Gram and Design Matrix

The **Gram Matrix** is:

$$Gram = X^T X$$

Where X is the **Design Matrix**:

- Its a square $p \times p$ matrix
- Every value of $Gram_{ij}$ is the dot product between column i and column j of X
- Captures how correlated the Predictors are with each other
- Used in the Hat Matrix

The **Design Matrix** is:

- Its an $n \times p$
- n: number of Observation(rows)
- p: number of features or variables (columns)
- the intercept column β_0 isn't included in the columns p, Its $k + p$
- Role in Multiple Linear Regression
    - X holds the data ,β is what we solve for
    - We use the Design matrix in both Training aka fitting β and prediction aka new unseen data

Conclusion:

- **Design Matrix** $\rightarrow$ our raw data arranged in a matrix for the ease of modeling and interpretation
- **Gram Matrix** $X^T X \rightarrow$ Derived from X useful for computing the solutions and estimations for linear regression