

CS532100 Numerical Optimization Homework 2

109062710 Bijon Setyawan Raya

Due Dec 10

1. Consider the linear least square problem:

$$\min_{\vec{x} \in \mathbb{R}^2} ||A\vec{x} - \vec{b}||^2$$

where

$$A = \begin{bmatrix} 4 & 8 \\ 2 & 4 \\ 1 & 2 \end{bmatrix}, \vec{b} = \begin{pmatrix} 21/4 \\ 0 \\ 0 \end{pmatrix}$$

- (a) (10%) Write its normal equation.

$$A^T A \vec{x} = A^T \vec{b} \quad (1)$$

$$\begin{bmatrix} 4 & 2 & 1 \\ 8 & 4 & 2 \end{bmatrix} \begin{bmatrix} 4 & 8 \\ 2 & 4 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 1 \\ 8 & 4 & 2 \end{bmatrix} \begin{bmatrix} 21/4 \\ 0 \\ 0 \end{bmatrix} \quad (2)$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad (3)$$

$$x_1 + 2x_2 = 1 \quad (4)$$

Then, we can express x_1 and x_2 as

$$\vec{x} = \begin{bmatrix} 1 - 2t \\ t \end{bmatrix} \quad (5)$$

- (b) (10%) Express $\vec{b} = \vec{b}_1 + \vec{b}_2$ such that \vec{b}_1 is in the subspace spanned by A 's column vectors, and \vec{b}_2 is orthogonal to A 's column vectors. From matrix A , we know that its basis is its first column which is $\begin{bmatrix} 4 & 2 & 1 \end{bmatrix}^T$. From its basis column, we can express \vec{b} in $\vec{b}_1 + \vec{b}_2$.
Let,

$$\vec{b}_1 = \begin{bmatrix} 4s \\ 2s \\ s \end{bmatrix}, \vec{b}_2 = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \quad (6)$$

Then we can express \vec{b} such as

$$\vec{b} = \vec{b}_1 + \vec{b}_2 \quad (7)$$

$$\begin{bmatrix} 21/4 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 4s + t_1 \\ 2s + t_2 \\ s + t_3 \end{bmatrix} \quad (8)$$

Since \vec{b}_1 is a column space of A, and \vec{b}_2 is orthogonal to \vec{b}_1 , then

$$\begin{bmatrix} 4s \\ 2s \\ s \end{bmatrix} \begin{bmatrix} t_1 & t_2 & t_3 \end{bmatrix} = 0 \quad (9)$$

$$4s \cdot t_1 + 2s \cdot t_2 + s \cdot t_3 = 0 \quad (10)$$

We can add $4s \cdot t_1 + 2s \cdot t_2 + s \cdot t_3 = 0$ at the last index of our matrix equality. Then

$$\begin{bmatrix} 4s + t_1 \\ 2s + t_2 \\ s + t_3 \\ 4s \cdot t_1 + 2s \cdot t_2 + s \cdot t_3 \end{bmatrix} = \begin{bmatrix} 21/4 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving the equality above, we then have $s = 1$, $t_1 = 5/4$, $t_2 = -2$, and $t_3 = -1$. Thus,

$$\vec{b}_1 = \begin{bmatrix} 4s \\ 2s \\ s \end{bmatrix}, \vec{b}_2 = \begin{bmatrix} 5/4 \\ -2 \\ -1 \end{bmatrix} \quad (11)$$

$$\begin{bmatrix} 21/4 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 5/4 \\ -2 \\ -1 \end{bmatrix} \quad (12)$$

- (c) (10%) Show that $\vec{z} \in \mathbb{R}^2$ is a least square solution for $A\vec{x} = \vec{b}$ if and only if \vec{z} is part of a solution to the larger linear system:

$$\begin{bmatrix} 0 & A^T \\ A & I \end{bmatrix} \begin{bmatrix} \vec{z} \\ \vec{y} \end{bmatrix} = \begin{bmatrix} 0 \\ \vec{b} \end{bmatrix}$$

We are going to show that $\vec{z} \in \mathbb{R}^2$ is a least square solution for $A\vec{x} = \vec{b}$ \iff \vec{z} is part of a solution to the larger linear system

- i. First, we are going to show that if $\vec{z} \in \mathbb{R}^2$ is a least square solution for $A\vec{x} = \vec{b}$ then \vec{z} is part of a solution to the larger linear system.

$$A\vec{x} = \vec{b} \quad (13)$$

$$A^T A\vec{x} = A^T \vec{b} \quad (14)$$

$$\begin{bmatrix} 4 & 2 & 1 \\ 8 & 4 & 2 \end{bmatrix} \begin{bmatrix} 4 & 8 \\ 2 & 4 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 1 \\ 8 & 4 & 2 \end{bmatrix} \begin{bmatrix} 21/4 \\ 0 \\ 0 \end{bmatrix} \quad (15)$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad (16)$$

$$x_1 + 2x_2 = 1 \quad (17)$$

Then, we can express x_1 and x_2 as

$$\vec{x} = \begin{bmatrix} 1 - 2t \\ t \end{bmatrix} \quad (18)$$

Plugging \vec{x} into \vec{z} in the larger linear system, we then have

$$\begin{bmatrix} 0 & A^T \\ A & I \end{bmatrix} \begin{bmatrix} \vec{z} \\ \vec{y} \end{bmatrix} = \begin{bmatrix} 0 \\ \vec{b} \end{bmatrix} \quad (19)$$

$$\begin{bmatrix} 0 & 0 & 4 & 2 & 1 \\ 0 & 0 & 8 & 4 & 2 \\ 4 & 8 & 1 & 0 & 0 \\ 2 & 4 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1-2t \\ t \\ \vec{y}_1 \\ \vec{y}_2 \\ \vec{y}_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 21/4 \\ 0 \\ 0 \end{bmatrix} \quad (20)$$

Solving \vec{y} , we have

$$\vec{y} = \begin{bmatrix} 5/4 \\ -2 \\ -1 \end{bmatrix} \quad (21)$$

From \vec{y} , we can tell that $\vec{z} \in \mathbb{R}^2$ is a least square solution for $A\vec{x} = \vec{b}$.

Thus, the premise, if $\vec{z} \in \mathbb{R}^2$ is a least square solution for $A\vec{x} = \vec{b}$ then \vec{z} is part of a solution to the larger linear system, holds.

- ii. Second, we are going to show that if \vec{z} is a part of the solution to the larger linear system then $\vec{z} \in \mathbb{R}^2$ is a least square solution for $A\vec{x} = \vec{b}$.

The larger linear system can be expressed as

$$\begin{bmatrix} 0 & A^T \\ A & I \end{bmatrix} \begin{bmatrix} \vec{z} \\ \vec{y} \end{bmatrix} = \begin{bmatrix} 0 \\ \vec{b} \end{bmatrix} \quad (22)$$

$$\begin{bmatrix} 0 & 0 & 4 & 2 & 1 \\ 0 & 0 & 8 & 4 & 2 \\ 4 & 8 & 1 & 0 & 0 \\ 2 & 4 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \vec{z}_1 \\ \vec{z}_2 \\ \vec{y}_1 \\ \vec{y}_2 \\ \vec{y}_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 21/4 \\ 0 \\ 0 \end{bmatrix} \quad (23)$$

Solving y_1 , y_2 , and y_3 , we express them as

$$\vec{y} = \begin{bmatrix} 5/4 \\ -2 \\ -1 \end{bmatrix} \quad (24)$$

After acquiring \vec{y} , we can acquire \vec{z} with the third, the fourth, and the fifth row in the matrix. Thus,

$$4\vec{z}_1 + 8\vec{z}_2 + \vec{y}_1 = \frac{21}{4} \quad (25)$$

$$2\vec{z}_1 + 4\vec{z}_2 + \vec{y}_2 + 2 = 0 \quad (26)$$

$$\vec{z}_1 + 2\vec{z}_2 + \vec{y}_3 = 0 \quad (27)$$

Solving the equalities above, we have

$$\vec{z}_1 + 2\vec{z}_2 = 1 \quad (28)$$

We can express the equality above in matrix as,

$$\vec{z} = \begin{bmatrix} 1 - 2t \\ t \end{bmatrix} \quad (29)$$

Clearly, \vec{z} is a part of the solution to the larger linear system. Thus, the premise, if \vec{z} is a part of the solution to the larger linear system then $\vec{z} \in \mathbb{R}^2$ is a least square solution for $A\vec{x} = \vec{b}$, holds.

By showing the premises above, we can conclude that the statement, $\vec{z} \in \mathbb{R}^2$ is a least square solution for $A\vec{x} = \vec{b}$ if and only if \vec{z} is part of a solution to the larger linear system, holds.

2. In Note05 (Page 16), memoryless BFGS iteration matrix H_{k+1} can be derived from considering the Hestenes–Stiefel form of the nonlinear conjugate gradient method. Recalling that $\vec{s}_k = \alpha_k \vec{p}_k$, we have that the search direction for this method is given by

$$\vec{p}_{k+1} = -\nabla f_{k+1} + \frac{\nabla f_{k+1}^T \vec{y}_k}{\vec{y}_k^T \vec{p}_k} \vec{p}_k \quad (30)$$

$$= -\nabla f_{k+1} + \frac{\nabla f_{k+1}^T \vec{y}_k}{\vec{y}_k^T \vec{s}_k} \vec{s}_k \quad (31)$$

$$= -(I - \frac{\vec{s}_k \vec{y}_k^T}{\vec{y}_k^T \vec{s}_k}) \nabla f_{k+1} \quad (32)$$

$$= -\hat{H}_{k+1} \nabla f_{k+1} \quad (33)$$

However, the matrix \hat{H}_{k+1} is neither symmetric nor positive definite.

- (a) (10%) Please show that the matrix \hat{H}_{k+1} is singular.
(You can only prove it for the case $\nabla f_k, \vec{p}_k, \vec{y}_k, \vec{s}_k \in \mathbb{R}^2$ for all $k \in \mathbb{N}$.)
[Your answer here!](#)
- (b) (0%) Please read the reference book (Page 180) to understand the derivation of the inverse hessian formula in Note05 (Page 16).
(you don't need to write anything in this subproblem.)

$$H_{k+1} = (I - \frac{\vec{s}_k \vec{y}_k^T}{\vec{y}_k^T \vec{s}_k})(I - \frac{\vec{y}_k \vec{s}_k^T}{\vec{y}_k^T \vec{s}_k}) + \frac{\vec{s}_k \vec{s}_k^T}{\vec{y}_k^T \vec{s}_k}$$

3. (10%) The total least square problem is to solve the following problem

$$\min_{\vec{x}, \|\vec{x}\|=1} \vec{x}^T A^T A \vec{x}$$

where A is an $m \times n$ matrix. Here we assume $m > n$. Let $A = U\Sigma V^T$ be the SVD of A , where U is the matrix of left singular vectors, V is the matrix of right singular vectors, and Σ is a diagonal matrix with diagonal elements $\sigma_1, \sigma_2, \dots, \sigma_n$. Moreover, U and V are orthogonal matrices, and $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$. Show the solution of the above problem is the σ^2 .

[Your answer here!](#)

4. Consider the following linear programming problem:

$$\begin{aligned}
 \max_{x_1, x_2} \quad & z = x_1 + x_2 \\
 \text{s.t.} \quad & x_1 + 2x_2 \leq 4 \\
 & 4x_1 + 2x_2 \leq 12 \\
 & -x_1 + x_2 \leq 1 \\
 & x_1, x_2 \geq 0
 \end{aligned} \tag{34}$$

(a) (10%) Please refer Note08 (Page 2) to draw the figure of the constraints by any means, and use that to solve the problem.

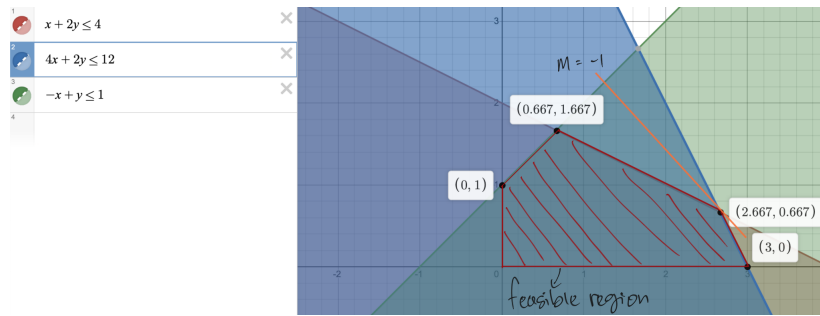


Figure 1: Feasible region of \vec{z} .

(b) (10%) Derive its dual problem and solve the dual problem by any means. Compare the solutions of the primal and the dual problems.
The dual problem is

$$\begin{aligned}
 \min_{y_1, y_2, y_3} \quad & 4y_1 + 12y_2 + y_3 \\
 \text{s.t.} \quad & y_1 + 4y_2 - y_3 \geq 1 \\
 & 2y_1 + 2y_2 + y_3 \geq 1 \\
 & y_1, y_2, y_3 \geq 0
 \end{aligned} \tag{35}$$

(c) (10%) Verify the complementarity slackness condition. [Answer here!](#)

(d) (10%) Transform the problem to the standard form.

$$\begin{aligned}
 \max_{x_1, x_2} \quad & z = x_1 + x_2 \\
 \text{s.t.} \quad & x_1 + 2x_2 + x_3 = 4 \\
 & 4x_1 + 2x_2 + x_4 = 12 \\
 & -x_1 + x_2 + x_5 = 1 \\
 & x_1, x_2, x_3, x_4, x_5 \geq 0
 \end{aligned} \tag{36}$$

(e) (10%) Solve it by the simplex method, as provided in Figure 1, using $\vec{x}_0 = (0, 0)$. Indicate $B_k, N_k, \vec{s}_k, \vec{d}_k, p_k, q_k, \gamma_k$ in each step. [Answer here!](#)

-
- (1) Given a basic feasible point \vec{x}_0 and the corresponding index set \mathcal{B}_0 and \mathcal{N}_0 .
 - (2) For $k = 0, 1, \dots$
 - (3) Let $B_k = A(:, \mathcal{B}_k)$, $N_k = A(:, \mathcal{N}_k)$, $\vec{x}_B = \vec{x}_k(\mathcal{B}_k)$, $\vec{x}_N = \vec{x}_k(\mathcal{N}_k)$,
and $\vec{c}_B = \vec{c}_k(\mathcal{B}_k)$, $\vec{c}_N = \vec{c}_k(\mathcal{N}_k)$.
 - (4) Compute $\vec{s}_k = \vec{c}_N - N_k^T (B_k^{-1})^T \vec{c}_B$ (pricing)
 - (5) If $\vec{s}_k \geq 0$, return the solution \vec{x}_k . (found optimal solution)
 - (6) Select $q_k \in \mathcal{N}_k$ such that $\vec{s}_k(i_{q_k}) < 0$,
where i_{q_k} is the index of q_k in \mathcal{N}_k
 - (7) Compute $\vec{d}_k = B_k^{-1} A_k(:, q_k)$. (search direction)
 - (8) If $\vec{d}_k \leq 0$, return **unbounded**. (unbounded case)
 - (9) Compute $[\gamma_k, i_p] = \min_{i, \vec{d}_k(i) > 0} \frac{\vec{x}_B(i)}{\vec{d}_k(i)}$ (ratio test)
(The first return value is the minimum ratio;
the second return value is the index of the minimum ratio.)
 - (10) $x_{k+1} \begin{pmatrix} \mathcal{B} \\ \mathcal{N} \end{pmatrix} = \begin{pmatrix} \vec{x}_B \\ \vec{x}_N \end{pmatrix} + \gamma_k \begin{pmatrix} -\vec{d}_k \\ \vec{e}_{i_{q_k}} \end{pmatrix}$
($\vec{e}_{i_{q_k}} = (0, \dots, 1, \dots, 0)^T$ is a unit vector with i_{q_k} th element 1.)
 - (11) Let the i_p th element in \mathcal{B} be p_k . (pivoting)
 $\mathcal{B}_{k+1} = (\mathcal{B}_k - \{p_k\}) \cup \{q_k\}$, $\mathcal{N}_{k+1} = (\mathcal{N}_k - \{q_k\}) \cup \{p_k\}$
-

Figure 2: The simplex method for solving (minimization) linear programming