Closed-form Solution for Least Squares Estimation (LSE)

Problem Definition

We aim to solve a linear equation:

$$Aw = b, \tag{1}$$

where:

- A is an $m \times n$ matrix.
- w is an $n \times 1$ parameter vector to be solved.
- b is an $m \times 1$ vector.

1 Case 1: When A is Invertible

If A is a square and invertible matrix, the solution can be directly computed as:

$$w = A^{-1}b. \tag{2}$$

However, in most cases, A is not invertible or even not a square matrix. Therefore, we need to use the **Least Squares Estimation (LSE)** method.

2 Case 2: Least Squares Problem

When A is not a square matrix or not invertible, we minimize the squared error:

$$\min_w \|Aw - b\|^2. \tag{3}$$

Solution: Normal Equations

By setting the gradient to zero, we obtain the **normal equations**:

$$A^T Aw = A^T b. \tag{4}$$

If $A^T A$ is invertible, the solution is:

$$w = (A^T A)^{-1} A^T b. \tag{5}$$

3 Case 3: Handling Singular $A^T A$ (Regularization)

If $A^T A$ is singular (i.e., not invertible), we use **regularization**:

$$w = (A^T A + \lambda I)^{-1} A^T b. \tag{6}$$

where $\lambda > 0$ ensures that $A^T A + \lambda I$ is positive definite and invertible.

LU Decomposition for Matrix Inversion

Using LU Decomposition to Compute A^{-1}

Instead of computing A^{-1} directly, we can use **LU decomposition** to efficiently solve for the inverse.

Step 1: LU Factorization

If A is an $n \times n$ invertible matrix, we can decompose it as:

$$A = LU, \tag{7}$$

where:

- L is a lower triangular matrix with 1s on the diagonal.
- U is an upper triangular matrix.

Step 2: Solving for A^{-1}

To compute A^{-1}, we solve:

$$AX = I. \tag{8}$$

Substituting $A = LU$, we obtain:

$$LUX = I. \tag{9}$$

Let:

$$UX = Y. \tag{10}$$

Then, we need to solve:

1. Solve $LY = I$ using **forward substitution**.
2. Solve $UX = Y$ using **backward substitution**.

The resulting X is the inverse A^{-1}.

Step 3: Forward Substitution for $LY = I$

Since L is a lower triangular matrix, we can solve for Y row by row. If L is:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix}, \tag{11}$$

and we are solving:

$$LY = I, \tag{12}$$

then for each column e_j of I, the corresponding column y_j of Y satisfies:

$$\begin{aligned} y_{1j} &= 1, \\ y_{2j} &= 1 - l_{21}y_{1j}, \\ y_{3j} &= 1 - l_{31}y_{1j} - l_{32}y_{2j}. \end{aligned} \tag{13}$$

This process proceeds from top to bottom (**forward substitution**).

Step 4: Backward Substitution for $UX = Y$

Once Y is determined, we solve:

$$UX = Y. \tag{14}$$

Since U is an upper triangular matrix:

$$U = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}, \tag{15}$$

the elements of X are computed using:

$$\begin{aligned} u_{33}x_{3j} &= y_{3j}, \\ u_{22}x_{2j} + u_{23}x_{3j} &= y_{2j}, \\ u_{11}x_{1j} + u_{12}x_{2j} + u_{13}x_{3j} &= y_{1j}. \end{aligned} \tag{16}$$

This process proceeds from bottom to top (**backward substitution**).

Final Result

After solving for all x_{ij}, we obtain the inverse:

$$A^{-1} = X. \tag{17}$$

Summary: Instead of computing A^{-1} directly, we solve two simple triangular systems:

1. $LY = I$ (forward substitution)
2. $UX = Y$ (backward substitution)

This method is numerically more stable and efficient than computing A^{-1} directly.

Gauss-Jordan

Proof: Gauss-Jordan Elimination Can Find A^{-1}

Since we have already established that A is invertible, we now show that the inverse A^{-1} can be found using **Gauss-Jordan elimination**.

Constructing the Augmented Matrix

The standard method for computing A^{-1} involves forming the augmented matrix:

$$[A \quad | \quad I]\,. \tag{18}$$

where A is an $n \times n$ invertible matrix and I is the identity matrix of the same size.

Row Reduction to Identity Matrix

Using elementary row operations, we perform Gaussian elimination to transform A into the identity matrix. The key steps are:

- Swap rows if necessary to obtain a nonzero pivot element.
- Scale the pivot row so that the leading coefficient becomes 1.
- Subtract multiples of the pivot row from other rows to create zeros below and above the pivot.

This row reduction process transforms the augmented matrix:

$$[A \quad | \quad I] \quad \longrightarrow \quad \left[I \quad | \quad A^{-1}\right]. \tag{19}$$

Summary

Since A is invertible, the row operations will successfully reduce A to I, and the right-hand side will yield A^{-1}. Therefore, we conclude that:

$$A^{-1} \text{ can be found using Gauss-Jordan elimination.} \tag{20}$$

Conclusion

(1) Direct Inversion (If A is Invertible)

$$w = A^{-1}b. \tag{21}$$

(2) Least Squares Solution (If A is Non-Square)

$$w = (A^T A)^{-1} A^T b. \tag{22}$$

(3) Regularized Least Squares (If $A^T A$ is Singular)

$$w = (A^T A + \lambda I)^{-1} A^T b. \tag{23}$$