MA-110 Linear Algebra

Nazar Khan

PUCIT

13. Least Squares Fitting

Pseudoinverse Matrix

- Consider the linear system $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A}$ is an $m \times n$ matrix.
- If $m = n$ and $\mathbf{A}$ is invertible, then $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.
- However, if $m \neq n$, then

$$\mathbf{Ax} = \mathbf{b} \implies \mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b} \implies \mathbf{x} = \underbrace{(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T}_{\mathbf{A}^\dagger}\mathbf{b} \quad (1)$$

where the $n \times m$ matrix $\mathbf{A}^\dagger = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ is called the *pseudoinverse* of $\mathbf{A}$.

Least Squares Fitting

- For both cases of $m = n$ and $m \neq n$, there is a possibility that the linear system $\mathbf{Ax} = \mathbf{b}$ is inconsistent.
- This happens very often in real-world applications where there is noise in measurements used to form $\mathbf{A}$ and/or $\mathbf{b}$.
- In such cases, $\mathbf{Ax} \approx \mathbf{b}$ and it makes more sense to find an $\mathbf{x}$ that minimizes the length of the *error vector* $\mathbf{Ax} - \mathbf{b}$.
- That is, the *best possible* $\mathbf{x}$ should be the one for which $\|\mathbf{Ax} - \mathbf{b}\|$, or equivalently, $\|\mathbf{Ax} - \mathbf{b}\|^2$ is the smallest.
- Such a solution for $\mathbf{x}$ is called the *least squares solution*.

Least Squares Fitting

- If $\mathbf{Ax} \neq \mathbf{b}$ then $\mathbf{b}$ cannot lie in the column space of $\mathbf{A}$.
- But $\mathbf{Ax}$ is always a linear combination of the columns of $\mathbf{A}$. Hence it lies in col($\mathbf{A}$).
- Vector $\mathbf{Ax}$ is closest to $\mathbf{b}$ when it is the projection of $\mathbf{b}$ onto col($\mathbf{A}$). That is,

$$\mathbf{Ax} = \mathbf{proj}_{\text{col}(\mathbf{A})}\mathbf{b} \tag{2}$$

$$\implies \mathbf{b} - \mathbf{Ax} = \mathbf{b} - \mathbf{proj}_{\text{col}(\mathbf{A})}\mathbf{b} \tag{3}$$

$$\implies \mathbf{A}^T(\mathbf{b} - \mathbf{Ax}) = \mathbf{A}^T(\mathbf{b} - \mathbf{proj}_{\text{col}(\mathbf{A})}\mathbf{b}) \tag{4}$$

where the R.H.S is equal to $\mathbf{0}$ since $\mathbf{b} - \mathbf{proj}_{\text{col}(\mathbf{A})}\mathbf{b})$ is orthogonal to col($\mathbf{A}$).

$$\implies \mathbf{A}^T(\mathbf{b} - \mathbf{Ax}) = \mathbf{0} \implies \mathbf{A}^T\mathbf{b} = \mathbf{A}^T\mathbf{x} \tag{5}$$

$$\implies \mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} = \mathbf{A}^\dagger\mathbf{b} \tag{6}$$

- So, the least squares solution is also obtained via the pseudoinverse.