

Multivariate Normal Theory

STA721 Linear Models Duke University

Merlise Clyde

September 2, 2019

Outline

- ▶ Expectations under Moments
- ▶ Spectral Theorem (Singular Value Decomposition)
- ▶ Multivariate Normal Distribution

Properties of OLS/MLEs

$\hat{\mathbf{Y}} = \hat{\boldsymbol{\mu}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ is an unbiased estimate of $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$

$$\begin{aligned} E[\hat{\mathbf{Y}}] &= E[\mathbf{P}_X \mathbf{Y}] \\ &= \mathbf{P}_X E[\mathbf{Y}] \\ &= \mathbf{P}_X \boldsymbol{\mu} \\ &= \boldsymbol{\mu} \end{aligned}$$

$E[\mathbf{e}] = \mathbf{0}$ if $\boldsymbol{\mu} \in C(\mathbf{X})$ (Ex. 1.11 in Christensen)

Will not be $\mathbf{0}$ if $\boldsymbol{\mu} \notin C(\mathbf{X})$ (useful for model checking)

Estimate of σ^2

MLE of σ^2 :

$$\hat{\sigma}^2 = \frac{\mathbf{e}^T \mathbf{e}}{n} = \frac{\mathbf{Y}^T (\mathbf{I} - \mathbf{P}_X) \mathbf{Y}}{n}$$

- ▶ What is the expectation of $\mathbf{e}^T \mathbf{e}$?
- ▶ Is this an unbiased estimate of σ^2 ?

Need expectations of quadratic forms $\mathbf{Y}^T \mathbf{A} \mathbf{Y}$ for \mathbf{A} an $n \times n$ matrix
 \mathbf{Y} a random vector in \mathbb{R}^n

Quadratic Forms

Without loss of generality we can assume that $\mathbf{A} = \mathbf{A}^T$

- ▶ $\mathbf{Y}^T \mathbf{A} \mathbf{Y}$ is a scalar
- ▶ $\mathbf{Y}^T \mathbf{A} \mathbf{Y} = (\mathbf{Y}^T \mathbf{A} \mathbf{Y})^T = \mathbf{Y}^T \mathbf{A}^T \mathbf{Y}$

$$\frac{\mathbf{Y}^T \mathbf{A} \mathbf{Y} + \mathbf{Y}^T \mathbf{A}^T \mathbf{Y}}{2} = \mathbf{Y}^T \mathbf{A} \mathbf{Y}$$
$$\mathbf{Y}^T \frac{(\mathbf{A} + \mathbf{A}^T)}{2} \mathbf{Y} = \mathbf{Y}^T \mathbf{A} \mathbf{Y}$$

- ▶ may take $\mathbf{A} = \mathbf{A}^T$

Expectations of Quadratic Forms

Theorem

Let \mathbf{Y} be a random vector in \mathbb{R}^n with $E[\mathbf{Y}] = \boldsymbol{\mu}$ and $\text{Cov}(\mathbf{Y}) = \boldsymbol{\Sigma}$.
Then $E[\mathbf{Y}^T \mathbf{A} \mathbf{Y}] = \text{tr} \mathbf{A} \boldsymbol{\Sigma} + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu}$.

Result useful for finding expected values of Mean Squares; no normality required! (See Christensen Thm 1.3.2)

Proof

Start with $(\mathbf{Y} - \boldsymbol{\mu})^T \mathbf{A} (\mathbf{Y} - \boldsymbol{\mu})$, expand and take expectations

$$\begin{aligned} E[(\mathbf{Y} - \boldsymbol{\mu})^T \mathbf{A} (\mathbf{Y} - \boldsymbol{\mu})] &= E[\mathbf{Y}^T \mathbf{A} \mathbf{Y} + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu} - \boldsymbol{\mu}^T \mathbf{A} \mathbf{Y} - \mathbf{Y}^T \mathbf{A} \boldsymbol{\mu}] \\ &= E[\mathbf{Y}^T \mathbf{A} \mathbf{Y}] + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu} - \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu} - \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu} \\ &= E[\mathbf{Y}^T \mathbf{A} \mathbf{Y}] - \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu} \end{aligned}$$

Rearrange

$$\begin{aligned} E[\mathbf{Y}^T \mathbf{A} \mathbf{Y}] &= E[(\mathbf{Y} - \boldsymbol{\mu})^T \mathbf{A} (\mathbf{Y} - \boldsymbol{\mu})] + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu} \\ &= E[\text{tr}(\mathbf{Y} - \boldsymbol{\mu})^T \mathbf{A} (\mathbf{Y} - \boldsymbol{\mu})] + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu} \\ &= E[\text{tr} \mathbf{A} (\mathbf{Y} - \boldsymbol{\mu})(\mathbf{Y} - \boldsymbol{\mu})^T] + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu} \\ &= \text{tr} E[\mathbf{A} (\mathbf{Y} - \boldsymbol{\mu})(\mathbf{Y} - \boldsymbol{\mu})^T] + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu} \\ &= \text{tr} \mathbf{A} E[(\mathbf{Y} - \boldsymbol{\mu})(\mathbf{Y} - \boldsymbol{\mu})^T] + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu} \\ &= \text{tr} \mathbf{A} \boldsymbol{\Sigma} + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu} \end{aligned}$$

$$\text{tr} \mathbf{A} \equiv \sum_{i=1}^n a_{ii}$$

Expectation of $\hat{\sigma}^2$

Use the theorem:

$$\begin{aligned}E[\mathbf{Y}^T(\mathbf{I} - \mathbf{P}_X)\mathbf{Y}] &= \text{tr}(\mathbf{I} - \mathbf{P}_X)\sigma^2\mathbf{I} + \boldsymbol{\mu}^T(\mathbf{I} - \mathbf{P}_X)\boldsymbol{\mu} \\&= \sigma^2\text{tr}(\mathbf{I} - \mathbf{P}_X) \\&= \sigma^2r(\mathbf{I} - \mathbf{P}_X) \\&= \sigma^2(n - r(\mathbf{X}))\end{aligned}$$

Therefore an unbiased estimate of σ^2 is

$$\frac{\mathbf{e}^T\mathbf{e}}{n - r(\mathbf{X})}$$

If \mathbf{X} is full rank ($r(\mathbf{X}) = p$) and $\mathbf{P}_X = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ then the

$$\begin{aligned}\text{tr}(\mathbf{P}_X) &= \text{tr}(\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T) \\&= \text{tr}(\mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}) \\&= \text{tr}(\mathbf{I}_p) = p\end{aligned}$$

Spectral Theorem

Theorem

If \mathbf{A} ($n \times n$) is a symmetric real matrix then there exists a \mathbf{U} ($n \times n$) such that $\mathbf{U}^T \mathbf{U} = \mathbf{U} \mathbf{U}^T = \mathbf{I}_n$ and a diagonal matrix $\mathbf{\Lambda}$ with elements λ_i such that $\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$

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

Projections

Projection Matrix

If \mathbf{P} is an orthogonal projection matrix, then its eigenvalues λ_i are either zero or one with $\text{tr}(\mathbf{P}) = \sum_i (\lambda_i) = r(\mathbf{P})$

- ▶ $\mathbf{P} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$
- ▶ $\mathbf{P} = \mathbf{P}^2 \Rightarrow \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T\mathbf{U}\mathbf{\Lambda}\mathbf{U}^T = \mathbf{U}\mathbf{\Lambda}^2\mathbf{U}^T$
- ▶ $\mathbf{\Lambda} = \mathbf{\Lambda}^2$ is true only for $\lambda_i = 1$ or $\lambda_i = 0$
- ▶ Since $r(\mathbf{P})$ is the number of non-zero eigenvalues,
 $r(\mathbf{P}) = \sum \lambda_i = \text{tr}(\mathbf{P})$

$$\mathbf{P} = [\mathbf{U}_P \mathbf{U}_{P^\perp}] \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{n-r} \end{bmatrix} \begin{bmatrix} \mathbf{U}_P^T \\ \mathbf{U}_{P^\perp}^T \end{bmatrix} = \mathbf{U}_P \mathbf{U}_P^T$$

$$\mathbf{P} = \sum_{i=1}^r \mathbf{u}_i \mathbf{u}_i^T$$

sum of r rank 1 projections.

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 μ and variance σ^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 $z_i \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$ for $i = 1, \dots, d$ and define

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

► Density of \mathbf{Z} :

$$\begin{aligned} f_{\mathbf{Z}}(\mathbf{z}) &= \prod_{j=1}^d \frac{1}{\sqrt{2\pi}} e^{-z_j^2/2} \\ &= (2\pi)^{-d/2} e^{-\frac{1}{2}(\mathbf{Z}^T \mathbf{Z})} \end{aligned}$$

► $\mathbb{E}[\mathbf{Z}] = \mathbf{0}$ and $\text{Cov}[\mathbf{Z}] = \mathbf{I}_d$

► $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}_d, \mathbf{I}_d)$

Multivariate Normal

For a d dimensional multivariate normal random vector, we write

$$\mathbf{Y} \sim N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

- ▶ $E[\mathbf{Y}] = \boldsymbol{\mu}$: d dimensional vector with means $E[Y_j]$
- ▶ $\text{Cov}[\mathbf{Y}] = \boldsymbol{\Sigma}$: $d \times d$ 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 $\boldsymbol{\Sigma}$ is positive definite ($\mathbf{x}'\boldsymbol{\Sigma}\mathbf{x} > 0$ for any $\mathbf{x} \neq 0$ in \mathbb{R}^d) then \mathbf{Y} has a density ¹

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

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

Multivariate Normal Density

- Density of $Z \sim N(\mathbf{0}, \mathbf{I}_d)$:

$$\begin{aligned}f_{\mathbf{Z}}(\mathbf{z}) &= \prod_{j=1}^d \frac{1}{\sqrt{2\pi}} e^{-z_j^2/2} \\&= (2\pi)^{-d/2} e^{-\frac{1}{2}(\mathbf{Z}^T \mathbf{Z})}\end{aligned}$$

- Write $\mathbf{Y} = \boldsymbol{\mu} + \mathbf{A}\mathbf{Z}$
- Solve for $\mathbf{Z} = g(\mathbf{Y})$
- Jacobian of the transformation $J(\mathbf{Z} \rightarrow \mathbf{Y}) = \left| \frac{\partial g}{\partial \mathbf{Y}} \right|$
- substitute $g(\mathbf{Y})$ for \mathbf{Z} in density and multiply by Jacobian

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{Z}}(\mathbf{z}) J(\mathbf{Z} \rightarrow \mathbf{Y})$$

Multivariate Normal Density

$$\mathbf{Y} = \boldsymbol{\mu} + \mathbf{A}\mathbf{Z} \quad \text{for } \mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d) \quad (1)$$

Proof.

- ▶ since $\boldsymbol{\Sigma} > 0$, \exists by the spectral theorem an \mathbf{A} ($d \times d$) such that $\mathbf{A} > 0$ and $\mathbf{A}\mathbf{A}^T = \boldsymbol{\Sigma}$ a (symmetric) square root of $\mathbf{A} > 0$ is $\mathbf{U}\boldsymbol{\Lambda}^{1/2}\mathbf{U}^T$
- ▶ $\mathbf{A} > 0 \Rightarrow \mathbf{A}^{-1}$ exists
- ▶ Multiply both sides (1) by \mathbf{A}^{-1} :

$$\mathbf{A}^{-1}\mathbf{Y} = \mathbf{A}^{-1}\boldsymbol{\mu} + \mathbf{A}^{-1}\mathbf{A}\mathbf{Z}$$

- ▶ Rearrange $\mathbf{A}^{-1}(\mathbf{Y} - \boldsymbol{\mu}) = \mathbf{Z}$
- ▶ Jacobian of transformation $d\mathbf{Z} = |\mathbf{A}^{-1}|d\mathbf{Y}$
- ▶ Substitute and simplify algebra

$$f(\mathbf{Y}) = (2\pi)^{-d/2} |\boldsymbol{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{Y} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \boldsymbol{\mu})\right)$$

Singular Case

$\mathbf{Y} = \boldsymbol{\mu} + \mathbf{AZ}$ with $\mathbf{Z} \in \mathbb{R}^d$ and \mathbf{A} is $n \times d$

- ▶ $E[\mathbf{Y}] = \boldsymbol{\mu}$
- ▶ $\text{Cov}(\mathbf{Y}) = \mathbf{AA}^T \geq 0$
- ▶ $\mathbf{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma} = \mathbf{AA}^T$

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

Definition

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

see linked videos using characteristic functions:

$$Y \sim N(\mu, \sigma^2) \Leftrightarrow \varphi_Y(t) \equiv E[e^{itY}] = e^{it\mu - t^2\sigma^2/2}$$

Linear Transformations are Normal

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

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

$\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T$ does not have to be positive definite!
(Proof in book or linked video)

Use to prove that all marginal distributions of normals are normal!

Conditional Normals

Partition \mathbf{Y}

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix}$$

as well as the mean and covariance matrix. Then the joint distribution is:

$$\begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}\right)$$

The conditional distribution of \mathbf{Y}_1 given $\mathbf{Y}_2 = \mathbf{y}_2$

$$\mathbf{Y}_1 \mid \mathbf{Y}_2 \sim N(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{y}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$$

Brute Force or Linear Transformations!

Equal in Distribution

Multiple ways to define the same normal:

- ▶ $\mathbf{Z}_1 \sim N(\mathbf{0}, \mathbf{I}_n)$, $\mathbf{Z}_1 \in \mathbb{R}^n$ and take $\mathbf{A} \ d \times n$
- ▶ $\mathbf{Z}_2 \sim N(\mathbf{0}, \mathbf{I}_p)$, $\mathbf{Z}_2 \in \mathbb{R}^p$ and take $\mathbf{B} \ d \times p$
- ▶ Define $\mathbf{Y} = \boldsymbol{\mu} + \mathbf{AZ}_1$
- ▶ Define $\mathbf{W} = \boldsymbol{\mu} + \mathbf{BZ}_2$

Theorem

If $\mathbf{Y} = \boldsymbol{\mu} + \mathbf{AZ}_1$ and $\mathbf{W} = \boldsymbol{\mu} + \mathbf{BZ}_2$ then $\mathbf{Y} \stackrel{D}{=} \mathbf{W}$ if and only if $\mathbf{AA}^T = \mathbf{BB}^T = \boldsymbol{\Sigma}$

see linked video

Zero Correlation and Independence

Theorem

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

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}\right)$$

then $\text{Cov}(\mathbf{Y}_1, \mathbf{Y}_2) = \boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21}^T = \mathbf{0}$ if and only if \mathbf{Y}_1 and \mathbf{Y}_2 are independent.

Independence Implies Zero Covariance

Proof.

$$\text{Cov}(\mathbf{Y}_1, \mathbf{Y}_2) = E[(\mathbf{Y}_1 - \boldsymbol{\mu}_1)(\mathbf{Y}_2 - \boldsymbol{\mu}_2)^T]$$

If \mathbf{Y}_1 and \mathbf{Y}_2 are independent

$$E[(\mathbf{Y}_1 - \boldsymbol{\mu}_1)(\mathbf{Y}_2 - \boldsymbol{\mu}_2)^T] = E[(\mathbf{Y}_1 - \boldsymbol{\mu}_1)E(\mathbf{Y}_2 - \boldsymbol{\mu}_2)^T] = \mathbf{0}\mathbf{0}^T = \mathbf{0}$$

therefore $\boldsymbol{\Sigma}_{12} = \mathbf{0}$



Zero Covariance Implies Independence

Assume $\Sigma_{12} = \mathbf{0}$

Proof

► Choose an

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix}$$

such that $\mathbf{A}_1 \mathbf{A}_1^T = \Sigma_{11}$, $\mathbf{A}_2 \mathbf{A}_2^T = \Sigma_{22}$

► Partition

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{0}_1 \\ \mathbf{0}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix} \right) \text{ and } \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

► then $\mathbf{Y} \stackrel{D}{=} \mathbf{AZ} + \boldsymbol{\mu} \sim N(\boldsymbol{\mu}, \Sigma)$

Continued

Proof.



$$\begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} \stackrel{D}{=} \begin{bmatrix} \mathbf{A}_1 \mathbf{Z}_1 + \boldsymbol{\mu}_1 \\ \mathbf{A}_2 \mathbf{Z}_2 + \boldsymbol{\mu}_2 \end{bmatrix}$$

- ▶ But \mathbf{Z}_1 and \mathbf{Z}_2 are independent
- ▶ Functions of \mathbf{Z}_1 and \mathbf{Z}_2 are independent
- ▶ Therefore \mathbf{Y}_1 and \mathbf{Y}_2 are independent



For Multivariate Normal Zero Covariance implies independence

Corollary

Corollary

If $\mathbf{Y} \sim N(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n)$ and $\mathbf{AB}^T = \mathbf{0}$ then \mathbf{AY} and \mathbf{BY} are independent.

Proof.

$$\begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} \mathbf{Y} = \begin{bmatrix} \mathbf{AY} \\ \mathbf{BY} \end{bmatrix}$$

- ▶ $\text{Cov}(\mathbf{W}_1, \mathbf{W}_2) = \text{Cov}(\mathbf{AY}, \mathbf{BY}) = \sigma^2 \mathbf{AB}^T$
- ▶ \mathbf{AY} and \mathbf{BY} are independent if $\mathbf{AB}^T = \mathbf{0}$

