2) + \dots + z_p(c_p^T x - d_p)$$

where $z = \begin{pmatrix} z_1 \\ \vdots \\ z_p \end{pmatrix} \in \mathbb{R}^p$ is a vector of Lagrange multipliers.

If $\hat{x}$ is a solution of CLP, then there exist Lagrange multipliers $\hat{z}$ s.t.

$$\frac{\partial L}{\partial x_i}(\hat{x}, \hat{z}) = 0 \quad \text{for } i = 1, 2, \ldots, n$$

$$\frac{\partial L}{\partial z_i}(\hat{x}, \hat{z}) = 0 \quad \text{for } i = 1, 2, \ldots, p$$

$\Big\}$ optimality conditions for CLP.

$$\frac{\partial L}{\partial z_i}(\hat{x}, \hat{z}) = 0 \quad \Rightarrow \quad c_i^T \hat{x} = d_i \quad \text{—— (2)} \quad \text{for } i = 1, 2, \ldots, p$$

(constraint satisfaction by $\hat{x}$).

$$\frac{\partial}{\partial x_i} L(x, z) = 2 \sum_{j=1}^{n} (A^T A)_{ij} x_j - 2(A^T b)_i + \sum_{j=1}^{p} z_j (c_j)_i$$

$$\therefore \quad \nabla_x L(x, z) = 2(A^T A)x - 2A^T b + C^T z \quad \text{—— (1)}$$

$$\frac{\partial L}{\partial x_i}(\hat{x}, \hat{z}) = 0 \;\; \& \;\; \frac{\partial L}{\partial z_i}(\hat{x}, \hat{z}) = 0 \quad \Rightarrow \quad \begin{bmatrix} 2A^T A & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{z} \end{bmatrix} = \begin{bmatrix} 2A^T b \\ d \end{bmatrix}$$

KKT conditions

(Karush, Kuhn, Tucker)

KKT matrix $\begin{bmatrix} 2A^TA & C^T \\ C & 0 \end{bmatrix} \in \mathbb{R}^{(n+p)\times(n+p)}$ is

invertible if and only if C has linearly independent rows and $\begin{bmatrix} A \\ C \end{bmatrix}$ has linearly independent columns.

Assume C has lin. indep rows & $\begin{bmatrix} A \\ C \end{bmatrix}$ has linearly indep. cols.

To prove. KKT matrix is invertible.

Let $\begin{pmatrix} \bar{x} \\ \bar{z} \end{pmatrix}$ be a vector such that

$$\begin{bmatrix} 2A^TA & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} \bar{x} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\Rightarrow \quad 2A^TA\bar{x} + C^T\bar{z} = 0$ — (1)

$\Rightarrow \quad 2\bar{x}^TA^TA\bar{x} + \bar{x}^TC^T\bar{z} = 0 \Rightarrow 2\|Ax\|_2^2 + (\cancel{Cx})^T\bar{z} = 0$

0

$=0$

$\Rightarrow \quad 2\|A\bar{x}\|_2^2 = 0$

$\Rightarrow \quad A\bar{x} = 0$ and $C\bar{x} = 0$

$\Rightarrow \quad \begin{bmatrix} A \\ C \end{bmatrix} \bar{x} = 0$

$\Rightarrow \quad \bar{x} = 0$

From (1), $\quad C^T \bar{z} = 0 \Rightarrow \bar{z} = 0 \quad \because$ columns of C^T (rows of C) are lin. indep.

$\Rightarrow$ KKT matrix invertible.

Converse:

$$\boxed{\begin{pmatrix} \hat{x} \\ \hat{z} \end{pmatrix} = \begin{pmatrix} 2A^TA & C^T \\ C & 0 \end{pmatrix}^{-1} \begin{pmatrix} A^Tb \\ d \end{pmatrix}}$$

Recursive LS problem: (12.4, pg. 242, vmls Book)

$$A = \begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_k^T & - \end{bmatrix} \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}$$

Matrix inversion lemma:

For a matrix $S \in \mathbb{R}^{n\times n}$ and vectors $u \in \mathbb{R}^{n\times 1}$, $v^T \in \mathbb{R}^{1\times n}$

$$(S + uv^T)^{-1} = S^{-1} - S^{-1} u (1 + v^T u)^{-1} v^T$$

$$= S^{-1} - \frac{(S^{-1} u) v^T}{(1 + v^T u)}$$

$$\rightarrow \quad A^{(k)} = \begin{bmatrix} a_1^T \\ \vdots \\ a_k^T \end{bmatrix}, \qquad b^{(k)} = \begin{bmatrix} b_1 \\ \vdots \\ b_k \end{bmatrix}$$

$$x^{(k)} = \left[(A^k)^T (A^k) \right]^{-1} (A^k)^T \, b^k$$

$$\rightarrow \quad A^{(k+1)} = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_k^T \\ a_{k+1}^T \end{bmatrix} = \begin{bmatrix} A^k \\ a_{k+1}^T \end{bmatrix}, \qquad b^{(k+1)} = \begin{bmatrix} b^{(k)} \\ b_{k+1} \end{bmatrix}$$

$$x^{(k+1)} = \left(\left[A^{(k+1)} \right]^T A^{(k+1)} \right)^{-1} (A^{(k+1)})^T \, b^{(k+1)}$$

$$\left(A^{(k+1)}\right)^T \left(A^{(k+1)}\right) = \begin{bmatrix} A^{(k)} \\ a_{k+1}^T \end{bmatrix}^T \begin{bmatrix} A^{(k)} \\ a_{k+1}^T \end{bmatrix}$$

$$= \begin{bmatrix} \left(A^{(k)}\right)^T & a_{k+1} \end{bmatrix} \begin{bmatrix} A^{(k)} \\ a_{k+1}^T \end{bmatrix}$$

$$= \left(A^{(k)}\right)^T \left(A^{(k)}\right) + a_{k+1} a_{k+1}^T$$

$$\left(\left(A^{(k+1)}\right)^T A^{(k+1)}\right)^{-1} = \left(\left(A^{(k)}\right)^T A^{(k)} + a_{k+1} a_{k+1}^T\right)^{-1}$$

$$= \left(\left(A^{(k)}\right)^T A^{(k)}\right)^{-1} - \frac{\left(\left(A^{(k)}\right)^T A^{(k)}\right)^{-1} a_{k+1} a_{k+1}^T}{1 + a_{k+1}^T a_{k+1}}$$

$$\left(A^{(k+1)}\right)^T b^{(k+1)} = \begin{bmatrix} A^{(k)} \\ a_{k+1}^T \end{bmatrix}^T \begin{bmatrix} b^{(k)} \\ b_{k+1} \end{bmatrix}$$

$$= \begin{bmatrix} \left(A^{(k)}\right)^T & a_{k+1} \end{bmatrix} \begin{bmatrix} b^{(k)} \\ b_{k+1} \end{bmatrix}$$

$$= \left(A^{(k)}\right)^T b^{(k)} + a_{k+1} b_{k+1}$$

$$x^{(k+1)} = \left[\left(A^{(k+1)}\right)^T A^{(k+1)}\right]^{-1} \left(A^{(k+1)}\right)^T b^{(k+1)}$$

$$x^{(k+1)} = \left\{ \left[\left(A^{(k)}\right)^T \left(A^{(k)}\right) \right]^{-1} - \frac{\left(\left(A^{(k)}\right)^T A^{(k)} \right)^{-1} a_{k+1} a_{k+1}^T}{1 + a_{k+1}^T a_{k+1}} \right\}$$

$$\left\{ \left(A^{(k)}\right)^T b^{(k)} + a_{k+1} b_{k+1} \right\}$$

$$x^{(k+1)} = x^{(k)} +$$