# CS532100 Numerical Optimization Homework 2

106033233 Sheng-Yen Chou

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.

Answers are put here.

$$\nabla_{\vec{x}}||A\vec{x}-\vec{b}||^2$$

$$=\nabla_{\vec{x}}(A\vec{x}-\vec{b})^T(A\vec{x}-\vec{b})$$

$$=2A^T(A\vec{x}-\vec{b})$$

Since the minimum is at $\nabla_{\vec{x}}||A\vec{x}-\vec{b}||^2=0$, thus, we can rewrite the formula as

$$2A^T(A\vec{x}-\vec{b})=0$$

$$2A^TA\vec{x}-2A^T\vec{b}=0$$

Then, we can get the normal equation

$$A^TA\vec{x}=A^T\vec{b}$$

(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.

Answers are put here.

Let $\vec{c}$ be the projection of $\vec{a}$ on the $\vec{b}$

Since $rank(\mathbf{A})=1$, let $\vec{b}$ project on the subspace of $\mathbf{A}$ and the projection is $\vec{c}$. Because the set $\mathcal{S}=\{\begin{pmatrix}4\\2\\1\end{pmatrix}\}$ is one of the basis of

**A**. Then we can project $\vec{b}$ onto the basis of **A** and we can get the projection $\vec{c}$ of $\vec{b}$ onto the basis $\mathcal{S}$. The dimension of the basis $\mathcal{S}$ is 1. As a result, we only need to compute the projection onto the vector $\vec{s}_1 = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$

$$\vec{c} = \frac{\langle \vec{b}, \vec{s}_1 \rangle}{||\vec{s}_1||^2} \vec{s}_1$$

$$= \frac{\langle \begin{pmatrix} \frac{21}{4} \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix} \rangle}{21} \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$$

Let $\vec{b}_1 = \vec{c}$. Then, to get the vertical vector between $\vec{c}$ and $\vec{b}$. Compute $\vec{b}_2 = \vec{b} - \vec{c} = \begin{pmatrix} \frac{5}{4} \\ -2 \\ -1 \end{pmatrix}$.

We can simply verify whether $\vec{b}_1$ and $\vec{b}_2$ are orthogonal by computing the inner product. The result is $\langle \vec{b}_1, \vec{b}_2 \rangle = 5 - 4 - 1 = 0$ and we know $\vec{b}_1$ and $\vec{b}_2$ are orthogonal while $\vec{b}_1 = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$ $\quad \vec{b}_2 = \begin{pmatrix} \frac{5}{4} \\ -2 \\ -1 \end{pmatrix}$.

(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:

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

Answers are put here.

RHS $\rightarrow$ LHS

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

According to the linear system, we can convert it to simultaneous equations.

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

We can rewrite the second formula as

$$I\vec{y} = \vec{b} - A\vec{z}$$

$$\vec{y} = \vec{b} - A\vec{z}$$

Plugin into $A^T\vec{y} = 0$

$$A^T(\vec{b} - A\vec{z}) = 0$$

$$-A^TA\vec{z} + A^T\vec{b} = 0$$

Then we can get the normal form of the least square problem.

$$A^TA\vec{z} = A^T\vec{b}$$

Remove $A^T$ from both sides and we can get

$$A\vec{z} = \vec{b}$$

Done.

LHS $\to$ RHS

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.

Due to the least square problem $\min_{\vec{x}\in\mathbb{R}^2} ||A\vec{x} - \vec{b}||^2$, We can get the normal form of the least square problem as the following formula.

$$A^TA\vec{z} = A^T\vec{b}$$

$$A^T\vec{b} - A^TA\vec{z} = 0$$

$$A^T(\vec{b} - A\vec{z}) = 0$$

According to the linear system, we can convert it to simultaneous equations.

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

According to the first formula, the solution of $\vec{y}$ can be $\vec{b} - A\vec{z}$. Then, plugin the solution of $\vec{y}$, the second formula $A\vec{z} + I\vec{y} = \vec{b}$ can be rewritten as

$$A\vec{z} + I(\vec{b} - A\vec{z}) = \vec{b}$$

$$A\vec{z} + \vec{b} = A\vec{z} + \vec{b}$$

The solution of $\vec{y} = \vec{b} - A\vec{z}$ perfectly satisfies the second equation. Then, we finish the proof.

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

$$\begin{aligned}\vec{p}_{k+1} &= -\nabla f_{k+1} + \frac{\nabla f_{k+1}^T \vec{y}_k}{\vec{y}^T \vec{p}_k}\vec{p}_k \\ &= -\nabla f_{k+1} + \frac{\nabla f_{k+1}^T \vec{y}_k}{\vec{y}^T \vec{s}_k}\vec{s}_k \\ &= -(I - \frac{\vec{s}_k \vec{y}_k^T}{\vec{y}^T \vec{s}_k})\nabla f_{k+1} \\ &= -\hat{H}_{k+1}\nabla f_{k+1}\end{aligned}$$

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 $k \in \mathbb{N}$.)

Answers are put here.

Lemma 1

For an identity matrix $I \in \mathbb{R}^n$, $\forall v \in \mathbb{R}^n$ are eigenvectors of the identity matrix $I$.

Proof:

Since $I$ is an identity matrix, thus

$$Iv = v$$

Lemma 2

For a rank 1 matrix $A = uv^T, A \in \mathbb{R}^{n\times n}, rank(A) = 1$ and $u, v \in \mathbb{R}^n$, it has only one eigenvalue $\lambda_1 = u^T v > 0, \lambda_i \in \mathbb{R}, 1 \le i \le n$ and other eigenvalues are $\lambda_2 = \lambda_3 = ... = \lambda_n = 0$. The corresponding eigenvectors is $x_1 = u$.

Proof:

We denote $\lambda_0$ as the set of eigenvectors corresponding to 0 eigenvalues $\lambda_0(A) = \{x | Ax = \lambda_{zero} x, \lambda_{zero} = 0\}$.

Step 1. Show that $\forall A \in \mathbb{R}^{n\times n}, \dim(\text{null}(A)) = k$, if and only if $k$ is the number of the zero eigenvalues of matrix $A$.

LHS $\rightarrow$ RHS:

By definition $\text{null}(A) = \{x | Ax = 0\}$. Thus, $\forall x, Ax = 0 = 0 \cdot x$. The corresponding eigenvalue of $A$ is 0.

As a result, $\text{null}(A) = \lambda_0(A)$ and $\dim(\text{null}(A)) = \dim(\lambda_0(A))$.

RHS $\rightarrow$ LHS:

By definition $\lambda_0(A) = \{x | Ax = \lambda_{zero} x, \lambda_{zero} = 0\} = \{x | Ax = 0\}$. Thus, $\lambda_0(A) = \text{null}(A)$ and $\dim(\lambda_0(A)) = \dim(\text{null}(A))$.

Step 2. Show that the non-zero eigenvalue of $A$ is $u^T v$ and the corresponding eigenvector is $u$.

$$Au = uv^T u = u(v^T u)$$

Since $(v^T u) \in \mathbb{R}$

$$= (v^T u)u = \lambda u$$

Step 3. Combine them

We've know that since rank$(A) = 1$ and dim(null$(A)) = n - 1$(by Gaussian elimination). Thus, $A$ has $n-1$ eigenvalues are 0 and the only one non-zero eigenvector is $u$ and the corresponding eigenvalue is $u^T v$.

From the description of the formula, we know

$$\hat{H}_{k+1} = I - \frac{\vec{s}_k \vec{y}_k^T}{\vec{y}^T \vec{s}_k}$$

Assume $\hat{H}_{k+1} \in \mathbb{R}^{2\times 2}$, $\vec{s}_k = \begin{pmatrix} s_{k1} & s_{k2} \end{pmatrix} \in \mathbb{R}^2$, and $\vec{y}_k = \begin{pmatrix} y_{k1} & y_{k2} \end{pmatrix} \in \mathbb{R}^2$. We can find out the eigenvalue and eigenvector of matrix $A$ as the following formula

$$Av = \lambda v$$

Where $\lambda$ is the eigenvalue and $v$ is the eigenvector. As for matrix $\hat{H}_{k+1}$

$$\hat{H}_{k+1} v = \lambda v$$

$$(\hat{H}_{k+1} - \lambda I)v = 0$$

Let matrix $K = \vec{s}_k \vec{y}_k^T$ and scalar $k = \vec{y}^T \vec{s}_k$

$$(I - \frac{K}{k} - \lambda I)v = 0$$

$$((1-\lambda)I - \frac{K}{k})v = 0$$

$$(K - (k - k\lambda)I)v = 0$$

Let $c = k(1-\lambda)$ and $\lambda = 1 - \frac{c}{k}$

$$(K - cI)v = 0$$

If $K$ has eigenvector, then $(K - cI)$ must be nonsingular matrix. The determinant of $K - cI$ must be 0.

$$det(K - cI) = 0$$

Since we've known rank$(K) = 1$ and $c$ can be seem as the eigenvalue of the matrix $K$. We denote the eigenvalue as $\lambda_K$ and the eigenvector as $v_K$

$$\lambda_K = c = \vec{s}_k^T \vec{y}_k, \quad v_K = \vec{s}_k$$

$$\lambda = 1 - \frac{c}{k} = 1 - \frac{\vec{s}_k^T \vec{y}_k}{\vec{s}_k^T \vec{y}_k} = 0$$

Thus, we known the eigenvalue of $\hat{H}_{k+1}$ must contains 0. As a result, the matrix $\hat{H}_{k+1}$ is singular.

(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 $\Sigma$ is a diagonal matrix with diagonal elements $\sigma_1, \sigma_2, \ldots, \sigma_n$. Moreover, $U$ and $V$ are orthogonal matrices, and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$. Show the solution of the above problem is the $\sigma_n^2$.

Answers are put here.

$$\vec{x}^T A^T A \vec{x}$$

$$= \vec{x}^T (U\Sigma V^T)^T (U\Sigma V^T) \vec{x}$$

$$= \vec{x}^T V\Sigma U^T U\Sigma V^T \vec{x}$$

$$= \vec{x}^T V\Sigma U^T U\Sigma V^T \vec{x}$$

Since $U$ and $V$ are orthogonal matrices, we get

$$= \vec{x}^T V\Sigma^T \Sigma V^T \vec{x}$$

Since $\Sigma$ is diagonal matrix, we can move $\Sigma^T\Sigma$ to the tail

$$= \vec{x}^T V V^T \vec{x} (\Sigma^T \Sigma)$$

$$= \vec{x}^T \vec{x} (\Sigma^T \Sigma)$$

We expand the vector and the matrix. Thus, vector

$$\vec{x} = \begin{pmatrix} x_1 & x_2 & x_3 & ... & x_n \end{pmatrix}$$

and the matrix

$$\Sigma = \begin{pmatrix} \sigma_1^2 & 0 & 0 & ... & 0 \\ 0 & \sigma_2^2 & 0 & ... & 0 \\ 0 & 0 & \sigma_3^2 & ... & 0 \\ ... & ... & ... & & 0 \\ 0 & 0 & 0 & ... & \sigma_n^2 \end{pmatrix}$$

$$= \begin{pmatrix} x_1^2 & x_2^2 & x_3^2 & ... & x_n^2 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 & 0 & ... & 0 \\ 0 & \sigma_2^2 & 0 & ... & 0 \\ 0 & 0 & \sigma_3^2 & ... & 0 \\ ... & ... & ... & & 0 \\ 0 & 0 & 0 & ... & \sigma_n^2 \end{pmatrix}$$

$$= \sum_{i=1}^{n} (x_i \sigma_i)^2$$

Trivially, since $||\vec{x}|| = 1$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$, the minimum should be $(0 \cdot \sigma_1)^2 + (0 \cdot \sigma_2)^2 + ... + (1 \cdot \sigma_n)^2$. Thus,

$$\sigma_n^2 = \min_{\vec{x}, \|\vec{x}\|=1} \sum_{i=1}^{n} (x_i \sigma_i)^2$$

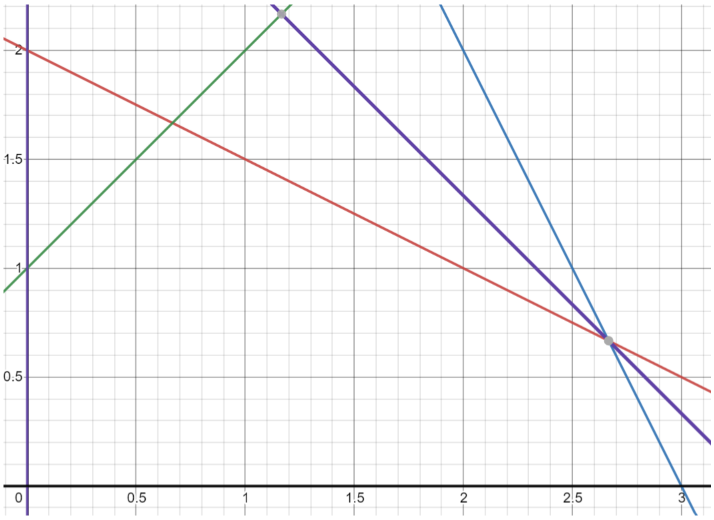
The minimum is $\sigma_n^2$.

4. Consider the following linear programming problem:

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

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

The red line is $x_1 + 2x_2 = 4$, the blue one is $4x_1 + 2x_2 = 12$, the green one is $-x_1 + x_2 = 1$, the purple one is $x_1 = 0$, and the black one is $x_2 = 0$.




The deep purple one is $x_1 + x_2 = \frac{10}{3}$. Trivially, the solution of $\max_{x1,x2} z = x1+x2$ must on the line $x_1+x_2 = z$ and the graph shows that the maximum should be $z = \frac{10}{3}$ at the point $(x_1, x_2) = (\frac{8}{3}, \frac{2}{3})$ because we cannot move the deep purple line to the further right side to get larger maximum $z$.

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

Answers are put here.

(1) Method 1

Convert to standard form

$$\begin{array}{ll} \min_{x_1,x_2} & z = -x_1 - x_2 \\ \text{s.t.} & -x_1 - 2x_2 \geq -4 \\ & -4x_1 - 2x_2 \geq -12 \\ & x_1 - x_2 \geq -1 \\ & x_1, x_2 \geq 0 \end{array}$$

Using Maximum Lower Bound Method

$$\begin{array}{rl} & (-x_1 - 2x_2)y_1 \geq -4y_1 \\ & (-4x_1 - 2x_2)y_2 \geq -12y_2 \\ + & (x_1 - x_2)y_3 \geq -1y_3 \\ \hline & (-y_1 - 4y_2 + y_3)x_1 + (-2y_1 - 2y_2 - y_3)x_2 \geq -4y_1 - 12y_2 - y_3 \end{array}$$

$$\max_{y_1,y_2,y_3} -4y_1 - 12y_2 - y_3$$

$$\begin{array}{ll} \text{subject to} & -y_1 - 4y_2 + y_3 \leq -1 \\ & -2y_1 - 2y_2 - y_3 \leq -1 \\ & y_1, y_2, y_3 \geq 0 \end{array}$$

Convert it to the minimization problem

$$\min_{y_1,y_2,y_3} \ 4y_1 + 12y_2 + y_3$$

$$\begin{aligned} \text{subject to} \quad & y_1 + 4y_2 - y_3 \ge 1 \\ & 2y_1 + 2y_2 + y_3 \ge 1 \\ & y_1, y_2, y_3 \ge 0 \end{aligned}$$

(2) Method 2

Directly convert, use Minimum Upper Bound Method

$$\begin{array}{rl} & (x_1 + 2x_2)y_1 \le 4y_1 \\ & (4x_1 + 2x_2)y_2 \le 12y_2 \\ + & (-x_1 + x_2)y_3 \le 1y_3 \\ \hline & (y_1 + 4y_2 - y_3)x_1 + (2y_1 + 2y_2 + y_3)x_2 \le 4y_1 + 12y_2 + y_3 \end{array}$$

$$\min_{y_1,y_2,y_3} p = 4y_1 + 12y_2 + y_3$$

$$\begin{aligned} \text{subject to} \quad & y_1 + 4y_2 - y_3 \ge 1 \\ & 2y_1 + 2y_2 + y_3 \ge 1 \\ & y_1, y_2, y_3 \ge 0 \end{aligned}$$

(3) Method 3

Consider complementary slackness, add slack variables

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

Transpose

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

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | constant |
|---|---|---|---|---|---|
| 1 | 2 | 1 | 0 | 0 | 4 |
| 4 | 2 | 0 | 1 | 0 | 12 |
| -1 | 1 | 0 | 0 | 1 | 1 |
| 1 | 1 | 0 | 0 | 0 | 1 |

The dual problem is

$$\min_{y_1,y_2,y_3} p = 4y_1 + 12y_2 + y_3$$

$$\begin{aligned} \text{subject to} \quad & y_1 + 4y_2 - y_3 \ge 1 \\ & 2y_1 + 2y_2 + y_3 \ge 1 \\ & y_1, y_2, y_3 \ge 0 \end{aligned}$$

Solution: Convert to dual problem to solve it

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

Consider complementary slackness, add slack variables

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

Rewrite objective function as $-x_1 - x_2 + z = 0$

Compute the result

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | z | |
|---|---|---|---|---|---|---|
| 1 | 2 | 1 | 0 | 0 | 0 | 4 |
| 4 | 2 | 0 | 1 | 0 | 0 | 12 |
| -1 | 1 | 0 | 0 | 1 | 0 | 1 |
| -1 | -1 | 0 | 0 | 0 | 1 | 0 |

Iteration 1

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | z | | ratio |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 1 | 0 | 0 | 0 | 4 | 4 |
| 4 | 2 | 0 | 1 | 0 | 0 | 12 | 3 |
| -1 | 1 | 0 | 0 | 1 | 0 | 1 | -1 |
| -1 | -1 | 0 | 0 | 0 | 1 | 0 | |

Choose column1, row2 as the pivot

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | z | | multiplier |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 1 | 0 | 0 | 0 | 4 | -1 |
| 1 | $\frac{1}{2}$ | 0 | $\frac{1}{4}$ | 0 | 0 | 3 | |
| -1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 |
| -1 | -1 | 0 | 0 | 0 | 1 | 0 | 1 |

After elimination

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | z | |
|---|---|---|---|---|---|---|
| 0 | $\frac{3}{2}$ | 1 | $-\frac{1}{4}$ | 0 | 0 | 1 |
| 1 | $\frac{1}{2}$ | 0 | $\frac{1}{4}$ | 0 | 0 | 3 |
| 0 | $\frac{3}{2}$ | 0 | $\frac{1}{4}$ | 1 | 0 | 4 |
| 0 | $-\frac{1}{2}$ | 0 | $\frac{1}{4}$ | 0 | 1 | 3 |

Iteration 2

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | z | | ratio |
|---|---|---|---|---|---|---|---|
| 0 | $\frac{3}{2}$ | 1 | $-\frac{1}{4}$ | 0 | 0 | 1 | $\frac{2}{3}$ |
| 1 | $\frac{1}{2}$ | 0 | $\frac{1}{4}$ | 0 | 0 | 3 | 6 |
| 0 | $\frac{3}{2}$ | 0 | $\frac{1}{4}$ | 1 | 0 | 4 | $\frac{8}{3}$ |
| 0 | $-\frac{1}{2}$ | 0 | $\frac{1}{4}$ | 0 | 1 | 3 | |

Choose column2, row1 as the pivot

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | z | | multiplier |
|---|---|---|---|---|---|---|---|
| 0 | 1 | $\frac{2}{3}$ | $-\frac{1}{6}$ | 0 | 0 | $\frac{2}{3}$ | |
| 1 | $\frac{1}{2}$ | 0 | $\frac{1}{4}$ | 0 | 0 | 3 | $-\frac{1}{2}$ |
| 0 | $\frac{3}{2}$ | 0 | $\frac{1}{4}$ | 1 | 0 | 4 | $-\frac{3}{2}$ |
| 0 | $-\frac{1}{2}$ | 0 | $\frac{1}{4}$ | 0 | 1 | 3 | $\frac{1}{2}$ |

After elimination

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | z | |
|---|---|---|---|---|---|---|
| 0 | 1 | $\frac{2}{3}$ | $-\frac{1}{6}$ | 0 | 0 | $\frac{2}{3}$ |
| 1 | 0 | $-\frac{1}{3}$ | $\frac{1}{3}$ | 0 | 0 | $\frac{8}{3}$ |
| 0 | 0 | -1 | $\frac{1}{2}$ | 1 | 0 | 3 |
| 0 | 0 | $\frac{1}{3}$ | $\frac{1}{6}$ | 0 | 1 | $\frac{10}{3}$ |

The result is $(y_1, y_2, y_3) = (\frac{1}{3}, \frac{1}{6}, 0)$ and the minimum value of $p = \frac{10}{3}$.
Other slack variables are $(y_4, y_5) = (0, 0)$

(c) (10%) Verify the complementarity slackness condition.

Answers are put here.

The dual problem is

$$\begin{aligned} \min_{y_1,y_2,y_3} \quad & p = 4y_1 + 12y_2 + y_3 \\ \text{subject to} \quad & y_1 + 4y_2 - y_3 \geq 1 \\ & 2y_1 + 2y_2 + y_3 \geq 1 \\ & y_1, y_2, y_3 \geq 0 \end{aligned}$$

Consider the complementary slackness

$$\begin{aligned} \min_{y_1,y_2,y_3} \quad & p = 4y_1 + 12y_2 + y_3 \\ \text{subject to} \quad & y_1 + 4y_2 - y_3 - y_4 = 1 \\ & 2y_1 + 2y_2 + y_3 - y_5 = 1 \\ & y_1, y_2, y_3, y_4, y_5 \geq 0 \end{aligned}$$

According to the final table of (b), the result of the dual problem(minimum) is $(y_1, y_2, y_3) = (\frac{1}{3}, \frac{1}{6}, 0)$ and the minimum value of $p = \frac{10}{3}$. Other slack variables are $(y_4, y_5) = (0, 0)$.
As for the primal problem, The result is $(x_1, x_2) = (\frac{8}{3}, \frac{2}{3})$ and the maximum value of $z = \frac{10}{3}$. Other slack variables are $(x_3, x_4, x_5) = (0, 0, 3)$.
Verify the complementary slackness condition $y_i s_i = 0$

$$\begin{aligned} y_1 s_1 &= y_1 x_3 = \tfrac{1}{3} \cdot 0 = 0 \\ y_2 s_2 &= y_2 x_4 = \tfrac{1}{6} \cdot 0 = 0 \\ y_3 s_3 &= y_3 x_5 = 0 \cdot 3 = 0 \end{aligned}$$

The result satisfies the complementary slackness condition.

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

Answers are put here.

Convert to standard form

$$\begin{array}{ll} \min_{x_1,x_2} & z=-x_1-x_2 \\ \text{s.t.} & -x_1-2x_2 \geq -4 \\ & -4x_1-2x_2 \geq -12 \\ & x_1-x_2 \geq -1 \\ & x_1,x_2 \geq 0 \end{array}$$

Then, introduce complementary slackness

$$\begin{array}{ll} \min_{x_1,x_2} & z=-x_1-x_2 \\ \text{s.t.} & -x_1-2x_2-x_3=-4 \\ & -4x_1-2x_2-x_4=-12 \\ & x_1-x_2-x_5=-1 \\ & x_1,x_2,x_3,x_4,x_5 \geq 0 \end{array}$$

(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.
Answers are put here.

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

Consider complementary slackness, add slack variables

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

Rewrite objective function as $-x_1-x_2+z=0$
Compute the result
(1) Method 1: Simplex Algorithm:
Iteration 1

$$B_1=\begin{pmatrix} A_{:,3} & A_{:,4} & A_{:,5} \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, N_1=\begin{pmatrix} A_{:,1} & A_{:,2} \end{pmatrix}=\begin{pmatrix} 1 & 2 \\ 4 & 2 \\ -1 & 1 \end{pmatrix}$$

$$\vec{c}_N=\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}=\begin{pmatrix} -1 \\ -1 \end{pmatrix}, \vec{c}_B=\begin{pmatrix} c_3 \\ c_4 \\ c_5 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\vec{x}_N=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \vec{x}_B=\begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix}=\begin{pmatrix} 4 \\ 12 \\ 1 \end{pmatrix}$$

pricing: $\vec{s}_1=\vec{c}_N-N_1^T(B_1^{-1})^T\vec{c}_B=\begin{pmatrix} -1 \\ -1 \end{pmatrix}-\begin{pmatrix} 1 & 4 & -1 \\ 2 & 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}=\begin{pmatrix} -1 \\ -1 \end{pmatrix},$

Since $\vec{s}_1$ not greater or equal to 0, continue.

$$q_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \vec{d}_1 = B_1^{-1} A_1(:, q_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ -1 \end{pmatrix}$$

search direction: $\begin{pmatrix} d_3 \\ d_4 \\ d_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ -1 \end{pmatrix}$ the whole direction: $\begin{pmatrix} d_3 \\ d_4 \\ d_5 \\ d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ -1 \\ 1 \\ 0 \end{pmatrix}$

ratio test : $\gamma_1, p_1 = 3, 2 = \min_{\gamma_1, p_1 > 0} \frac{\vec{x}_B}{\vec{d}_1} = (\frac{4}{1}, \frac{12}{4}, \frac{1}{-1})$

Swap column 1 and 4. Thus, $\mathcal{B}_1 = \{3, 1, 5\}$ and $\mathcal{N}_1 = \{4, 2\}$

$$\vec{x}_1 = \begin{pmatrix} x_3 \\ x_4 \\ x_5 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 12 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} -1 \\ -4 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 4 \\ 3 \\ 0 \end{pmatrix}$$

Iteration 2

$$B_2 = \begin{pmatrix} A_{:,3} & A_{:,1} & A_{:,5} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 4 & 0 \\ 0 & -1 & 1 \end{pmatrix}, N_2 = \begin{pmatrix} A_{:,4} & A_{:,2} \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 1 & 2 \\ 0 & 1 \end{pmatrix}$$

$$\vec{c}_N = \begin{pmatrix} c_4 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \vec{c}_B = \begin{pmatrix} c_3 \\ c_1 \\ c_5 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

$$\vec{x}_N = \begin{pmatrix} x_4 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \vec{x}_B = \begin{pmatrix} x_3 \\ x_1 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}$$

pricing: $\vec{s}_2 = \vec{c}_N - N_2^T (B_2^{-1})^T \vec{c}_B = \begin{pmatrix} 0 \\ -1 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$

Since $\vec{s}_2$ not greater or equal to 0, continue.

$$= \begin{pmatrix} 0 \\ -1 \end{pmatrix} - \begin{pmatrix} -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{3}{2} & \frac{1}{2} & \frac{3}{2} \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} - \begin{pmatrix} -\frac{1}{4} \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \\ -\frac{1}{2} \end{pmatrix}$$

$$q_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \vec{d}_2 = B_2^{-1} A_2(:, q_2) = \begin{pmatrix} 1 & -\frac{1}{4} & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & \frac{1}{4} & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ \frac{1}{2} \\ \frac{3}{2} \end{pmatrix}$$

search direction: $\begin{pmatrix} d_3 \\ d_1 \\ d_5 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ \frac{1}{2} \\ \frac{3}{2} \end{pmatrix}$ the whole direction: $\begin{pmatrix} d_3 \\ d_1 \\ d_5 \\ d_4 \\ d_2 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ \frac{1}{2} \\ \frac{3}{2} \\ 0 \\ 1 \end{pmatrix}$

ratio test : $\gamma_2, p_2 = \frac{2}{3}, 1 = \min_{\gamma_2, p_2 > 0} \frac{\vec{x}_B}{\vec{d}_2} = (\frac{2}{3}, 6, \frac{8}{3})$

Swap column 2 and 3. Thus, $\mathcal{B}_0 = \{2, 1, 5\}$ and $\mathcal{N}_0 = \{4, 3\}$

$$\vec{x}_2 = \begin{pmatrix} x_3 \\ x_1 \\ x_5 \\ x_4 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 4 \\ 0 \\ 0 \end{pmatrix} + \frac{2}{3} \begin{pmatrix} -\frac{3}{2} \\ -\frac{1}{2} \\ -\frac{3}{2} \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{8}{3} \\ 3 \\ 0 \\ \frac{2}{3} \end{pmatrix}$$

Iteration 3

$$B_3 = \begin{pmatrix} A_{:,2} & A_{:,1} & A_{:,5} \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 \\ 2 & 4 & 0 \\ 1 & -1 & 1 \end{pmatrix}, N_3 = \begin{pmatrix} A_{:,4} & A_{:,3} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\vec{c}_N = \begin{pmatrix} c_4 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \vec{c}_B = \begin{pmatrix} c_2 \\ c_1 \\ c_5 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}$$

$$\vec{x}_N = \begin{pmatrix} x_4 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{2}{3} \end{pmatrix}, \vec{x}_B = \begin{pmatrix} x_3 \\ x_1 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{8}{3} \\ 3 \end{pmatrix}$$

pricing: $\vec{s}_3 = \vec{c}_N - N_3^T (B_3^{-1})^T \vec{c}_B = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -1 \\ -\frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}$

$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} -\frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & -\frac{1}{3} & -1 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} -\frac{1}{6} \\ -\frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{6} \\ \frac{1}{3} \end{pmatrix} \geq 0$$

Since $\vec{s}_3 \geq 0$, return the solution $\vec{x}_3$. Then, we done.

The result is $(x_1, x_2) = (\frac{8}{3}, \frac{2}{3})$ and the maximum value of $z = \frac{10}{3}$. Other slack variables are $(x_3, x_4, x_5) = (0, 0, 3)$.

(2) Method 2: Tabular Method

Initial tabular

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | z | |
|---|---|---|---|---|---|---|
| 1 | 2 | 1 | 0 | 0 | 0 | 4 |
| 4 | 2 | 0 | 1 | 0 | 0 | 12 |
| -1 | 1 | 0 | 0 | 1 | 0 | 1 |
| -1 | -1 | 0 | 0 | 0 | 1 | 0 |

Iteration 1:

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | z | | ratio |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 1 | 0 | 0 | 0 | 4 | 4 |
| 4 | 2 | 0 | 1 | 0 | 0 | 12 | 3 |
| -1 | 1 | 0 | 0 | 1 | 0 | 1 | -1 |
| -1 | -1 | 0 | 0 | 0 | 1 | 0 | |

Choose column1, row2 as the pivot

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | z | | multiplier |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 1 | 0 | 0 | 0 | 4 | -1 |
| 1 | $\frac{1}{2}$ | 0 | $\frac{1}{4}$ | 0 | 0 | 3 | |
| -1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 |
| -1 | -1 | 0 | 0 | 0 | 1 | 0 | 1 |

After elimination

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | z | |
|---|---|---|---|---|---|---|
| 0 | $\frac{3}{2}$ | 1 | $-\frac{1}{4}$ | 0 | 0 | 1 |
| 1 | $\frac{1}{2}$ | 0 | $\frac{1}{4}$ | 0 | 0 | 3 |
| 0 | $\frac{3}{2}$ | 0 | $\frac{1}{4}$ | 1 | 0 | 4 |
| 0 | $-\frac{1}{2}$ | 0 | $\frac{1}{4}$ | 0 | 1 | 3 |

Iteration 2:

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | z | | ratio |
|---|---|---|---|---|---|---|---|
| 0 | $\frac{3}{2}$ | 1 | $-\frac{1}{4}$ | 0 | 0 | 1 | $\frac{2}{3}$ |
| 1 | $\frac{1}{2}$ | 0 | $\frac{1}{4}$ | 0 | 0 | 3 | 6 |
| 0 | $\frac{3}{2}$ | 0 | $\frac{1}{4}$ | 1 | 0 | 4 | $\frac{8}{3}$ |
| 0 | $-\frac{1}{2}$ | 0 | $\frac{1}{4}$ | 0 | 1 | 3 | |

Choose column2, row1 as the pivot

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | z | | multiplier |
|---|---|---|---|---|---|---|---|
| 0 | 1 | $\frac{2}{3}$ | $-\frac{1}{6}$ | 0 | 0 | $\frac{2}{3}$ | |
| 1 | $\frac{1}{2}$ | 0 | $\frac{1}{4}$ | 0 | 0 | 3 | $-\frac{1}{2}$ |
| 0 | $\frac{3}{2}$ | 0 | $\frac{1}{4}$ | 1 | 0 | 4 | $-\frac{3}{2}$ |
| 0 | $-\frac{1}{2}$ | 0 | $\frac{1}{4}$ | 0 | 1 | 3 | $\frac{1}{2}$ |

After elimination

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | z | |
|---|---|---|---|---|---|---|
| 0 | 1 | $\frac{2}{3}$ | $-\frac{1}{6}$ | 0 | 0 | $\frac{2}{3}$ |
| 1 | 0 | $-\frac{1}{3}$ | $\frac{1}{3}$ | 0 | 0 | $\frac{8}{3}$ |
| 0 | 0 | -1 | $\frac{1}{2}$ | 1 | 0 | 3 |
| 0 | 0 | $\frac{1}{3}$ | $\frac{1}{6}$ | 0 | 1 | $\frac{10}{3}$ |

Done

The result is $(x_1, x_2) = (\frac{8}{3}, \frac{2}{3})$ and the maximum value of $z = \frac{10}{3}$. Other slack variables are $(x_3, x_4, x_5) = (0, 0, 3)$.

Reference

Simplex Method-Minimization Problem-Part 2
Introduction to the Simplex Method: Standard Maximization (2 variables)

| | |
|---|---|
| (1) | Given a basic feasible point $\vec{x}_0$ and the corresponding index set $\mathcal{B}_0$ and $\mathcal{N}_0$. |
| (2) | For $k = 0, 1, \ldots$ |
| (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 $\vec{s}_k(i_q) < 0$, where $i_q$ 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 $[\gamma_k, i_p] = \min\limits_{i, \vec{d}_k(i) > 0} \dfrac{\vec{x}_B(i)}{\vec{d}_k(i)}$ (ratio test) (The first return value is the minimum ratio; the second return value is the index of the minimum ratio.) |
| (10) | $x_{k+1} \begin{pmatrix} \mathcal{B} \\ \mathcal{N} \end{pmatrix} = \begin{pmatrix} \vec{x}_B \\ \vec{x}_N \end{pmatrix} + \gamma_k \begin{pmatrix} -\vec{d}_k \\ \vec{e}_{i_q} \end{pmatrix}$ ($\vec{e}_{i_q} = (0, \ldots, 1, \ldots, 0)^T$ is a unit vector with $i_q$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 1: The simplex method for solving (minimization) linear programming