Extrapolation:

$$x^* \notin \text{Range}[X]$$

where

$$\text{Range}[X] = [X_{\cdot 1,\min}, X_{\cdot 1,\max}] \times [X_{\cdot 2,\min}, X_{\cdot 2,\max}] \times \cdots \times [X_{\cdot p,\min}, X_{\cdot p,\max}]$$

each being the interval of the min and max of the ith prediction.

Let $\vec{y} \in \mathbb{R}^n$ and $X \in \mathbb{R}^{n\times(p+1)}$. The linear multivariate least squares fitting is given by

$$\vec{b} = (X^T X)^{-1} X^T \vec{y}$$

such that

$$\hat{\vec{y}} = X\vec{b} = \begin{bmatrix} \leftarrow & \vec{X}_1 & \rightarrow \\ \leftarrow & \vec{X}_2 & \rightarrow \\ & \vdots & \\ \leftarrow & \vec{X}_n & \rightarrow \end{bmatrix} \vec{b} = X(X^T X)^{-1} X^T \vec{y} = H\vec{y}$$

The rank of X is $p+1$ because $\hat{y}_i = b_0 + b_1 x_{i1} + \cdots + b_p x_{ip}$ which lines in a $p+1$ dimension space. We call $p+1$ the degrees of freedom because it is the dimension of the column space of X.

Recall that $\vec{y} = g(\vec{x}) + \vec{e} = \hat{\vec{y}} + \vec{e}$, Therefore $\vec{e} = \vec{y} - \hat{\vec{y}}$.

$$SSE = \sum_{i=1}^{n} e_i^2 = \|\vec{e}\|^2$$

Furthermore,

$$MSE = \frac{1}{n-(p+1)} = SSE$$

where $p+1$ is the degree of freedom. Then

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n-(p+1)} SSE} = \frac{\|\vec{e}\|}{\sqrt{n-(p+1)}}$$

Same as before,

$$R^2 = \frac{S_y^2 - S_e^2}{S_y^2}$$

Orthogonal Projection: Project two vectors in $\mathbb{R}^d$ separated by an angle θ between a point joining it. By dropping $\vec{a}$ vertically down on $\vec{v}$, that is the $\text{proj}_{\vec{v}}(\vec{a})$. Its length will be called l. By the law of cosines,

$$\cos\theta = \frac{\vec{a}\cdot\vec{c}}{\|\vec{a}\|\|\vec{v}\|}$$

Therefore

$$l = \|\vec{a}\|\cos\theta = \frac{\vec{a}\cdot\vec{v}}{\|\vec{v}\|}$$

Then

$$\text{proj}_{\vec{v}}(\vec{a}) = l\frac{\vec{v}}{\|\vec{v}\|} = \frac{\vec{a}\cdot\vec{v}}{\|\vec{v}\|^2}\vec{v} = \frac{\vec{a}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v} = \frac{\vec{v}\vec{v}^T\vec{a}}{\vec{v}^T\vec{v}} = \frac{\vec{v}\vec{v}^T}{\vec{v}^T\vec{v}}\vec{a} = H\vec{a}$$

Note that

$$HH = \left(\frac{\vec{v}\cdot\vec{v}^T}{\|\vec{v}\|^2}\right)\left(\frac{\vec{v}\vec{v}^T}{\|\vec{v}\|^2}\right) = \frac{\vec{v}\|\vec{v}\|^2\vec{v}^T}{\|\vec{v}\|^4} = \frac{\vec{v}\vec{v}^T}{\|\vec{v}\|^2} = H$$

Now let's project $\vec{a}$ onto $\mathbb{V} = \text{colsp}(V)$, where V is linearly independent ($V = \begin{bmatrix}\vec{v}_1 & \dots & \vec{v}_k\end{bmatrix}$). Then there exists $\vec{w}$ such that

$$\text{proj}_{\vec{v}}(\vec{a}) = V\vec{w}$$