ADVANCED BAYESIAN MODELING

Matrix Formulation and Normality

Matrix Form

Convenient to let

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1k} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nk} \end{pmatrix} \qquad y = \begin{pmatrix} y_1 \\ \vdots \\ \vdots \\ y_n \end{pmatrix} \qquad \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$$

Then

$$X_i = \begin{pmatrix} x_{i1} & \cdots & x_{ik} \end{pmatrix} \quad \text{is row } i \text{ of } X$$

and linear regression becomes

$$\mathrm{E}(y_i \mid \theta, X_i) = X_i \beta$$

The full regression is given by

$$\mathrm{E}(y \mid \theta, X) \;=\; \begin{pmatrix} \mathrm{E}(y_1 \mid \beta, X_1) \\ \vdots \\ \vdots \\ \mathrm{E}(y_n \mid \beta, X_n) \end{pmatrix} \;=\; X\beta$$

This is the (conditional) **mean vector** of y.

For example, in a *simple* linear regression of y on scalar x,

$$X \;=\; \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \qquad \beta \;=\; \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \qquad \text{and} \qquad X\beta \;=\; \begin{pmatrix} \beta_1 + \beta_2 x_1 \\ \vdots \\ \vdots \\ \beta_1 + \beta_2 x_n \end{pmatrix}$$

The (conditional) **covariance matrix** of y is

$$\mathrm{var}(y \mid \theta, X) \;=\; \begin{pmatrix} \mathrm{var}(y_1 \mid \theta, X) & \mathrm{cov}(y_1, y_2 \mid \theta, X) & \cdots & \mathrm{cov}(y_1, y_n \mid \theta, X) \\ \mathrm{cov}(y_2, y_1 \mid \theta, X) & \mathrm{var}(y_2 \mid \theta, X) & & \vdots \\ \vdots & & & \vdots \\ \mathrm{cov}(y_n, y_1 \mid \theta, X) & \cdots & \cdots & \mathrm{var}(y_n \mid \theta, X) \end{pmatrix}$$

For ordinary linear regression, this is just

$$\mathrm{var}(y \mid \theta, X) \;=\; \sigma^2 I$$

where I is the $n \times n$ identity matrix.

Multivariate Normal

The **multivariate normal** distribution

$$\mathrm{N}(\mu, \Sigma)$$

generalizes the (univariate) normal to vectors: If

$$z \sim \mathrm{N}(\mu, \Sigma)$$

then z has mean vector μ, covariance matrix Σ, and normal elements.

See BDA3, Sec. A.1, for density and properties.

Normal-Theory Model

In ordinary *normal-theory* linear regression, we assume

$$y_i \mid \theta, X \;\sim\; \text{indep. } \mathrm{N}(X_i\beta,\, \sigma^2) \qquad\qquad i = 1, \ldots, n$$

which is equivalent to

$$y \mid \theta, X \;\sim\; \mathrm{N}(X\beta,\, \sigma^2 I)$$

For this model,

$$\theta \;=\; (\beta, \sigma^2)$$

In the ordinary normal-theory model, the errors

$$\varepsilon_i \quad = \quad y_i - X_i\beta$$

are conditionally $\mathrm{iid}\ \mathrm{N}(0, \sigma^2)$ (given θ, X).

Checking fit might involve checking errors for

- Non-zero mean structure (depending on unused variables)
- Non-normality, especially outliers
- Non-constant variance structure (depending on X)
- Non-zero covariances (e.g., over time)

Prediction

Consider *unobserved* $\tilde{y}$ from the same kind of regression (incl. error properties).

For the ordinary normal-theory model,

$$\tilde{y} \mid \theta, \tilde{X} \;\sim\; \mathrm{N}(\tilde{X}\beta, \, \sigma^2 I)$$

and conditionally independent of y.

Note: $\tilde{X}$ *is* observed.

Usage: Predict $\tilde{y}$ using known $\tilde{X}$ and fitted model.