<div align="center">

EE2100: Matrix Analysis

Review Notes - 24

</div>

Topics covered :

1. Overdetermined system of linear equations

---

1. A system of linear equations where the number of equations (say m) are greater than the number of unknowns (say n) are commonly referred to as overdetermined system of linear equations. An overdetermined system of linear equations can be represented as

$$\mathbf{Ax} = \mathbf{b} \text{ where } \mathbf{A} \in \mathcal{R}^{m \times n}, \ \mathbf{x} \in \mathcal{R}^n \text{ and } \mathbf{b} \in \mathcal{R}^m \tag{1}$$

2. The system of linear equations $\mathbf{Ax} = \mathbf{b}$, have at least one solution if $\mathbf{b} \in \text{Col}(\mathbf{A})$. The column space of the matrix is $\text{Span}\{\mathbf{a}_1, \cdots \mathbf{a}_n\}$ where $\mathbf{a} \in \mathcal{R}^m$. Since the number of column vectors are less than m, they cannot span entire $\mathcal{R}^m$ (even if the column vectors are linearly independent). As a result, $\mathbf{Ax} = \mathbf{b}$ may or may not have any solution for any random $\mathbf{b} \in \mathcal{R}^m$.

3. The maximum rank of the matrix $\mathbf{A} \in \mathcal{R}^{m \times n}$ is n (when all the column vectors are linearly independent) and the minimum nullity is 0 (by rank-nullity theorem).

4. Since, $\mathbf{Ax} = \mathbf{b}$ may or may not have any solution for any random $\mathbf{b} \in \mathcal{R}^m$, there is a need to analyze two different scenarios.

   (a) When $\mathbf{b} \in \text{Col}(\mathbf{A})$: In this scenario, the system of linear equations have at least one solution.

   (b) When $\mathbf{b} \notin \text{Col}(\mathbf{A})$: In this scenario, the system of linear equations do not have any solution. However, in most of the scenarios, it is often of practical interest to obtain an approximate solution (commonly referred to as least squares solution).

5. **Approximate or least squares solution**: An approximate solution to overdetermined system of linear equations when $\mathbf{b} \notin \text{Col}(\mathbf{A})$ can be obtained by first projecting $\mathbf{b}$ into the $\text{Col}(\mathbf{A})$ (say $\hat{\mathbf{b}}$) and subsequently finding the solution to $\mathbf{A}\hat{\mathbf{x}} = \hat{\mathbf{b}}$ (where $\hat{\mathbf{x}}$ is the approximated solution).

   If $\hat{\mathbf{b}}$ denotes the projection of $\mathbf{b}$ onto the $\text{Col}(\mathbf{A})$, then $\mathbf{b} - \hat{\mathbf{b}} \perp \text{Col}(\mathbf{A})$ (i.e., to every vector in the $\text{Col}(\mathbf{A})$). Accordingly,

$$\begin{aligned}
\mathbf{b} - \hat{\mathbf{b}} \perp \mathbf{a}_1 &\implies \mathbf{a}_1^T\left(\mathbf{b} - \hat{\mathbf{b}}\right) = 0 \\
\mathbf{b} - \hat{\mathbf{b}} \perp \mathbf{a}_2 &\implies \mathbf{a}_2^T\left(\mathbf{b} - \hat{\mathbf{b}}\right) = 0 \\
&\vdots \qquad\qquad = \vdots \\
\mathbf{b} - \hat{\mathbf{b}} \perp \mathbf{a}_n &\implies \mathbf{a}_n^T\left(\mathbf{b} - \hat{\mathbf{b}}\right) = 0
\end{aligned} \tag{2}$$

The set of equations given by (2) can be represented as

$$
\begin{bmatrix}
-- & \mathbf{a}_1^T & -- \\
-- & \mathbf{a}_2^T & -- \\
-- & \vdots & -- \\
-- & \mathbf{a}_n^T & --
\end{bmatrix}
\begin{bmatrix}
b_1 - \hat{b}_1 \\
b_2 - \hat{b}_2 \\
\vdots \\
b_n - \hat{b}_n
\end{bmatrix}
=
\begin{bmatrix}
0 \\
0 \\
\vdots \\
0
\end{bmatrix}
\tag{3}
$$

The approximate solution $\hat{\mathbf{x}}$ can be computed as

$$
\mathbf{A}^T \left(\mathbf{b} - \hat{\mathbf{b}} \right) = \mathbf{0}
$$
$$
\mathbf{A}^T \left(\mathbf{b} - \mathbf{A}\hat{\mathbf{x}} \right) = \mathbf{0} \tag{4}
$$
$$
\left(\mathbf{A}\mathbf{A}^T \right) \hat{\mathbf{x}} = \mathbf{A}^T \mathbf{b}
$$

The matrix $\mathbf{A}^T\mathbf{A} \in \mathcal{R}^{n \times n}$ is often referred to as Gram matrix. The approximate solution is obtained by solving the system of linear equations (either by Gaussian Elimination or LU decomposition) given by $\left(\mathbf{A}\mathbf{A}^T \right) \hat{\mathbf{x}} = \mathbf{A}^T \mathbf{b}$.

6. It can be easily shown that the Gram matrix (i.e., $\mathbf{A}^T\mathbf{A}$) has the following properties.

   - The null space/nullity of $\mathbf{A}^T\mathbf{A}$ is the same as null space/nullity of $\mathbf{A}$.

   - The matrix $\mathbf{A}^T\mathbf{A}$ is a full ranked matrix if the columns of $\mathbf{A}$ are linearly independent.

7. An alternate way to figure out the approximate solution is by solving the least squares optimization problem given by

$$
\hat{\mathbf{x}} : \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 \tag{5}
$$

8. It is important to note that the system of linear equations $\left(\mathbf{A}\mathbf{A}^T \right) \hat{\mathbf{x}} = \mathbf{A}^T \mathbf{b}$ also gives a solution (exact in this case) to $\mathbf{A}\mathbf{x} = \mathbf{b}$ when $\mathbf{b} \in \text{Col}\left(\mathbf{A} \right)$.