Brief Notes #7
Conditional Second-Moment Analysis

- **Important result for jointly normally distributed variables X_1 and X_2**

If X_1 and X_2 are jointly normally distributed with mean values m_1 and m_2, variances σ_1^2 and σ_2^2, and correlation coefficient ρ, then $(X_1 \mid X_2 = x_2)$ is also normally distributed with mean and variance:

$$\begin{cases} m_{1|2}(x_2) = m_1 + \rho \dfrac{\sigma_1}{\sigma_2}(x_2 - m_2) \\ \\ \sigma_{1|2}^2(x_2) = \sigma_1^2(1-\rho^2) \end{cases} \tag{1}$$

Notice that the conditional variance does not depend on x_2.

The results in Eq. 1 hold strictly when X_1 and X_2 are jointly normal, but may be used in approximation for other distributions or when one knows only the first two moments of the vector $\underline{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$.

- **Extension to many observations and many predictions**

Let $\underline{X} = \begin{bmatrix} \underline{X}_1 \\ \underline{X}_2 \end{bmatrix}$, where $\underline{X}_1$ and $\underline{X}_2$ are sub-vectors of $\underline{X}$. Suppose $\underline{X}$ has multivariate normal distribution with mean value vector and covariance matrix:

$$\underline{m} = \begin{bmatrix} \underline{m}_1 \\ \underline{m}_2 \end{bmatrix}, \quad \text{and} \quad \underline{\Sigma} = \begin{bmatrix} \underline{\Sigma}_{11} & \underline{\Sigma}_{12} \\ \underline{\Sigma}_{21} & \underline{\Sigma}_{22} \end{bmatrix} \qquad (\underline{\Sigma}_{12} = \underline{\Sigma}_{21}^T).$$

Then, given $\underline{X}_2 = \underline{x}_2$, the conditional vector $(\underline{X}_1 \mid \underline{X}_2 = \underline{x}_2)$ has jointly normal distributions with parameters:

$$\begin{cases} \underline{m}_{1|2}(\underline{x}_2) = \underline{m}_1 + \underline{\Sigma}_{12}\underline{\Sigma}_{22}^{-1}(\underline{x}_2 - \underline{m}_2) \\ \\ \underline{\Sigma}_{1|2}(\underline{x}_2) = \underline{\Sigma}_{11} - \underline{\Sigma}_{12}\underline{\Sigma}_{22}^{-1}\underline{\Sigma}_{12}^T \end{cases} \tag{2}$$

Notice again that $\underline{\Sigma}_{1|2}$ does not depend on $\underline{x}_2$.

As for the scalar case, Eq. 2 may be used in approximation when $\underline{X}$ does not have multivariate normal distribution or when the distribution of $\underline{X}$ is not known, except for the mean vector $\underline{m}$ and covariance matrix $\underline{\Sigma}$.