Multivariate Normal Theory

STA721 Linear Models Duke University

Merlise Clyde

September 5, 2017
Outline

- Multivariate Normal Distribution
- Linear Transformations
- Distribution of estimates under normality
Properties of MLE’s Recap

- \( \hat{Y} = \hat{\mu} = P_X Y \) is an unbiased estimate of \( \mu = X \beta \)
- \( \mathbb{E}[e] = 0 \) if \( \mu \in C(X) \)

\[
\mathbb{E}[e] = \mathbb{E}[(I - P_X)Y]
\]

- MLE of \( \sigma^2 \):

\[
\hat{\sigma}^2 = \frac{e^T e}{n} = \frac{Y^T(I - P_X)Y}{n}
\]

Is not an unbiased estimate of \( \sigma^2 \), but

\[
\hat{\sigma}^2 \equiv \frac{e^T e}{n - p} = \frac{Y^T(I - P_X)Y}{n - p}
\]

where \( p \) equals the rank of \( X \) is an unbiased estimate.
Sampling Distributions

- Distribution of $\hat{\beta}$
- Distribution of $\mathbf{P}_x \mathbf{Y}$
- Distribution of $\mathbf{e}$
Univariate Normal

Definition
We say that $Z$ has a standard Normal distribution

$$Z \sim N(0, 1)$$

with mean 0 and variance 1 if it has density

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

If $Y = \mu + \sigma Z$ then $Y \sim N(\mu, \sigma^2)$ with mean $\mu$ and variance $\sigma^2$

$$f_Y(y) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{1}{2} \left( \frac{y-\mu}{\sigma} \right)^2}$$
Standard Multivariate Normal

Let $z_i \overset{iid}{\sim} \mathcal{N}(0, 1)$ for $i = 1, \ldots, d$ and define

$$ Z \equiv \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_d \end{bmatrix} $$

▶ Density of $Z$:

$$ f_Z(z) = \prod_{j=1}^{d} \frac{1}{\sqrt{2\pi}} e^{-z_i^2/2} = (2\pi)^{-d/2} e^{-\frac{1}{2}(Z^T Z)} $$

▶ $E[Z] = 0$ and $\text{Cov}[Z] = I_d$

▶ $Z \sim \mathcal{N}(0_d, I_d)$
Multivariate Normal

For a \(d\) dimensional multivariate normal random vector, we write
\[ Y \sim N_d(\mu, \Sigma) \]

- \(E[Y] = \mu\): \(d\) dimensional vector with means \(E[Y_j]\)
- \(\text{Cov}[Y] = \Sigma\): \(d \times d\) matrix with diagonal elements that are the variances of \(Y_j\) and off diagonal elements that are the covariances \(E[(Y_j - \mu_j)(Y_k - \mu_k)]\)

Density

If \(\Sigma\) is positive definite (\(x'\Sigma x > 0\) for any \(x \neq 0\) in \(\mathbb{R}^d\)) then \(Y\) has a density \(^1\)

\[ p(Y) = (2\pi)^{-d/2} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2}(Y - \mu)^T \Sigma^{-1} (Y - \mu)\right) \]

\(^1\)with respect to Lebesgue measure on \(\mathbb{R}^d\)
Theorem

If \( A \ (n \times n) \) is a symmetric real matrix then there exists a \( U \ (n \times n) \) such that \( U^T U = UU^T = I_n \) and a diagonal matrix \( \Lambda \) with elements \( \lambda_i \) such that \( A = U \Lambda U^T \)

- \( U \) is an orthogonal matrix; \( U^{-1} = U^T \)
- The columns of \( U \) from an Orthonormal Basis for \( \mathbb{R}^n \)
- rank of \( A \) equals the number of non-zero eigenvalues \( \lambda_i \)
- Columns of \( U \) associated with non-zero eigenvalues form an ONB for \( C(A) \) (eigenvectors of \( A \))
- \( A^p = U \Lambda^p U^T \) (matrix powers)
- a (symmetric) square root of \( A > 0 \) is \( U \Lambda^{1/2} U^T \)
Multivariate Normal Density

- Density of \( Z \sim N(0, I_d) \):

\[
f_Z(z) = \prod_{j=1}^{d} \frac{1}{\sqrt{2\pi}} e^{-z_i^2/2} = (2\pi)^{-d/2} e^{-\frac{1}{2}(Z^T Z)}
\]

- Write \( Y = \mu + AZ \)
- Solve for \( Z = g(Y) \)
- Jacobian of the transformation \( J(Z \rightarrow Y) = |\frac{\partial g}{\partial Y}| \)
- Substitute \( g(Y) \) for \( Z \) in density and multiply by Jacobian

\[
f_Y(y) = f_Z(z) J(Z \rightarrow Y)
\]
Multivariate Normal Density

\[ Y = \mu + AZ \quad \text{for } Z \sim \mathcal{N}(0, I_d) \quad (1) \]

Proof.

- since \( \Sigma > 0 \), \( \exists \) by the spectral theorem an \( A \) \((d \times d)\) such that \( A > 0 \) and \( AA^T = \Sigma \)
- \( A > 0 \) \( \Rightarrow \) \( A^{-1} \) exists
- Multiply both sides (1) by \( A^{-1} \):
  \[ A^{-1}Y = A^{-1}\mu + A^{-1}AZ \]
- Rearrange \( A^{-1}(Y - \mu) = Z \)
- Jacobian of transformation \( dZ = |A^{-1}|dY \)
- Substitute and simplify algebra
  \[ f(Y) = (2\pi)^{-d/2}|\Sigma|^{-1/2}\exp\left(-\frac{1}{2}(Y - \mu)^T\Sigma^{-1}(Y - \mu)\right) \]
Singular Case

\[ Y = \mu + AZ \text{ with } Z \in \mathbb{R}^d \text{ and } A \text{ is } n \times d \]

1. \[ E[Y] = \mu \]
2. \[ \text{Cov}(Y) = AA^T \geq 0 \]
3. \[ Y \sim N(\mu, \Sigma) \text{ where } \Sigma = AA^T \]

If \( \Sigma \) is singular then there is no density (on \( \mathbb{R}^n \)), but claim that \( Y \) still has a multivariate normal distribution!

**Definition**

\( Y \in \mathbb{R}^n \) has a multivariate normal distribution \( N(\mu, \Sigma) \) if for any \( v \in \mathbb{R}^n \) \( v^T Y \) has a normal distribution with mean \( v^T \mu \) and variance \( v^T \Sigma v \)

see linked videos using characteristic functions:

\[ Y \sim N(\mu, \sigma^2) \iff \varphi_y(t) \equiv E[e^{itY}] = e^{it\mu - t^2\sigma^2/2} \]
Linear Transformations are Normal

If $\mathbf{Y} \sim N_n(\mu, \Sigma)$ then for $\mathbf{A} \ m \times n$

$$\mathbf{A}\mathbf{Y} \sim N_m(\mathbf{A}\mu, \mathbf{A}\Sigma\mathbf{A}^T)$$

$\mathbf{A}\Sigma\mathbf{A}^T$ does not have to be positive definite! (Proof in book or linked video)
Distribution of $\hat{Y}$ and $e$ (marginally)
Multiple ways to define the same normal:

- $Z_1 \sim N(0, I_n)$, $Z_1 \in \mathbb{R}^n$ and take $A$ $d \times n$
- $Z_2 \sim N(0, I_p)$, $Z_2 \in \mathbb{R}^p$ and take $B$ $d \times p$
- Define $Y = \mu + AZ_1$
- Define $W = \mu + BZ_2$

**Theorem**

If $Y = \mu + AZ_1$ and $W = \mu + BZ_2$ then $Y \overset{D}{=} W$ if and only if $AA^T = BB^T = \Sigma$

see linked video
Zero Correlation and Independence

Theorem

For a random vector \( \mathbf{Y} \sim N(\mu, \Sigma) \) partitioned as

\[
\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} \sim N\left( \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right)
\]

then \( \text{Cov}(\mathbf{Y}_1, \mathbf{Y}_2) = \Sigma_{12} = \Sigma_{21} = 0 \) if and only if \( \mathbf{Y}_1 \) and \( \mathbf{Y}_2 \) are independent.
Independence Implies Zero Covariance

Proof.

\[ \text{Cov}(Y_1, Y_2) = \mathbb{E}[(Y_1 - \mu_1)(Y_2 - \mu_2)^T] \]

If \( Y_1 \) and \( Y_2 \) are independent

\[ \mathbb{E}[(Y_1 - \mu_1)(Y_2 - \mu_2)^T] = \mathbb{E}[(Y_1 - \mu_1)\mathbb{E}(Y_2 - \mu_2)^T] = 00^T = 0 \]

therefore \( \Sigma_{12} = 0 \)
Zero Covariance Implies Independence

Assume $\Sigma_{12} = 0$

Proof

- Choose an $A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$ such that $A_1A_1^T = \Sigma_{11}$, $A_2A_2^T = \Sigma_{22}$

- Partition $Z = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \sim N \left( \begin{bmatrix} 0_1 \\ 0_2 \end{bmatrix}, \begin{bmatrix} I_1 & 0 \\ 0 & I_2 \end{bmatrix} \right)$ and $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$

- Then $Y \overset{D}{=} AZ + \mu \sim N(\mu, \Sigma)$
Continued

Proof.

\[
\begin{bmatrix}
Y_1 \\
Y_2
\end{bmatrix} \overset{D}{=} \begin{bmatrix}
A_1 Z_1 + \mu_1 \\
A_2 Z_2 + \mu_2
\end{bmatrix}
\]

- But \( Z_1 \) and \( Z_2 \) are independent
- Functions of \( Z_1 \) and \( Z_2 \) are independent
- Therefore \( Y_1 \) and \( Y_2 \) are independent

For Multivariate Normal Zero Covariance implies independence
Corollary

If \( Y \sim N(\mu, \sigma^2 I_n) \) and \( AB^T = 0 \) then \( AY \) and \( BY \) are independent.

Proof.

\[
\begin{bmatrix}
W_1 \\
W_2
\end{bmatrix} = \begin{bmatrix}
A \\
B
\end{bmatrix} Y = \begin{bmatrix}
AY \\
BY
\end{bmatrix}
\]

\[
\text{Cov}(W_1, W_2) = \text{Cov}(AY, BY) = \sigma^2 AB^T
\]

\( AY \) and \( BY \) are independent if \( AB^T = 0 \)
Joint Distribution of $\hat{Y}$ and $e$
More Distribution Theory

Distributions unconditional on $\sigma^2$

- $\chi^2$ distributions ($\hat{\sigma}^2$)
- $t$ distribution ($\hat{Y}$, $e$, $\hat{\beta}$)