Lecture 5: Multivariate Normal Distribution.

Math 586

February 5, 2009

BLUEs/ BLUPs

These are best unbiased linear estimators/ predictors $\hat{Y}$ of unknown quantity Y. *Unbiased* means that $\mathbb{E}\,(\hat{Y}\,|\,\texttt{data}) = \mathbb{E}\,(Y\,|\,\texttt{data})$.

Example:
BLUP of Y_1 based on Y_2, Y_3.
Let $\mathbb{E}\,(Y_j) = 0$, $j = 1, 2, 3$. Note: $\hat{Y}_1 := a_2 Y_2 + a_3 Y_3$ is unbiased (why?)
Minimize mean square error:

$$MSE = \mathbb{E}\,(\hat{Y}_1 - Y_1)^2 = Var(\hat{Y}_1 - Y_1) = Var(a_2 Y_2 + a_3 Y_3 - Y_1) =$$

$$= \sigma_1^2 + a_2^2\sigma_2^2 + a_3^2\sigma_3^2 + 2a_2 a_3\, Cov(Y_2, Y_3) - 2a_2\, Cov(Y_1, Y_2) - 2a_3\, Cov(Y_1, Y_3)$$

Differentiate with respect to a_2:

$$2a_2\sigma_2^2 + 2a_3 Cov(Y_2, Y_3) - 2\, Cov(Y_1, Y_2) = 0$$

Differentiate with respect to a_3:

$$2a_3\sigma_3^2 + 2a_2 Cov(Y_2, Y_3) - 2\, Cov(Y_1, Y_3) = 0$$

Get 2 equations with 2 unknowns (kriging equations): more to come!

Question: when is the best linear predictor i.e. $\hat{Y}_1 = a_2 Y_2 + a_3 Y_3 + ... + a_n Y_n$ also the **best** predictor i.e. $\mathbb{E}\,(Y_1\,|\,Y_2, Y_3, ..., Y_n)$?

Important case: $\mathbf{Y}$ is Multivariate Normal.

Multivariate Normal (MVN).

Univariate:

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}}\ \exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\} = \frac{1}{\sigma\sqrt{2\pi}}\ \exp\left\{-\frac{1}{2}(y-\mu)\sigma^{-2}(y-\mu)\right\}$$

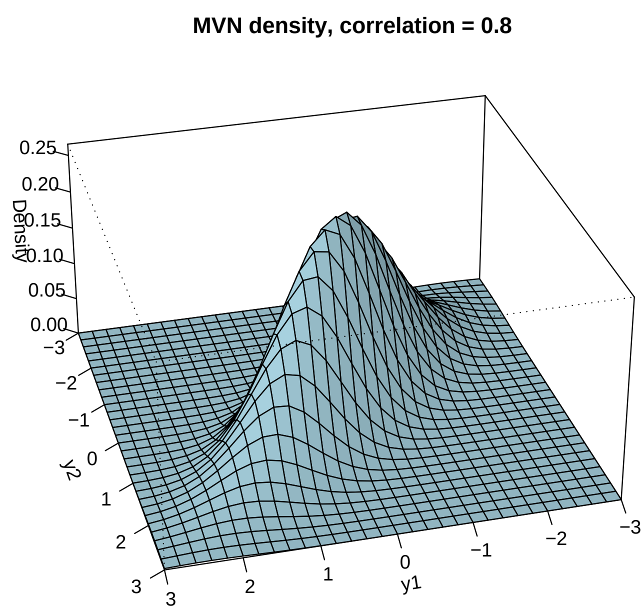
Generalize: $\mathbf{Y} \in \mathbb{R}^n$, $\mathbb{E}\,(\mathbf{Y}) = \boldsymbol{\mu} = (\mu_1, ..., \mu_n)'$.

Let $Var(\mathbf{Y}) = \mathbf{\Sigma} = (\sigma_{ij})$ be $n \times n$ matrix called *variance* or *variance-covariance* matrix of vector $\mathbf{Y}$, so that $\sigma_{ij} = Cov(Y_i, Y_j)$. Let also $det(\mathbf{\Sigma}) = |\mathbf{\Sigma}|$ be the determinant.
Then

$$f(\mathbf{y}) = \frac{1}{\sqrt{|\mathbf{\Sigma}|}(2\pi)^{n/2}} \exp\left\{ -\frac{1}{2}(\mathbf{y} - \mu)' \, \mathbf{\Sigma}^{-1}(\mathbf{y} - \mu) \right\}$$

where $'$ is the transposition and $\mathbf{\Sigma}^{-1}$ = inverse matrix.

Example: when $Y_1, Y_2, ..., Y_n$ are independent and $Var(Y_j) = \sigma_j^2$ then $\mathbf{\Sigma} = diag\{\sigma_1^2, ..., \sigma_n^2\}$. In this case, can prove that the MVN density is the product of marginals.
Vice versa, if $\mathbf{Y}$ is MVN and its variance matrix is diagonal, then all $Y_1, Y_2, ..., Y_n$ are mutually independent.



Alternative Definition: $\mathbf{Y}$ is MVN iff $\sum_{i=1}^n b_i Y_i$ is a univariate Normal for every set of $\{b_i\}_{i=1}^n$ (not all 0's). Often useful.

Properties of MVN

a. Each component Y_i is univariate Normal with mean μ_i and variance σ_{ii}.

b. Any subset of vector $\mathbf{Y}$ is also MVN, with variance matrix being a sub-matrix of $\mathbf{\Sigma}$.

c. If $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2)$ is MVN then conditional $f(\mathbf{Y_1} | \mathbf{Y_2} = \mathbf{y}_2)$ is also MVN.

d. $\mathbb{E}(\mathbf{Y_1} \,|\, \mathbf{Y_2})$ is linear in $\mathbf{Y}_2$.

e. Any linear combination of *independent* MVN is also MVN, i.e. for n-vector $\boldsymbol{\mu}_0$ and constants $\beta_1, ..., \beta_m$ (not all 0),

$$\mathbf{Y} := \boldsymbol{\mu}_0 + \sum_{i=1}^{m} \beta_i \mathbf{Y}_i$$

is n-dim. MVN.

Lemma. For any random vector $\mathbf{Y}$ its variance matrix $\boldsymbol{\Sigma}$ is symmetric and positive semidefinite, i.e. for each n-vector $\mathbf{b}$, $\mathbf{b}'\,\boldsymbol{\Sigma}\,\mathbf{b} \geq 0$.
Proof: Let r.v. $X = \mathbf{b}'\mathbf{Y} = \sum_i b_i Y_i$. Then

$$0 \leq Var(X) = \sum_i b_i^2 Var(Y_i) + \sum_i \sum_{j \neq i} b_i b_j Cov(Y_i, Y_j) = \mathbf{b}'\,\boldsymbol{\Sigma}\,\mathbf{b}.$$

Examples:
1) Let $\mathbf{Y} = (Y_1, Y_2, Y_3, Y_4)'$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} \end{bmatrix}$$

then $(Y_1, Y_3)'$ is MVN with variance matrix

$$\boldsymbol{\Sigma}_{13} = \begin{bmatrix} \sigma_{11} & \sigma_{13} \\ \sigma_{31} & \sigma_{33} \end{bmatrix}$$

and $Y_2 - Y_3$ is univariate Normal with variance ... (see Lemma above, use $\mathbf{b} = (0, 1, -1, 0)'$).

2) Let $\mathbf{Y} = (Y_1, Y_2)$ with variance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

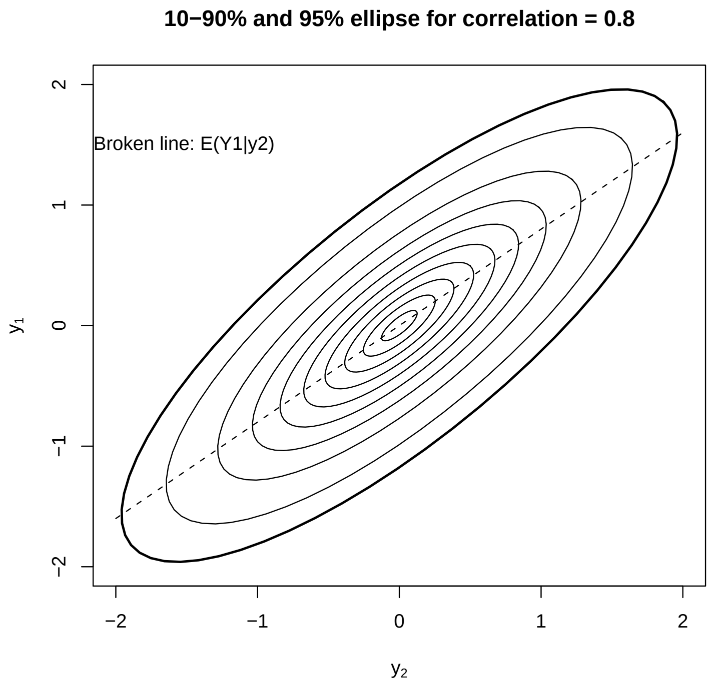
Then

$$\mathbb{E}(Y_1 \,|\, Y_2 = y_2) = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y_2 - \mu_2), \tag{1}$$

$$Var(Y_1 \,|\, Y_2 = y_2) = \sigma_1^2(1 - \rho^2) \tag{2}$$

Surprisingly, the conditional variance does not depend on y_2!

Exercise. Prove (1), (2) by using the ratio formula for conditional density.



General case $\mathbf{Y} = (\mathbf{Y}_1', \mathbf{Y}_2')'$ with mean vector $\boldsymbol{\mu} = (\boldsymbol{\mu}_1', \boldsymbol{\mu}_2')'$ and variance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

Then $f(\mathbf{Y_1} \,|\, \mathbf{Y_2} = \mathbf{y}_2)$ is MVN with mean

$$\mathbb{E}\left(\mathbf{Y}_1 \,|\, \mathbf{Y}_2 = \mathbf{y}_2\right) = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{y}_2 - \boldsymbol{\mu}_2)$$

and conditional var/covar

$$\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$$

does **not** depend on $\mathbf{y}_2$.

Note: if $\boldsymbol{\mu} = 0$ then $\mathbb{E}\left(\mathbf{Y}_1 \,|\, \mathbf{Y}_2 = \mathbf{y}_2\right) = \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{y}_2$. If we seek $\mathbf{B}$ such that

$$\mathbb{E}\left(\mathbf{Y}_1 \,|\, \mathbf{Y}_2 = \mathbf{y}_2\right) = \mathbf{B}\mathbf{y}_2$$

then $\mathbf{B} = \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}$. Thus, the bext unbiased estimator of $\mathbf{Y}_1$ given $\mathbf{y}_2$ is a *linear* function of $\mathbf{y}_2$.

Note: careful. Sometimes the distribution can degenerate, for example normal distribution on a line $y_2 = a + b\,y_1$ is not MVN. Make sure that the variance matrix $\boldsymbol{\Sigma}$ is positive-definite, in particular $|\boldsymbol{\Sigma}| \neq 0$.