

## Math 164 - Assignment #6, Fall 2024

- **Due Date:** 11:59 pm on November 15, 2024
- 

1. Suppose that we perform an experiment to calculate the gravitational constant  $g$  as follows. We drop a ball from a certain height and measure its distance from the original point at certain time instants. The results of the experiment are shown in the following table.

Time (s)	1.00	2.00	3.00
Distance (m)	5.00	19.5	44.0

The equation relating the distance  $s$  and the time  $t$  at which  $s$  is measured is given by

$$s = \frac{1}{2}gt^2.$$

- (a) Find a least-squares estimate of  $g$  using the experimental results from the table above.
  - (b) Suppose that we take an additional measurement at time 4.00 and obtain a distance of 78.5. Use the RLS algorithm to calculate an updated least-squares estimate of  $g$ .
2. Let  $[x_1, y_1]^\top, \dots, [x_p, y_p]^\top, p \geq 2$ , be points in  $\mathbb{R}^2$ . We wish to find the straight line of best fit through these points ("best" in the sense that the total squared error is minimized); that is, we wish to find  $a^*, b^* \in \mathbb{R}$  to minimize

$$f(a, b) = \sum_{i=1}^p (ax_i + b - y_i)^2.$$

Assume that the  $x_i, i = 1, \dots, p$ , are not all equal. Show that there exist unique parameters  $a^*$  and  $b^*$  for the line of best fit, and find the parameters in terms of the following quantities:

$$\begin{aligned}\bar{X} &= \frac{1}{p} \sum_{i=1}^p x_i \\ \bar{Y} &= \frac{1}{p} \sum_{i=1}^p y_i \\ \overline{X^2} &= \frac{1}{p} \sum_{i=1}^p x_i^2 \\ \overline{Y^2} &= \frac{1}{p} \sum_{i=1}^p y_i^2 \\ \overline{XY} &= \frac{1}{p} \sum_{i=1}^p x_i y_i.\end{aligned}$$

3. We are given a point  $[x_0, y_0]^\top \in \mathbb{R}^2$ . Consider the straight line on the  $\mathbb{R}^2$  plane given by the equation  $y = mx$ . Using a least-squares formulation, find the point on the straight line that is closest to the given point  $[x_0, y_0]$ , where the measure of closeness is in terms of the Euclidean norm on  $\mathbb{R}^2$ . *Hint:* The given line can be expressed as the range of the matrix  $\mathbf{A} = [1, m]^\top$ .

4. Let  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{b} \in \mathbb{R}^m$ ,  $m \geq n$ , and  $\text{rank } \mathbf{A} = n$ . Consider the constrained optimization problem

$$\begin{aligned} & \text{minimize} && \frac{1}{2} \mathbf{x}^\top \mathbf{x} - \mathbf{x}^\top \mathbf{b} \\ & \text{subject to} && \mathbf{x} \in \mathcal{R}(\mathbf{A}), \end{aligned}$$

where  $\mathcal{R}(\mathbf{A})$  denotes the range of  $\mathbf{A}$ . Derive an expression for the global minimizer of this problem in terms of  $\mathbf{A}$  and  $\mathbf{b}$ .

5. Let  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{b} \in \mathbb{R}^m$ ,  $m \leq n$ ,  $\text{rank } \mathbf{A} = m$ , and  $\mathbf{x}_0 \in \mathbb{R}^n$ . Consider the problem

$$\begin{aligned} & \text{minimize} && \|\mathbf{x} - \mathbf{x}_0\| \\ & \text{subject to} && \mathbf{Ax} = \mathbf{b} \end{aligned}$$

Show that this problem has a unique solution given by

$$\mathbf{x}^* = \mathbf{A}^\top (\mathbf{AA}^\top)^{-1} \mathbf{b} + \left( \mathbf{I}_n - \mathbf{A}^\top (\mathbf{AA}^\top)^{-1} \mathbf{A} \right) \mathbf{x}_0$$

6. Prove the following properties of generalized inverses:

(a)  $(\mathbf{A}^\top)^\dagger = (\mathbf{A}^\dagger)^\top$ .

(b)  $(\mathbf{A}^\dagger)^\dagger = \mathbf{A}$ .

1. Suppose that we perform an experiment to calculate the gravitational constant  $g$  as follows. We drop a ball from a certain height and measure its distance from the original point at certain time instants. The results of the experiment are shown in the following table.

Time (s)	1.00	2.00	3.00
Distance (m)	5.00	19.5	44.0

The equation relating the distance  $s$  and the time  $t$  at which  $s$  is measured is given by

$$s = \frac{1}{2}gt^2.$$

- (a) Find a least-squares estimate of  $g$  using the experimental results from the table above.
- (b) Suppose that we take an additional measurement at time 4.00 and obtain a distance of 78.5. Use the RLS algorithm to calculate an updated least-squares estimate of  $g$ .

**Solution:**

Recall that Theorem 12.1 states that the unique vector  $x^*$  that minimizes  $\|Ax - b\|^2$  is given by the solution to the equation

$$A^T Ax = A^T b \implies x^* = (A^T A)^{-1} A^T b$$

To solve this problem, we must first write it in the form of a least squares problem.

$$s = \frac{t^2}{2}g$$

$$A = \begin{bmatrix} 1/2 \\ 2 \\ 9/2 \end{bmatrix}$$

$$b = \begin{bmatrix} 5 \\ 19.5 \\ 44 \end{bmatrix}$$

So the least squares estimate of  $g$  is now given by

$$g^* = ([1/2 \quad 2 \quad 9/2] \begin{bmatrix} 1/2 \\ 2 \\ 9/2 \end{bmatrix})^{-1} [1/2 \quad 2 \quad 9/2] \begin{bmatrix} 5 \\ 19.5 \\ 44 \end{bmatrix} = 9.776$$

2. Let  $[x_1, y_1]^\top, \dots, [x_p, y_p]^\top$ ,  $p \geq 2$ , be points in  $\mathbb{R}^2$ . We wish to find the straight line of best fit through these points (“best” in the sense that the total squared error is minimized); that is, we wish to find  $a^*, b^* \in \mathbb{R}$  to minimize

$$f(a, b) = \sum_{i=1}^p (ax_i + b - y_i)^2.$$

Assume that the  $x_i$ ,  $i = 1, \dots, p$ , are not all equal. Show that there exist unique parameters  $a^*$  and  $b^*$  for the line of best fit, and find the parameters in terms of the following quantities:

$$\bar{X} = \frac{1}{p} \sum_{i=1}^p x_i, \quad \bar{Y} = \frac{1}{p} \sum_{i=1}^p y_i, \quad \overline{X^2} = \frac{1}{p} \sum_{i=1}^p x_i^2,$$

$$\overline{Y^2} = \frac{1}{p} \sum_{i=1}^p y_i^2, \quad \overline{XY} = \frac{1}{p} \sum_{i=1}^p x_i y_i.$$

**Solution:**

We can rewrite the problem as:

$$\begin{aligned} f(a, b) &= \sum_{i=1}^p (ax_i + b - y_i)^2 \\ &= \left\| \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_p & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} \right\|^2 = \|Ax - b\|^2 \end{aligned}$$

By Theorem 12.1, the matrix  $A$  must be invertible. Since  $x_i$ ’s are distinct, the columns of the matrix are linearly independent, making it full rank and thus invertible.

Therefore, there exists a unique pair of parameters  $a^*, b^*$  that minimizes the total squared error.

Now, let’s apply theorem 12.1 to find the minimizer  $x^*$

$$\begin{aligned} x^* &= (A^T A)^{-1} A^T b \\ A &= \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_p & 1 \end{bmatrix} \\ b &= \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} \\ x^* &= \begin{pmatrix} \overline{X^2} & \bar{X} \\ \bar{X} & 1 \end{pmatrix}^{-1} \begin{bmatrix} \overline{XY} \\ \bar{Y} \end{bmatrix} \\ \begin{bmatrix} a^* \\ b^* \end{bmatrix} &= \frac{1}{\overline{X^2} - \bar{X}^2} \begin{bmatrix} \overline{XY} - (\bar{X})(\bar{Y}) \\ \overline{X^2 Y} - (\bar{X})\overline{XY} \end{bmatrix} \end{aligned}$$

3. We are given a point  $[x_0, y_0]^\top \in \mathbb{R}^2$ . Consider the straight line on the  $\mathbb{R}^2$  plane given by the equation  $y = mx$ . Using a least-squares formulation, find the point on the straight line that is closest to the given point  $[x_0, y_0]$ , where the measure of closeness is in terms of the Euclidean norm on  $\mathbb{R}^2$ .

**Hint:** The given line can be expressed as the range of the matrix  $A = [1, m]^\top$ .

**Solution:**

By the hint, we know that every point on the line can be expressed as

$$v = cA = c \begin{bmatrix} 1 \\ m \end{bmatrix}, c \in \mathbb{R}$$

We can now transform the given problem into the form  $\|Ax - b\|^2$ , and apply Theorem 12.1

The measure of closeness is in terms of the Euclidean norm on  $\mathbb{R}^2$ , which is given by

$$\|x - x_0\|^2 = \left\| \begin{bmatrix} 1 \\ m \end{bmatrix} c - \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \right\|^2$$

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

$$A = \begin{bmatrix} 1 \\ m \end{bmatrix}$$

$$b = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

$$c^* = \frac{x_0 + y_0 m}{1 + m^2}$$

$$v^* = c^* A = \begin{bmatrix} \frac{x_0 + y_0 m}{1 + m^2} \\ \frac{m(x_0 + y_0 m)}{1 + m^2} \end{bmatrix}$$

4. Let  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $m \geq n$ , and  $\text{rank} A = n$ . Consider the constrained optimization problem

$$\text{minimize } \frac{1}{2} x^\top x - x^\top b \quad \text{subject to } x \in R(A),$$

where  $R(A)$  denotes the range of  $A$ . Derive an expression for the global minimizer of this problem in terms of  $A$  and  $b$ .

**Solution:**

We can rewrite the problem as a least squares problem, and then apply Theorem 12.1

Observe that

$$\frac{1}{2} x^\top x - x^\top b = \frac{1}{2} \|x - b\|^2$$

And if  $x \in R(A)$ , then  $\exists y \in \mathbb{R}^n$  such that  $x = Ay$

The problem is now to find the unique vector  $x^*$  that minimizes:  $\frac{1}{2} \|Ay - b\|^2$

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

$$x^* = Ay^* = A (A^T A)^{-1} A^T b$$

Note: another interesting way to solve this problem is to express the objective function in terms of  $y$ , and minimize it by setting the gradient to 0 and solving for  $y^*$ .

5. Let  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $m \leq n$ ,  $\text{rank } A = m$ , and  $x_0 \in \mathbb{R}^n$ . Consider the problem

$$\text{minimize } \|x - x_0\| \quad \text{subject to } Ax = b.$$

Show that this problem has a unique solution given by

$$x^* = A^\top (AA^\top)^{-1} b + \left( I_n - A^\top (AA^\top)^{-1} A \right) x_0.$$

**Solution:**

To solve this, we will rewrite the problem as a system of linear equations minimize  $\|x\|$  subject to  $Ax = b$ . We can then apply Theorem 12.2 Let

$$y = x - x_0$$

$$Ax = A(y + x_0) = b \implies Ay = b - Ax_0$$

Then the problem is now to find the unique vector  $y^*$  that minimizes:  $\|Ay = b - Ax_0\|$

By Theorem 12.1, the unique solution  $x^*$  to  $Ax = b$  that minimizes the norm  $\|x\|$  is given by

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

Applying this theorem to the system  $Ay = b - Ax_0$ , we find that

$$\begin{aligned} y^* &= A^T (AA^T)^{-1} (b - Ax_0) \\ x^* &= y^* + x_0 = A^T (AA^T)^{-1} (b - Ax_0) + x_0 \\ &= A^T (AA^T)^{-1} b - A^T (AA^T)^{-1} Ax_0 + x_0 \\ &= A^T (AA^T)^{-1} b + (I_n - A^T (AA^T)^{-1} A) x_0 \end{aligned}$$

6. Prove the following properties of generalized inverses:

(a)  $(A^\top)^\dagger = (A^\dagger)^\top.$

(b)  $(A^\dagger)^\dagger = A.$

**Solution:**

- (a) Recall that the pseudoinverse,  $A^\dagger \in \mathbb{R}^{m \times n}$  of  $A$  can be defined as

$$A^\dagger = (A^T A)^{-1} A^T = (A A^T)^{-1} A^T$$

So, using this definition, we have:

$$(A^T)^\dagger = (A A^T)^{-1} A^T$$

$$(A^\dagger)^T = (A A^T)^{-1} A^T$$

$$(A^T)^\dagger = (A^\dagger)^T$$

- (b)