Canonical Correlation Analysis

Bidit Sadhukhan, Anirban Dey

Ramkrishna Mission Vivekananda Educational Research Institute

Motivation

Suppose there is a firm that surveyed a random sample of $n = 50$ of its employees in an attempt to determine which factors influence sales performance. Two collections of variables were measured:

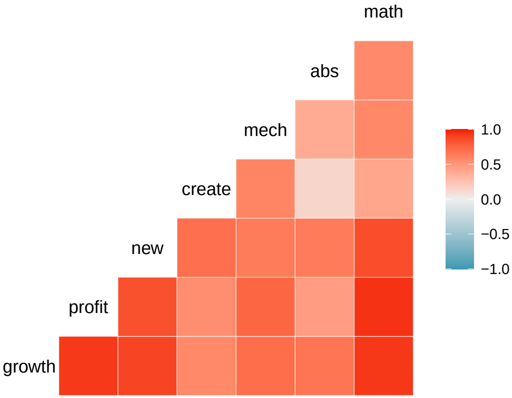
- Sales Performance:
  - Sales Growth
  - Sales Profitability
  - New Account Sales
- Test Scores as a Measure of Intelligence
  - Creativity
  - Mechanical Reasoning
  - Abstract Reasoning
  - Mathematics

Correlation Matrix

There are $p = 3$ variables in the first group relating to Sales Performance and $q = 4$ variables in the second group relating to Test Scores.

	growth	profit	new	create	mech	abs	math
growth	1.00	0.93	0.88	0.57	0.71	0.67	0.93
profit	0.93	1.00	0.84	0.54	0.75	0.47	0.94
new	0.88	0.84	1.00	0.70	0.64	0.64	0.85
create	0.57	0.54	0.70	1.00	0.59	0.15	0.41
mech	0.71	0.75	0.64	0.59	1.00	0.39	0.57
abs	0.67	0.47	0.64	0.15	0.39	1.00	0.57
math	0.93	0.94	0.85	0.41	0.57	0.57	1.00

Correlation Plots for Sales Data



Motivation for large p & q

- What if p and q are large? There will be pq such scatter plots and a correlation matrix of pq dimension.

Motivation for large p & q

- What if p and q are large? There will be pq such scatter plots and a correlation matrix of pq dimension.
- How can we interpret for large p,q?

Purpose

Canonical Correlation Analysis (CCA) connects two sets of variables by finding linear combinations of variables that maximally correlate.

- Data reduction: explain covariation between two sets of variables using small number of linear combinations

Purpose

Canonical Correlation Analysis (CCA) connects two sets of variables by finding linear combinations of variables that maximally correlate.

- Data reduction: explain covariation between two sets of variables using small number of linear combinations
- Data interpretation: find features (i.e., canonical variates) that are important for explaining covariation between sets of variables

Canonical Variates- Notations

Notations:

$$\mathbf{X} = [X_1, X_2, \dots, X_p]^T$$

$$\mathbf{Y} = [Y_1, Y_2, \dots, Y_q]^T$$

and $p \leq q$.

Define:

$$U = a^T X$$

&

$$V = b^T Y$$

Canonical Variates

U and V are called Canonical Variates

- (U_1, V_1) is called as the first canonical covariate pair

Canonical Variates

U and V are called Canonical Variates

- (U_1, V_1) is called as the first canonical covariate pair
- (U_i, V_i) is the i^{th} canonical covariate pair.

Canonical Variates

U and V are called Canonical Variates

- (U_1, V_1) is called as the first canonical covariate pair
- (U_i, V_i) is the i^{th} canonical covariate pair.
- $p \leq q$, there are p canonical covariate pair.

Canonical Variates- Properties

- $E(X) = \mu_x$, $Cov(X) = \Sigma_x$ then

$$E(U) = E(a^T X) = a^T \mu_x$$

$$Cov(U) = Cov(a^T X) = a^T \Sigma_x a$$

Canonical Variates- Properties

- $E(X) = \mu_x$, $Cov(X) = \Sigma_x$ then

$$E(U) = E(a^T X) = a^T \mu_x$$

$$Cov(U) = Cov(a^T X) = a^T \Sigma_x a$$

- $E(Y) = \mu_y$, $Cov(Y) = \Sigma_y$ then

$$E(V) = E(b^T Y) = b^T \mu_y$$

$$Cov(V) = Cov(b^T X) = b^T \Sigma_y b$$

Canonical Variates- Properties

- $E(X) = \mu_x$, $Cov(X) = \Sigma_x$ then

$$E(U) = E(a^T X) = a^T \mu_x$$

$$Cov(U) = Cov(a^T X) = a^T \Sigma_x a$$

- $E(Y) = \mu_y$, $Cov(Y) = \Sigma_y$ then

$$E(V) = E(b^T Y) = b^T \mu_y$$

$$Cov(V) = Cov(b^T X) = b^T \Sigma_y b$$

- $Cov(X, Y) = \Sigma_{xy}$ then $Cov(U, V) = a^T \Sigma_{xy} b$

Relation between Canonical Covariates & Original

- $$Cor(X,Y) = \rho_{xy} = \frac{\mathsf{Cov}(XY)}{\sqrt{\mathsf{var}(X)} \cdot \sqrt{\mathsf{Var}(Y)}}$$

Relation between Canonical Covariates & Original

- $$Cor(X,Y) = \rho_{xy} = \frac{\mathsf{Cov}(XY)}{\sqrt{\mathsf{var}(X)} \cdot \sqrt{\mathsf{Var}(Y)}}$$

- $$Cor(U,V) = \rho_{uv} = \frac{\mathsf{Cov}(UV)}{\sqrt{\mathsf{Var}(U)} \cdot \sqrt{\mathsf{Var}(V)}} = \frac{a^T \Sigma_{xy} b}{\sqrt{a^T \Sigma_x a} \cdot \sqrt{b^T \Sigma_y b}}$$

Canonical Variates Defined

First Canonical Covariate:

- Define

$$U_1 = a_1^T X$$

$$V_1 = b_1^T Y$$

Canonical Variates Defined

First Canonical Covariate:

- Define

$$U_1 = a_1^T X$$

$$V_1 = b_1^T Y$$

- Conditions

Canonical Variates Defined

First Canonical Covariate:

- Define

$$U_1 = a_1^T X$$

$$V_1 = b_1^T Y$$

- Conditions

1.

$$\mathsf{Var}(U_1) = \mathsf{Var}(V_1) = 1$$

Canonical Variates Defined

First Canonical Covariate:

- Define

$$U_1 = a_1^T X$$
$$V_1 = b_1^T Y$$

- Conditions

1. $$\mathsf{Var}(U_1) = \mathsf{Var}(V_1) = 1$$

2. $$\rho_1^*(U_1, V_1) = \max_{a,b} \rho(a^T X, b^T Y)$$

(U_1, V_1) is the first pair of canonical variable.

Second Canonical Covariate

- Define

$$U_2 = a_2^T X$$
$$V_2 = b_2^T Y$$

- Conditions

1.
$$\mathsf{Var}(U_2) = \mathsf{Var}(V_2) = 1$$

2.
$$\rho_2^*(U_2, V_2) = \max_{a,b} \rho(a^T X, b^T Y)$$

3.
$$Cov(U_1, U_2) = Cov(V_1, V_2) = 0$$

4.
$$Cov(U_1, V_2) = Cov(U_2, V_1) = 0$$

i-th canonical covariate

- Define

$$U_i = a_i^T X$$
$$V_i = b_i^T Y$$

- Conditions

1. $$\text{Var}(U_i) = \text{Var}(V_i) = 1$$

2. $$\rho_i^*(U_i, V_i) = \max_{a,b} \rho(a^T X, b^T Y)$$

3. $$Cov(U_i, U_k) = Cov(V_i, V_k) = 0$$

4. $$Cov(U_i, V_k) = Cov(U_k, V_i) = 0$$

The i-th pair of canonical variates is given by

$$U_i = e_i^T \Sigma_X^{-1/2} X \quad \text{(represented as } a_i^T X)$$

$$V_i = f_i^T \Sigma_Y^{-1/2} Y \quad \text{(represented as } b_i^T Y)$$

where:

- e_i is the i_{th} eigenvector of $\Sigma_X^{-1/2}\Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX}\Sigma_X^{-1/2}$

The i_{th} canonical correlation is given by:

$$Cor(U_i, V_i) = \rho_i^*$$

where ρ_i^{*2} is the i_{th} eigenvalue of $\Sigma_X^{-1/2}\Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX}\Sigma_X^{-1/2}$ Note: ρ_i^{*2} is also the i_{th} eigenvalue of $\Sigma_Y^{-1/2}\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY}\Sigma_Y^{-1/2}$

The i-th pair of canonical variates is given by

$$U_i = e_i^T \Sigma_X^{-1/2} X \quad \text{(represented as } a_i^T X\text{)}$$
$$V_i = f_i^T \Sigma_Y^{-1/2} Y \quad \text{(represented as } b_i^T Y\text{)}$$

where:

- e_i is the i_{th} eigenvector of $\Sigma_X^{-1/2}\Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX}\Sigma_X^{-1/2}$
- f_i is the i_{th} eigenvector of $\Sigma_Y^{-1/2}\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY}\Sigma_Y^{-1/2}$

The i_{th} canonical correlation is given by:

$$Cor(U_i, V_i) = \rho_i^*$$

where ρ_i^{*2} is the i_{th} eigenvalue of $\Sigma_X^{-1/2}\Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX}\Sigma_X^{-1/2}$ Note: ρ_i^{*2} is also the i_{th} eigenvalue of $\Sigma_Y^{-1/2}\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY}\Sigma_Y^{-1/2}$

Properties of Canonical Variates

Define: $Corr(U_i, V_i) = \rho_i^*, \quad \ni i = 1...p, \quad \& \; p \leq q$

- $\rho_i^{*2} \leq \rho_{i-1}^{*2} \leq \rho_{i-2}^{*2} ... \rho_3^{*2} \leq \rho_2^{*2} \leq \rho_1^{*2}$

Properties of Canonical Variates

Define: $Corr(U_i, V_i) = \rho_i^*, \quad \ni i = 1...p, \quad \& \ p \leq q$

- $\rho_i^{*2} \leq \rho_{i-1}^{*2} \leq \rho_{i-2}^{*2} ... \rho_3^{*2} \leq \rho_2^{*2} \leq \rho_1^{*2}$
- $\rho_i^{*2}, \rho_{i-1}^{*2}, .., \rho_2^{*2}, \rho_1^{*2}$ are the non-zero eigen values of $\Sigma_{11}^{-0.5}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-0.5}$ and$\Sigma_{22}^{-0.5}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-0.5}$

Properties of Canonical Variates

Define: $Corr(U_i, V_i) = \rho_i^*, \quad \ni i = 1...p, \quad \& \; p \leq q$

- $\rho_i^{*2} \leq \rho_{i-1}^{*2} \leq \rho_{i-2}^{*2} ... \rho_3^{*2} \leq \rho_2^{*2} \leq \rho_1^{*2}$
- $\rho_i^{*2}, \rho_{i-1}^{*2}, .., \rho_2^{*2}, \rho_1^{*2}$ are the non-zero eigen values of $\Sigma_{11}^{-0.5}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-0.5}$ and$\Sigma_{22}^{-0.5}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-0.5}$
- Canonical correlations are unchanged by standardization.

Correlation of Original and Canonical Variables

The canonical variates and original variables have correlation matrices:

- $Cor(U, X) = Cov(AX, \widetilde{\Sigma}_X^{-1/2} X) = A\Sigma_X \widetilde{\Sigma}_X^{-1/2}$

given that $Var(U_k) = Var(V_k) = 1 \forall k, l$ where - $\widetilde{\Sigma}_X = diag(\sqrt{X})$ is a diagonal matrix containing X variances. - $\widetilde{\Sigma}_Y = diag(\sqrt{Y})$ is a diagonal matrix containing Y variances.

Correlation of Original and Canonical Variables

The canonical variates and original variables have correlation matrices:

- $Cor(U, X) = Cov(AX, \widetilde{\Sigma}_X^{-1/2} X) = A\Sigma_X \widetilde{\Sigma}_X^{-1/2}$
- $Cor(U, Y) = Cov(AX; \widetilde{\Sigma}_Y^{-1/2} Y) = A\Sigma_{XY} \widetilde{\Sigma}_Y^{-1/2}$

given that $Var(U_k) = Var(V_k) = 1 \forall k, l$ where - $\widetilde{\Sigma}_X = diag(\sqrt{X})$ is a diagonal matrix containing X variances. - $\widetilde{\Sigma}_Y = diag(\sqrt{Y})$ is a diagonal matrix containing Y variances.

Correlation of Original and Canonical Variables

The canonical variates and original variables have correlation matrices:

- $Cor(U, X) = Cov(AX, \widetilde{\Sigma}_X^{-1/2} X) = A\Sigma_X \widetilde{\Sigma}_X^{-1/2}$
- $Cor(U, Y) = Cov(AX; \widetilde{\Sigma}_Y^{-1/2} Y) = A\Sigma_{XY} \widetilde{\Sigma}_Y^{-1/2}$
- $Cor(V, X) = Cov(BY; \widetilde{\Sigma}_X^{-1/2} X) = B\Sigma_{YX} \widetilde{\Sigma}_X^{-1/2}$

given that $Var(U_k) = Var(V_k) = 1 \forall k, l$ where - $\widetilde{\Sigma}_X = diag(\sqrt{X})$ is a diagonal matrix containing X variances. - $\widetilde{\Sigma}_Y = diag(\sqrt{Y})$ is a diagonal matrix containing Y variances.

Correlation of Original and Canonical Variables

The canonical variates and original variables have correlation matrices:

- $Cor(U, X) = Cov(AX, \widetilde{\Sigma}_X^{-1/2} X) = A\Sigma_X \widetilde{\Sigma}_X^{-1/2}$
- $Cor(U, Y) = Cov(AX; \widetilde{\Sigma}_Y^{-1/2} Y) = A\Sigma_{XY} \widetilde{\Sigma}_Y^{-1/2}$
- $Cor(V, X) = Cov(BY; \widetilde{\Sigma}_X^{-1/2} X) = B\Sigma_{YX} \widetilde{\Sigma}_X^{-1/2}$
- $Cor(V, Y) = Cov(BY; \widetilde{\Sigma}_Y^{-1/2} Y) = B\Sigma_Y \widetilde{\Sigma}_Y^{-1/2}$

given that $Var(U_k) = Var(V_k) = 1 \forall k, l$ where -
$\widetilde{\Sigma}_X = diag(\sqrt{X})$ is a diagonal matrix containing X variances. -
$\widetilde{\Sigma}_Y = diag(\sqrt{Y})$ is a diagonal matrix containing Y variances.

Sample Canonical Correlation

Assume that $\mathbf{z}_i = (\mathbf{x}_i', \mathbf{y}_i')'$

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_X \\ \boldsymbol{\mu}_Y \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_X & \boldsymbol{\Sigma}_{XY} \\ \boldsymbol{\Sigma}_{YX} & \boldsymbol{\Sigma}_Y \end{pmatrix}$$

and let the sample mean vector and covariance matrix be denoted by

$$\bar{\mathbf{z}} = \begin{pmatrix} \bar{\mathbf{x}} \\ \bar{\mathbf{y}} \end{pmatrix} \quad \text{and} \quad \mathbf{S} = \begin{pmatrix} \mathbf{S}_X & \mathbf{S}_{XY} \\ \mathbf{S}_{YX} & \mathbf{S}_Y \end{pmatrix}$$

where

- $\bar{\mathbf{x}} = (1/n)\sum_{i=1}^n \mathbf{x}_i$ and $\bar{\mathbf{y}} = (1/n)\sum_{i=1}^n \mathbf{y}_i$
- $\mathbf{S}_X = \frac{1}{n-1}\sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'$
- $\mathbf{S}_Y = \frac{1}{n-1}\sum_{i=1}^n (\mathbf{y}_i - \bar{\mathbf{y}})(\mathbf{y}_i - \bar{\mathbf{y}})'$
- $\mathbf{S}_{XY} = \mathbf{S}_{YX}' = \frac{1}{n-1}\sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{y}_i - \bar{\mathbf{y}})'$

Sample Properties

Note that $U = \mathbf{a}'\boldsymbol{X}$ and $V = \mathbf{b}'\boldsymbol{Y}$ have sample properties

$$\widehat{\mathrm{Var}}(U) = \mathbf{a}'\mathbf{S}_X\mathbf{a}$$
$$\widehat{\mathrm{Var}}(V) = \mathbf{b}'\mathbf{S}_Y\mathbf{b}$$
$$\widehat{\mathrm{Cov}}(U, V) = \mathbf{a}'\mathbf{S}_{XY}\mathbf{b}$$

The first pair of sample canonical variates (U_1, V_1) is defined via the pair of linear combination vectors $\{\mathbf{a}_1, \mathbf{b}_1\}$ that maximize

$$\widehat{\mathrm{Cor}}(U, V) = \frac{\widehat{\mathrm{Cov}}(U, V)}{\sqrt{\widehat{\mathrm{Var}}(U)}\sqrt{\widehat{\mathrm{Var}}(V)}} = \frac{\mathbf{a}'\mathbf{S}_{XY}\mathbf{b}}{\sqrt{\mathbf{a}'\mathbf{S}_X\mathbf{a}}\sqrt{\mathbf{b}'\mathbf{S}_Y\mathbf{b}}}$$

subject to U_1 and V_1 having unit variance.

Remaining canonical variates (U_ℓ, V_ℓ) maximize the above subject to having unit variance and being uncorrelated with (U_k, V_k) for all $k < \ell$.

Sample Canonical Correlation

The sample estimate of the k-th pair of canonical variates is given by

$$\hat{U}_k = \underbrace{\hat{\mathbf{u}}_k' \mathbf{S}_X^{-1/2}}_{\hat{\mathbf{a}}_k'} \boldsymbol{X} \quad \text{and} \quad \hat{V}_k = \underbrace{\hat{\mathbf{v}}_k \mathbf{S}_Y^{-1/2}}_{\hat{\mathbf{b}}_k'} \boldsymbol{Y}$$

where

- $\hat{\mathbf{u}}_k$ is the k-th eigenvector of $\mathbf{S}_X^{-1/2} \mathbf{S}_{XY} \mathbf{S}_Y^{-1} \mathbf{S}_{YX} \mathbf{S}_X^{-1/2}$
- $\hat{\mathbf{v}}_k$ is the k-th eigenvector of $\mathbf{S}_Y^{-1/2} \mathbf{S}_{YX} \mathbf{S}_X^{-1} \mathbf{S}_{XY} \mathbf{S}_Y^{-1/2}$

The sample estimate of the k-th canonical correlation is given by

$$\widehat{\mathrm{Cor}}(U_k, V_k) = \hat{\rho}_k$$

where $\hat{\rho}_k^2$ is the k-th eigenvalue of $\mathbf{S}_X^{-1/2} \mathbf{S}_{XY} \mathbf{S}_Y^{-1} \mathbf{S}_{YX} \mathbf{S}_X^{-1/2}$
[$\hat{\rho}_k^2$ is also the k-th eigenvalue of $\mathbf{S}_Y^{-1/2} \mathbf{S}_{YX} \mathbf{S}_X^{-1} \mathbf{S}_{XY} \mathbf{S}_Y^{-1/2}$]

Continued

$\hat{\boldsymbol{U}} = \hat{\mathbf{A}}'\boldsymbol{X}$ and $\hat{\boldsymbol{V}} = \hat{\mathbf{B}}'\boldsymbol{Y}$ where $\hat{\mathbf{A}} = [\hat{\mathbf{a}}_1, \ldots, \hat{\mathbf{a}}_p]$ and $\hat{\mathbf{B}} = [\hat{\mathbf{b}}_1, \ldots, \hat{\mathbf{b}}_q]$.

- $\hat{\boldsymbol{U}} = (\hat{U}_1, \ldots, \hat{U}_p)'$ contains the p canonical variates from $\boldsymbol{X}$
- $\hat{\boldsymbol{V}} = (\hat{V}_1, \ldots, \hat{V}_q)'$ contains the q canonical variates from $\boldsymbol{Y}$
- If $p \leq q$, we are interested in first p canonical variates from $\boldsymbol{Y}$

The sample canonical variates and original variables have covariances

$$\begin{aligned}
\widehat{\mathrm{Cov}}(\hat{\boldsymbol{U}}, \boldsymbol{X}) &= \widehat{\mathrm{Cov}}(\hat{\mathbf{A}}'\boldsymbol{X}, \boldsymbol{X}) = \hat{\mathbf{A}}'\mathbf{S}_X \\
\widehat{\mathrm{Cov}}(\hat{\boldsymbol{U}}, \boldsymbol{Y}) &= \widehat{\mathrm{Cov}}(\hat{\mathbf{A}}'\boldsymbol{X}, \boldsymbol{Y}) = \hat{\mathbf{A}}'\mathbf{S}_{XY} \\
\widehat{\mathrm{Cov}}(\hat{\boldsymbol{V}}, \boldsymbol{X}) &= \widehat{\mathrm{Cov}}(\hat{\mathbf{B}}'\boldsymbol{Y}, \boldsymbol{X}) = \hat{\mathbf{B}}'\mathbf{S}_{YX} \\
\widehat{\mathrm{Cov}}(\hat{\boldsymbol{V}}, \boldsymbol{Y}) &= \widehat{\mathrm{Cov}}(\hat{\mathbf{B}}'\boldsymbol{Y}, \boldsymbol{Y}) = \hat{\mathbf{B}}'\mathbf{S}_Y
\end{aligned}$$

Continued

The sample canonical variates and original variables have correlations

$$\widehat{\mathrm{Cor}}(\hat{\boldsymbol{U}}, \boldsymbol{X}) = \widehat{\mathrm{Cov}}(\hat{\mathbf{A}}'\boldsymbol{X}, \tilde{\mathbf{S}}_X^{-1/2}\boldsymbol{X}) = \hat{\mathbf{A}}'\mathbf{S}_X\tilde{\mathbf{S}}_X^{-1/2}$$
$$\widehat{\mathrm{Cor}}(\hat{\boldsymbol{U}}, \boldsymbol{Y}) = \widehat{\mathrm{Cov}}(\hat{\mathbf{A}}'\boldsymbol{X}, \tilde{\mathbf{S}}_Y^{-1/2}\boldsymbol{Y}) = \hat{\mathbf{A}}'\mathbf{S}_{XY}\tilde{\mathbf{S}}_Y^{-1/2}$$
$$\widehat{\mathrm{Cor}}(\hat{\boldsymbol{V}}, \boldsymbol{X}) = \widehat{\mathrm{Cov}}(\hat{\mathbf{B}}'\boldsymbol{Y}, \tilde{\mathbf{S}}_X^{-1/2}\boldsymbol{X}) = \hat{\mathbf{B}}'\mathbf{S}_{YX}\tilde{\mathbf{S}}_X^{-1/2}$$
$$\widehat{\mathrm{Cor}}(\hat{\boldsymbol{V}}, \boldsymbol{Y}) = \widehat{\mathrm{Cov}}(\hat{\mathbf{B}}'\boldsymbol{Y}, \tilde{\mathbf{S}}_Y^{-1/2}\boldsymbol{Y}) = \hat{\mathbf{B}}'\mathbf{S}_Y\tilde{\mathbf{S}}_Y^{-1/2}$$

given that $\mathrm{Var}(\hat{U}_k) = \mathrm{Var}(\hat{V}_\ell) = 1$ for all k, ℓ.

- $\tilde{\mathbf{S}}_X = \mathrm{diag}(\mathbf{S}_X)$ is a diagonal matrix containing $\boldsymbol{X}$ variances
- $\tilde{\mathbf{S}}_Y = \mathrm{diag}(\mathbf{S}_Y)$ is a diagonal matrix containing $\boldsymbol{Y}$ variances

Gemometrical Interpretation of CCA

- Let's look at a geometric interpretation of CCA.
- First, some notation:
  - Let A be the matrix whose k-th row is the k-th canonical direction $e_k^T \Sigma_Y^{-1/2}$.
  - Let E be the matrix whose k-th *column* is the eigenvector e_k. Note that $E^T E = I_p$.
  - We thus have $A = E^T \Sigma_Y^{-1/2}$.
- We get all canonical variates U_k by transforming $\mathbf{Y}$ using A:

$$\mathbf{U} = A\mathbf{Y}.$$

Continued

- Now, using the spectral decomposition of Σ_Y, we can write
$$A = E^T \Sigma_Y^{-1/2} = E^T P_Y \Lambda_Y^{-1/2} P_Y^T,$$
where P_Y contains the eigenvectors of Σ_Y and Λ_Y is the diagonal matrix with its eigenvalues.
- Therefore, we can see that
$$\mathbf{U} = A\mathbf{Y} = E^T P_Y \Lambda_Y^{-1/2} P_Y^T \mathbf{Y}.$$

Continued

- Let's look at this expression in stages:
  - $P_Y^T \mathbf{Y}$: This is the matrix of **principal components** of $\mathbf{Y}$.
  - $\Lambda_Y^{-1/2}\left(P_Y^T \mathbf{Y}\right)$: We standardize the principal components to have unit variance.
  - $P_Y\left(\Lambda_Y^{-1/2} P_Y^T \mathbf{Y}\right)$: We rotate the standardized PCs using a transformation that **only involves** Σ_Y.
  - $E^T\left(P_Y \Lambda_Y^{-1/2} P_Y^T \mathbf{Y}\right)$: We rotate the result using a transformation that **involves the whole covariance matrix** Σ.

Canonical Correlations for Sales Data

```
Canonical Correlations:
[1] 0.9944827 0.8781065 0.3836057
```

Canonical Coefficients for Sales Data

```
Canonical Coefficients for Sales Variables:

        [,1]   [,2]   [,3]
growth 0.009  0.025 -0.054
profit 0.003 -0.035  0.015
new    0.011  0.034  0.055

Canonical Coefficients for Test Scores:

        [,1]   [,2]   [,3]
create 0.010  0.027  0.035
mech   0.004 -0.029 -0.020
abs    0.013  0.071 -0.040
math   0.009 -0.010  0.002

     growth      profit         new
0.008911125 0.002989377 0.011179739
```

$Thank$

$You!$