Regression-based methods for finding coupled patterns

Michael K. Tippett[1] Timothy DelSole[2]
Simon J. Mason[1] Anthony G. Barnston[1]

[1]International Research Institute for Climate and Society
The Earth Institute, Columbia University

[2]George Mason University
Center for Ocean-Land-Atmosphere Studies (COLA)

88th AMS Annual Meeting

Main idea

Connect

Coupled pattern methods and multivariate regression

Results

- Pattern methods diagonalize the regression.
  - Diagonal regression easy to understand.
  - Different pattern methods diagnose different properties.
- Pattern methods are SVDs of the regression.
  - Different methods = different norms.
  - Changing norm = linear transformation of data.

Known for CCA.
New for MCA, RDA, PPA.

Main idea

Connect

Coupled pattern methods and multivariate regression

Results

- Pattern methods diagonalize the regression.
  - Diagonal regression easy to understand.
  - Different pattern methods diagnose different properties.
- Pattern methods are SVDs of the regression.
  - Different methods = different norms.
  - Changing norm = linear transformation of data.

Known for CCA.
New for MCA, RDA, PPA.

Main idea

Connect

Coupled pattern methods and multivariate regression

Results

- Pattern methods diagonalize the regression.
  - Diagonal regression easy to understand.
  - Different pattern methods diagnose different properties.
- Pattern methods are SVDs of the regression.
  - Different methods = different norms.
  - Changing norm = linear transformation of data.

Known for CCA.
New for MCA, RDA, PPA.

Main idea

Connect

Coupled pattern methods and multivariate regression

Results

- Pattern methods diagonalize the regression.
  - Diagonal regression easy to understand.
  - Different pattern methods diagnose different properties.
- Pattern methods are SVDs of the regression.
  - Different methods = different norms.
  - Changing norm = linear transformation of data.

Known for CCA.
New for MCA, RDA, PPA.

Main idea

Connect

Coupled pattern methods and multivariate regression

Results

- Pattern methods diagonalize the regression.
  - Diagonal regression easy to understand.
  - Different pattern methods diagnose different properties.
- Pattern methods are SVDs of the regression.
  - Different methods = different norms.
  - Changing norm = linear transformation of data.

Known for CCA.
New for MCA, RDA, PPA.

Main idea

Connect

Coupled pattern methods and multivariate regression

Results

- Pattern methods diagonalize the regression.
  - Diagonal regression easy to understand.
  - Different pattern methods diagnose different properties.
- Pattern methods are SVDs of the regression.
  - Different methods = different norms.
  - Changing norm = linear transformation of data.

Known for CCA.

New for MCA, RDA, PPA.

Main idea

Connect

Coupled pattern methods and multivariate regression

Results

- Pattern methods diagonalize the regression.
  - Diagonal regression easy to understand.
  - Different pattern methods diagnose different properties.
- Pattern methods are SVDs of the regression.
  - Different methods = different norms.
  - Changing norm = linear transformation of data.

Known for CCA.
New for MCA, RDA, PPA.

Main idea

Connect

Coupled pattern methods and multivariate regression

Results

- Pattern methods diagonalize the regression.
  - Diagonal regression easy to understand.
  - Different pattern methods diagnose different properties.
- Pattern methods are SVDs of the regression.
  - Different methods = different norms.
  - Changing norm = linear transformation of data.

Known for CCA.

New for MCA, RDA, PPA.

Coupled pattern methods

Two data sets. (x and y)
Optimally decompose the data.

- CCA – canonical correlation analysis
  Components with maximum *correlation*.
- MCA – maximum covariance analysis.
  Components with maximum *covariance*.

Predictor and predictand. (asymmetric)

- RDA – redundancy analysis.
  Predictor components that explain the most *variance*.
- PPA – principal predictor analysis
  Predictor components that explain the most *standardized variance*.

Coupled pattern methods

Two data sets. (x and y)
Optimally decompose the data.

- CCA – canonical correlation analysis
  Components with maximum *correlation*.
- MCA – maximum covariance analysis.
  Components with maximum *covariance*.

Predictor and predictand. (asymmetric)

- RDA – redundancy analysis.
  Predictor components that explain the most *variance*.
- PPA – principal predictor analysis
  Predictor components that explain the most *standardized variance*.

Coupled pattern methods

Two data sets. (x and y)
Optimally decompose the data.

- CCA – canonical correlation analysis
  Components with maximum *correlation*.
- MCA – maximum covariance analysis.
  Components with maximum *covariance*.

Predictor and predictand. (asymmetric)

- RDA – redundancy analysis.
  Predictor components that explain the most *variance*.
- PPA – principal predictor analysis
  Predictor components that explain the most *standardized variance*.

Coupled pattern methods

Two data sets. (x and y)
Optimally decompose the data.

- CCA – canonical correlation analysis
  Components with maximum *correlation*.
- MCA – maximum covariance analysis.
  Components with maximum *covariance*.

Predictor and predictand. (asymmetric)

- RDA – redundancy analysis.
  Predictor components that explain the most *variance*.
- PPA – principal predictor analysis
  Predictor components that explain the most *standardized variance*.

Coupled pattern methods

Two data sets. (x and y)
Optimally decompose the data.

- CCA – canonical correlation analysis
  Components with maximum *correlation*.
- MCA – maximum covariance analysis.
  Components with maximum *covariance*.

Predictor and predictand. (asymmetric)

- RDA – redundancy analysis.
  Predictor components that explain the most *variance*.
- PPA – principal predictor analysis
  Predictor components that explain the most *standardized variance*.

Coupled pattern methods

Two data sets. (x and y)
Optimally decompose the data.

- CCA – canonical correlation analysis
  Components with maximum *correlation*.
- MCA – maximum covariance analysis.
  Components with maximum *covariance*.

Predictor and predictand. (asymmetric)

- RDA – redundancy analysis.
  Predictor components that explain the most *variance*.
- PPA – principal predictor analysis
  Predictor components that explain the most *standardized variance*.

Multivariate regression

$$y = Ax, \quad A = \langle yx^T \rangle \langle xx^T \rangle^{-1} = \text{regression coefficients}$$

How to diagnose the regression?
Diagonalize A. A is not square. SVD of A:

$$\begin{aligned} A &= USV^T \quad m \times n \\ &= [u_1, \ldots, u_m]\text{diag}(s_1, s_2, \ldots, s_p)[v_1, \ldots, v_n]^T \\ &= \sum_{i=1}^{p} s_i u_i v_i^T \end{aligned}$$

U, V are orthogonal. S is diagonal. $s_1 \geq s_2 \geq \ldots s_p \geq 0$.
Regression is diagonal in the basis of the columns of U and V

$$Av_j = \sum_{i=1}^{p} s_i u_i \boxed{v_i^T v_j} = s_j u_j$$

Multivariate regression

$$y = Ax, \quad A = \langle yx^T \rangle \langle xx^T \rangle^{-1} = \text{regression coefficients}$$

How to diagnose the regression?

Diagonalize A. A is not square. SVD of A:

$$\begin{aligned} A &= USV^T \quad m \times n \\ &= [u_1, \ldots, u_m]\text{diag}(s_1, s_2, \ldots, s_p)[v_1, \ldots, v_n]^T \\ &= \sum_{i=1}^{p} s_i u_i v_i^T \end{aligned}$$

U, V are orthogonal. S is diagonal. $s_1 \geq s_2 \geq \ldots s_p \geq 0$.
Regression is diagonal in the basis of the columns of U and V

$$Av_j = \sum_{i=1}^{p} s_i u_i \boxed{v_i^T v_j} = s_j u_j$$

Multivariate regression

$$y = Ax, \quad A = \langle yx^T \rangle \langle xx^T \rangle^{-1} = \text{regression coefficients}$$

How to diagnose the regression?
Diagonalize A. A is not square. SVD of A:

$$\begin{aligned} A &= USV^T \quad m \times n \\ &= [u_1, \dots, u_m] \mathrm{diag}(s_1, s_2, \dots, s_p) [v_1, \dots, v_n]^T \\ &= \sum_{i=1}^{p} s_i u_i v_i^T \end{aligned}$$

U, V are orthogonal. S is diagonal. $s_1 \geq s_2 \geq \dots s_p \geq 0$.
Regression is diagonal in the basis of the columns of U and V

$$Av_j = \sum_{i=1}^{p} s_i u_i \boxed{v_i^T v_j} = s_j u_j$$

Multivariate regression

$$y = Ax, \quad A = \langle yx^T \rangle \langle xx^T \rangle^{-1} = \text{regression coefficients}$$

How to diagnose the regression?
Diagonalize A. A is not square. SVD of A:

$$\begin{aligned} A &= USV^T \quad m \times n \\ &= [u_1, \ldots, u_m] \text{diag}(s_1, s_2, \ldots, s_p) [v_1, \ldots, v_n]^T \\ &= \sum_{i=1}^{p} s_i u_i v_i^T \end{aligned}$$

U, V are orthogonal. S is diagonal. $s_1 \geq s_2 \geq \ldots s_p \geq 0$.
Regression is diagonal in the basis of the columns of U and V

$$Av_j = \sum_{i=1}^{p} s_i u_i \boxed{v_i^T v_j} = s_j u_j$$

Multivariate regression

$$y = Ax, \quad A = \langle yx^T \rangle \langle xx^T \rangle^{-1} = \text{regression coefficients}$$

How to diagnose the regression?
Diagonalize A. A is not square. SVD of A:

$$\begin{aligned} A &= USV^T \quad m \times n \\ &= [u_1, \ldots, u_m] \text{diag}(s_1, s_2, \ldots, s_p) [v_1, \ldots, v_n]^T \\ &= \sum_{i=1}^{p} s_i u_i v_i^T \end{aligned}$$

U, V are orthogonal. S is diagonal. $s_1 \geq s_2 \geq \ldots s_p \geq 0$.
Regression is diagonal in the basis of the columns of U and V

$$Av_j = \sum_{i=1}^{p} s_i u_i \boxed{v_i^T v_j} = s_j u_j$$

Multivariate regression

$$y = Ax, \quad A = \langle yx^T \rangle \langle xx^T \rangle^{-1} = \text{regression coefficients}$$

How to diagnose the regression?
Diagonalize A. A is not square. SVD of A:

$$\begin{aligned} A &= USV^T \quad m \times n \\ &= [u_1, \ldots, u_m] \text{diag}(s_1, s_2, \ldots, s_p) [v_1, \ldots, v_n]^T \\ &= \sum_{i=1}^{p} s_i u_i v_i^T \end{aligned}$$

U, V are orthogonal. S is diagonal. $s_1 \geq s_2 \geq \ldots s_p \geq 0$.
Regression is diagonal in the basis of the columns of U and V

$$Av_j = \sum_{i=1}^{p} s_i u_i \boxed{v_i^T v_j} = s_j u_j$$

Multivariate regression

$$y = Ax, \quad A = \langle yx^T \rangle \langle xx^T \rangle^{-1} = \text{regression coefficients}$$

How to diagnose the regression?
Diagonalize A. A is not square. SVD of A:

$$\begin{aligned} A &= USV^T \quad m \times n \\ &= [u_1, \ldots, u_m]\text{diag}(s_1, s_2, \ldots, s_p)[v_1, \ldots, v_n]^T \\ &= \sum_{i=1}^{p} s_i u_i v_i^T \end{aligned}$$

U, V are orthogonal. S is diagonal. $s_1 \geq s_2 \geq \ldots s_p \geq 0$.
Regression is diagonal in the basis of the columns of U and V

$$Av_j = \sum_{i=1}^{p} s_i u_i \boxed{v_i^T v_j} = s_j u_j$$

Multivariate regression

$$y = Ax, \quad A = \langle yx^T \rangle \langle xx^T \rangle^{-1} = \text{regression coefficients}$$

How to diagnose the regression?
Diagonalize A. A is not square. SVD of A:

$$\begin{aligned} A &= USV^T \quad m \times n \\ &= [u_1, \ldots, u_m] \text{diag}(s_1, s_2, \ldots, s_p) [v_1, \ldots, v_n]^T \\ &= \sum_{i=1}^{p} s_i u_i v_i^T \end{aligned}$$

U, V are orthogonal. S is diagonal. $s_1 \geq s_2 \geq \ldots s_p \geq 0$.
Regression is diagonal in the basis of the columns of U and V

$$Av_j = \sum_{i=1}^{p} s_i u_i \boxed{v_i^T v_j} = s_j u_j$$

SVD and optimization

SVD solves an optimization problem

$$s_1 = \max_v \frac{\|Av\|}{\|v\|}$$

or

$$s_1 = \max_{u,v} \frac{u^T A v}{\|u\| \|v\|}$$

Usual vector norm $\|x\|^2 = x^T x$.

If x is an initial condition, and y is a final condition, s_1 measures the maximum amplification.
Growth of instabilities, ENSO, ensemble perturbations (Farrell, Penland, Palmer etc.).

Depends on two norms, one for x and one for y.

SVD and optimization

SVD solves an optimization problem

$$s_1 = \max_v \frac{\|Av\|}{\|v\|}$$

or

$$s_1 = \max_{u,v} \frac{u^T A v}{\|u\|\|v\|}$$

Usual vector norm $\|x\|^2 = x^T x$.

If x is an initial condition, and y is a final condition, s_1 measures the maximum amplification.
Growth of instabilities, ENSO, ensemble perturbations (Farrell, Penland, Palmer etc.).

Depends on two norms, one for x and one for y.

SVD and optimization

SVD solves an optimization problem

$$s_1 = \max_v \frac{\|Av\|}{\|v\|}$$

or

$$s_1 = \max_{u,v} \frac{u^T Av}{\|u\|\|v\|}$$

Usual vector norm $\|x\|^2 = x^T x$.

If x is an initial condition, and y is a final condition,
s_1 measures the maximum amplification.
Growth of instabilities, ENSO, ensemble perturbations (Farrell, Penland, Palmer etc.).

Depends on two norms, one for x and one for y.

SVD and optimization

SVD solves an optimization problem

$$s_1 = \max_v \frac{\|Av\|}{\|v\|}$$

or

$$s_1 = \max_{u,v} \frac{u^T A v}{\|u\| \|v\|}$$

Usual vector norm $\|x\|^2 = x^T x$.

If x is an initial condition, and y is a final condition,
s_1 measures the maximum amplification.
Growth of instabilities, ENSO, ensemble perturbations (Farrell, Penland, Palmer etc.).

Depends on two norms, one for x and one for y.

Norms and changing variables

Usual vector norms

$$\|v\|^2 = v^T v\,, \quad \|u\|^2 = u^T u$$

New norms

$$\|v\|_x^2 \equiv v^T L^T L v = \|Lv\|^2$$

$$\|u\|_y^2 \equiv u^T M^T M u = \|Mu\|^2$$

New variables Lv and Mu.

$$\|\text{old variables}\|_{\text{new}} = \|\text{new variables}\|_{\text{old}}$$

New norms are norms if L and M are invertible.

Norms and changing variables

Usual vector norms

$$\|v\|^2 = v^T v\,, \quad \|u\|^2 = u^T u$$

New norms

$$\|v\|_x^2 \equiv v^T L^T L v = \|Lv\|^2$$

$$\|u\|_y^2 \equiv u^T M^T M u = \|Mu\|^2$$

New variables Lv and Mu.

$$\|\text{old variables}\|_{\text{new}} = \|\text{new variables}\|_{\text{old}}$$

New norms are norms if L and M are invertible.

Norms and changing variables

Usual vector norms

$$\|v\|^2 = v^T v\,, \quad \|u\|^2 = u^T u$$

New norms

$$\|v\|_x^2 \equiv v^T L^T L v = \|Lv\|^2$$

$$\|u\|_y^2 \equiv u^T M^T M u = \|Mu\|^2$$

New variables Lv and Mu.

$$\|\text{old variables}\|_{\text{new}} = \|\text{new variables}\|_{\text{old}}$$

New norms are norms if L and M are invertible.

Norms and changing variables

Usual vector norms

$$\|v\|^2 = v^T v \,, \quad \|u\|^2 = u^T u$$

New norms

$$\|v\|_x^2 \equiv v^T L^T L v = \|Lv\|^2$$

$$\|u\|_y^2 \equiv u^T M^T M u = \|Mu\|^2$$

New variables Lv and Mu.

$$\|\text{old variables}\|_{\text{new}} = \|\text{new variables}\|_{\text{old}}$$

New norms are norms if L and M are invertible.

Norms and changing variables

Usual vector norms

$$\|v\|^2 = v^T v\,, \quad \|u\|^2 = u^T u$$

New norms

$$\|v\|_x^2 \equiv v^T L^T L v = \|Lv\|^2$$

$$\|u\|_y^2 \equiv u^T M^T M u = \|Mu\|^2$$

New variables Lv and Mu.

$$\|\text{old variables}\|_{\text{new}} = \|\text{new variables}\|_{\text{old}}$$

New norms are norms if L and M are invertible.

SVD, norms and changing variables

SVD with new norm

$$\max_v \frac{\|Av\|_y}{\|v\|_x} = \max_v \frac{\|MAv\|}{\|Lv\|} = \max_{v'} \frac{\|MAL^{-1}v'\|}{\|v'\|}$$

$$(v' = Lv)$$

$$\mathrm{SVD}_{\text{new norms}}(A) = \mathrm{SVD}_{\text{old norms}}(MAL^{-1})$$

MAL^{-1} is the regression matrix between Lx and My.

$$My = MAx = \left(MAL^{-1}\right) Lx$$

$$\mathrm{SVD}_{\text{new norms}}(x \to y) = \mathrm{SVD}_{\text{old norms}}(Lx \to My)$$

SVD, norms and changing variables

SVD with new norm

$$\max_v \frac{\|Av\|_y}{\|v\|_x} = \max_v \frac{\|MAv\|}{\|Lv\|} = \max_{v'} \frac{\|MAL^{-1}v'\|}{\|v'\|}$$

$$(v' = Lv)$$

$$\text{SVD}_{\text{new norms}}(A) = \text{SVD}_{\text{old norms}}(MAL^{-1})$$

MAL^{-1} is the regression matrix between Lx and My.

$$My = MAx = \left(MAL^{-1}\right) Lx$$

$$\text{SVD}_{\text{new norms}}(x \to y) = \text{SVD}_{\text{old norms}}(Lx \to My)$$

SVD, norms and changing variables

SVD with new norm

$$\max_v \frac{\|Av\|_y}{\|v\|_x} = \max_v \frac{\|MAv\|}{\|Lv\|} = \max_{v'} \frac{\|MAL^{-1}v'\|}{\|v'\|}$$

$$(v' = Lv)$$

$$\text{SVD}_{\text{new norms}}(A) = \text{SVD}_{\text{old norms}}(MAL^{-1})$$

MAL^{-1} is the regression matrix between Lx and My.

$$My = MAx = \left(MAL^{-1}\right) Lx$$

$$\text{SVD}_{\text{new norms}}(x \to y) = \text{SVD}_{\text{old norms}}(Lx \to My)$$

SVD, norms and changing variables

SVD with new norm

$$\max_v \frac{\|Av\|_y}{\|v\|_x} = \max_v \frac{\|MAv\|}{\|Lv\|} = \max_{v'} \frac{\|MAL^{-1}v'\|}{\|v'\|}$$

$$(v' = Lv)$$

$$\text{SVD}_{\text{new norms}}(A) = \text{SVD}_{\text{old norms}}(MAL^{-1})$$

MAL^{-1} is the regression matrix between Lx and My.

$$My = MAx = \left(MAL^{-1}\right) Lx$$

$$\text{SVD}_{\text{new norms}}(x \to y) = \text{SVD}_{\text{old norms}}(Lx \to My)$$

SVD, norms and changing variables

SVD with new norm

$$\max_v \frac{\|Av\|_y}{\|v\|_x} = \max_v \frac{\|MAv\|}{\|Lv\|} = \max_{v'} \frac{\|MAL^{-1}v'\|}{\|v'\|}$$

$$(v' = Lv)$$

$$\text{SVD}_{\text{new norms}}(A) = \text{SVD}_{\text{old norms}}(MAL^{-1})$$

MAL^{-1} is the regression matrix between Lx and My.

$$My = MAx = \left(MAL^{-1}\right) Lx$$

$$\text{SVD}_{\text{new norms}}(x \to y) = \text{SVD}_{\text{old norms}}(Lx \to My)$$

SVD, norms and changing variables

SVD with new norm

$$\max_v \frac{\|Av\|_y}{\|v\|_x} = \max_v \frac{\|MAv\|}{\|Lv\|} = \max_{v'} \frac{\|MAL^{-1}v'\|}{\|v'\|}$$

$$(v' = Lv)$$

$$\text{SVD}_{\text{new norms}}(A) = \text{SVD}_{\text{old norms}}(MAL^{-1})$$

MAL^{-1} is the regression matrix between Lx and My.

$$My = MAx = \left(MAL^{-1}\right) Lx$$

$$\text{SVD}_{\text{new norms}}(x \to y) = \text{SVD}_{\text{old norms}}(Lx \to My)$$

SVD, norms and changing variables

SVD with new norm

$$\max_v \frac{\|Av\|_y}{\|v\|_x} = \max_v \frac{\|MAv\|}{\|Lv\|} = \max_{v'} \frac{\|MAL^{-1}v'\|}{\|v'\|}$$

$$(v' = Lv)$$

$$\text{SVD}_{\text{new norms}}(A) = \text{SVD}_{\text{old norms}}(MAL^{-1})$$

MAL^{-1} is the regression matrix between Lx and My.

$$My = MAx = \left(MAL^{-1}\right) Lx$$

$$\text{SVD}_{\text{new norms}}(x \to y) = \text{SVD}_{\text{old norms}}(Lx \to My)$$

Norms for the regression matrix

What norms (change of variables) make the SVD of A measure:

- Correlation.
- Explained variance.
- Standardized explained variance.
- Covariance.

Norms for the regression matrix

What norms (change of variables) make the SVD of A measure:

- Correlation.
- Explained variance.
- Standardized explained variance.
- Covariance.

Norms for the regression matrix

What norms (change of variables) make the SVD of A measure:

- Correlation.
- Explained variance.
- Standardized explained variance.
- Covariance.

Norms for the regression matrix

What norms (change of variables) make the SVD of A measure:

- Correlation.
- Explained variance.
- Standardized explained variance.
- Covariance.

Norms for the regression matrix

What norms (change of variables) make the SVD of A measure:

- Correlation.
- Explained variance.
- Standardized explained variance.
- Covariance.

Univariate regression

$$y = ax \qquad a = \langle yx \rangle / \langle x^2 \rangle$$

New variables: $x' = lx$ and $y' = my$
New regression: $y' = a'x'$ where $a' = mal^{-1}$

Correlation

- $x' = x/\sqrt{\langle x^2 \rangle}, y' = y/\sqrt{\langle y^2 \rangle}$
  $a' = a\sqrt{\frac{\langle x^2 \rangle}{\langle y^2 \rangle}} = \frac{\langle xy \rangle}{\sqrt{\langle x^2 \rangle \langle y^2 \rangle}}$

Square-root of explained variance

- $x' = x/\sqrt{\langle x^2 \rangle}$
  $a' = a\sqrt{\langle x^2 \rangle} = \frac{\langle xy \rangle}{\sqrt{\langle x^2 \rangle}}$

Covariance

- $x' = x/\langle x^2 \rangle$
  $a' = a\langle x^2 \rangle = \langle yx \rangle$

Univariate regression

$$y = ax \qquad a = \langle yx\rangle / \langle x^2\rangle$$

New variables: $x' = lx$ and $y' = my$

New regression: $y' = a'x'$ where $a' = mal^{-1}$

Correlation

- $x' = x/\sqrt{\langle x^2\rangle}$, $y' = y/\sqrt{\langle y^2\rangle}$
  $a' = a\sqrt{\frac{\langle x^2\rangle}{\langle y^2\rangle}} = \frac{\langle xy\rangle}{\sqrt{\langle x^2\rangle\langle y^2\rangle}}$

Square-root of explained variance

- $x' = x/\sqrt{\langle x^2\rangle}$
  $a' = a\sqrt{\langle x^2\rangle} = \frac{\langle xy\rangle}{\sqrt{\langle x^2\rangle}}$

Covariance

- $x' = x/\langle x^2\rangle$
  $a' = a\langle x^2\rangle = \langle yx\rangle$

Univariate regression

$$y = ax \qquad a = \langle yx \rangle / \langle x^2 \rangle$$

New variables: $x' = lx$ and $y' = my$
New regression: $y' = a'x'$ where $a' = mal^{-1}$

Correlation

- $x' = x/\sqrt{\langle x^2 \rangle}$, $y' = y/\sqrt{\langle y^2 \rangle}$
  $a' = a\sqrt{\frac{\langle x^2 \rangle}{\langle y^2 \rangle}} = \frac{\langle xy \rangle}{\sqrt{\langle x^2 \rangle \langle y^2 \rangle}}$

Square-root of explained variance

- $x' = x/\sqrt{\langle x^2 \rangle}$
  $a' = a\sqrt{\langle x^2 \rangle} = \frac{\langle xy \rangle}{\sqrt{\langle x^2 \rangle}}$

Covariance

- $x' = x/\langle x^2 \rangle$
  $a' = a\langle x^2 \rangle = \langle yx \rangle$

Univariate regression

$$y = ax \qquad a = \langle yx \rangle / \langle x^2 \rangle$$

New variables: $x' = lx$ and $y' = my$
New regression: $y' = a'x'$ where $a' = mal^{-1}$

Correlation

- $x' = x/\sqrt{\langle x^2 \rangle}$, $y' = y/\sqrt{\langle y^2 \rangle}$
  $a' = a\sqrt{\frac{\langle x^2 \rangle}{\langle y^2 \rangle}} = \frac{\langle xy \rangle}{\sqrt{\langle x^2 \rangle \langle y^2 \rangle}}$

Square-root of explained variance

- $x' = x/\sqrt{\langle x^2 \rangle}$
  $a' = a\sqrt{\langle x^2 \rangle} = \frac{\langle xy \rangle}{\sqrt{\langle x^2 \rangle}}$

Covariance

- $x' = x/\langle x^2 \rangle$
  $a' = a\langle x^2 \rangle = \langle yx \rangle$

Univariate regression

$$y = ax \qquad a = \langle yx \rangle / \langle x^2 \rangle$$

New variables: $x' = lx$ and $y' = my$
New regression: $y' = a'x'$ where $a' = mal^{-1}$

Correlation

- $x' = x/\sqrt{\langle x^2 \rangle}$, $y' = y/\sqrt{\langle y^2 \rangle}$
  $a' = a\sqrt{\frac{\langle x^2 \rangle}{\langle y^2 \rangle}} = \frac{\langle xy \rangle}{\sqrt{\langle x^2 \rangle \langle y^2 \rangle}}$

Square-root of explained variance

- $x' = x/\sqrt{\langle x^2 \rangle}$
  $a' = a\sqrt{\langle x^2 \rangle} = \frac{\langle xy \rangle}{\sqrt{\langle x^2 \rangle}}$

Covariance

- $x' = x/\langle x^2 \rangle$
  $a' = a\langle x^2 \rangle = \langle yx \rangle$

Multivariate

Claim: multivariate = replace variances by covariances.

CCA

$x \to \langle xx^T \rangle^{-1/2} x = Lx$, $y \to \langle yy^T \rangle^{-1/2} y = My$

(replace data by PCs, "whitening", unit variance, uncorrelated)

Proof 1: new norm SVD → CCA eigenvalue equations.

$$\max_v \frac{\|Av\|_y^2}{\|v\|_x^2} = \max_v \frac{\|MAv\|^2}{\|Lv\|^2} = \max_v \frac{\|\langle yy^T\rangle^{-1/2} Av\|^2}{\|\langle yy^T\rangle^{-1/2} v\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T\rangle^{-1/2} A \langle xx^T\rangle^{1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T\rangle^{-1/2} \langle yx^T\rangle \langle xx^T\rangle^{-1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{v'^T \langle xx^T\rangle^{-1/2} \langle xy^T\rangle \langle yy^T\rangle^{-1} \langle yx^T\rangle \langle xx^T\rangle^{-1/2} v'}{v'^T v'}$$

Eigenvalues of $\langle xx^T\rangle^{-1} \langle xy^T\rangle \langle yy^T\rangle^{-1} \langle yx^T\rangle$ = CCA

Multivariate

Claim: multivariate = replace variances by covariances.

CCA

$x \rightarrow \langle xx^T \rangle^{-1/2} x = Lx$, $y \rightarrow \langle yy^T \rangle^{-1/2} y = My$

(replace data by PCs, "whitening", unit variance, uncorrelated)

Proof 1: new norm SVD → CCA eigenvalue equations.

$$\max_v \frac{\|Av\|_y^2}{\|v\|_x^2} = \max_v \frac{\|MAv\|^2}{\|Lv\|^2} = \max_v \frac{\|\langle yy^T \rangle^{-1/2} Av\|^2}{\|\langle yy^T \rangle^{-1/2} v\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} A \langle xx^T \rangle^{1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{v'^T \langle xx^T \rangle^{-1/2} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'}{v'^T v'}$$

Eigenvalues of $\langle xx^T \rangle^{-1} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle$ = CCA

Multivariate

Claim: multivariate = replace variances by covariances.

CCA

$$x \to \langle xx^T \rangle^{-1/2} x = Lx,\; y \to \langle yy^T \rangle^{-1/2} y = My$$

(replace data by PCs, “whitening”, unit variance, uncorrelated)

Proof 1: new norm SVD → CCA eigenvalue equations.

$$\begin{aligned}
\max_v \frac{\|Av\|_y^2}{\|v\|_x^2} &= \max_v \frac{\|MAv\|^2}{\|Lv\|^2} = \max_v \frac{\|\langle yy^T \rangle^{-1/2} Av\|^2}{\|\langle yy^T \rangle^{-1/2} v\|^2} \\
&= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} A \langle xx^T \rangle^{1/2} v'\|^2}{\|v'\|^2} \\
&= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'\|^2}{\|v'\|^2} \\
&= \max_{v'} \frac{v'^T \langle xx^T \rangle^{-1/2} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'}{v'^T v'}
\end{aligned}$$

Eigenvalues of $\langle xx^T \rangle^{-1} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle$ = CCA

Multivariate

Claim: multivariate = replace variances by covariances.

CCA

$x \to \langle xx^T \rangle^{-1/2} x = Lx,\ y \to \langle yy^T \rangle^{-1/2} y = My$

(replace data by PCs, “whitening”, unit variance, uncorrelated)

Proof 1: new norm SVD $\to$ CCA eigenvalue equations.

$$\max_v \frac{\|Av\|_y^2}{\|v\|_x^2} = \max_v \frac{\|MAv\|^2}{\|Lv\|^2} = \max_v \frac{\|\langle yy^T\rangle^{-1/2} Av\|^2}{\|\langle yy^T\rangle^{-1/2} v\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T\rangle^{-1/2} A \langle xx^T\rangle^{1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T\rangle^{-1/2} \langle yx^T\rangle \langle xx^T\rangle^{-1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{v'^T \langle xx^T\rangle^{-1/2} \langle xy^T\rangle \langle yy^T\rangle^{-1} \langle yx^T\rangle \langle xx^T\rangle^{-1/2} v'}{v'^T v'}$$

Eigenvalues of $\langle xx^T\rangle^{-1} \langle xy^T\rangle \langle yy^T\rangle^{-1} \langle yx^T\rangle$ = CCA

Multivariate

Claim: multivariate = replace variances by covariances.

CCA

$x \to \langle xx^T \rangle^{-1/2} x = Lx$, $y \to \langle yy^T \rangle^{-1/2} y = My$

(replace data by PCs, "whitening", unit variance, uncorrelated)

Proof 1: new norm SVD → CCA eigenvalue equations.

$$\max_v \frac{\|Av\|_y^2}{\|v\|_x^2} = \max_v \frac{\|MAv\|^2}{\|Lv\|^2} = \max_v \frac{\|\langle yy^T \rangle^{-1/2} Av\|^2}{\|\langle yy^T \rangle^{-1/2} v\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} A \langle xx^T \rangle^{1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{v'^T \langle xx^T \rangle^{-1/2} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'}{v'^T v'}$$

Eigenvalues of $\langle xx^T \rangle^{-1} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle$ = CCA

Multivariate

Claim: multivariate = replace variances by covariances.

CCA

$x \to \langle xx^T \rangle^{-1/2} x = Lx$, $y \to \langle yy^T \rangle^{-1/2} y = My$

(replace data by PCs, “whitening”, unit variance, uncorrelated)
Proof 1: new norm SVD → CCA eigenvalue equations.

$$\max_v \frac{\|Av\|_y^2}{\|v\|_x^2} = \max_v \frac{\|MAv\|^2}{\|Lv\|^2} = \max_v \frac{\|\langle yy^T \rangle^{-1/2} Av\|^2}{\|\langle yy^T \rangle^{-1/2} v\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} A \langle xx^T \rangle^{1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{v'^T \langle xx^T \rangle^{-1/2} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'}{v'^T v'}$$

Eigenvalues of $\langle xx^T \rangle^{-1} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle$ = CCA

Multivariate

Claim: multivariate = replace variances by covariances.

CCA

$x \to \langle xx^T \rangle^{-1/2} x = Lx$, $y \to \langle yy^T \rangle^{-1/2} y = My$

(replace data by PCs, "whitening", unit variance, uncorrelated)

Proof 1: new norm SVD $\to$ CCA eigenvalue equations.

$$\max_v \frac{\|Av\|_y^2}{\|v\|_x^2} = \max_v \frac{\|MAv\|^2}{\|Lv\|^2} = \max_v \frac{\|\langle yy^T \rangle^{-1/2} Av\|^2}{\|\langle yy^T \rangle^{-1/2} v\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} A \langle xx^T \rangle^{1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{v'^T \langle xx^T \rangle^{-1/2} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'}{v'^T v'}$$

Eigenvalues of $\langle xx^T \rangle^{-1} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle$ = CCA

Multivariate

Claim: multivariate = replace variances by covariances.

CCA

$x \to \langle xx^T \rangle^{-1/2} x = Lx$, $y \to \langle yy^T \rangle^{-1/2} y = My$
(replace data by PCs, "whitening", unit variance, uncorrelated)
Proof 1: new norm SVD $\to$ CCA eigenvalue equations.

$$\max_v \frac{\|Av\|_y^2}{\|v\|_x^2} = \max_v \frac{\|MAv\|^2}{\|Lv\|^2} = \max_v \frac{\|\langle yy^T\rangle^{-1/2} Av\|^2}{\|\langle yy^T\rangle^{-1/2} v\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T\rangle^{-1/2} A \langle xx^T\rangle^{1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T\rangle^{-1/2} \langle yx^T\rangle \langle xx^T\rangle^{-1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{v'^T \langle xx^T\rangle^{-1/2} \langle xy^T\rangle \langle yy^T\rangle^{-1} \langle yx^T\rangle \langle xx^T\rangle^{-1/2} v'}{v'^T v'}$$

Eigenvalues of $\langle xx^T\rangle^{-1} \langle xy^T\rangle \langle yy^T\rangle^{-1} \langle yx^T\rangle$ = CCA

Multivariate

Claim: multivariate = replace variances by covariances.

CCA

$x \to \langle xx^T \rangle^{-1/2} x = Lx$, $y \to \langle yy^T \rangle^{-1/2} y = My$

(replace data by PCs, "whitening", unit variance, uncorrelated)

Proof 1: new norm SVD $\to$ CCA eigenvalue equations.

$$\max_v \frac{\|Av\|_y^2}{\|v\|_x^2} = \max_v \frac{\|MAv\|^2}{\|Lv\|^2} = \max_v \frac{\|\langle yy^T \rangle^{-1/2} Av\|^2}{\|\langle yy^T \rangle^{-1/2} v\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} A \langle xx^T \rangle^{1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{v'^T \langle xx^T \rangle^{-1/2} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'}{v'^T v'}$$

Eigenvalues of $\langle xx^T \rangle^{-1} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle$ = CCA

Multivariate

Claim: multivariate = replace variances by covariances.

CCA

$x \to \langle xx^T \rangle^{-1/2} x = Lx$, $y \to \langle yy^T \rangle^{-1/2} y = My$

(replace data by PCs, "whitening", unit variance, uncorrelated)

Proof 1: new norm SVD $\to$ CCA eigenvalue equations.

$$\max_v \frac{\|Av\|_y^2}{\|v\|_x^2} = \max_v \frac{\|MAv\|^2}{\|Lv\|^2} = \max_v \frac{\|\langle yy^T \rangle^{-1/2} Av\|^2}{\|\langle yy^T \rangle^{-1/2} v\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} A \langle xx^T \rangle^{1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{v'^T \langle xx^T \rangle^{-1/2} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'}{v'^T v'}$$

Eigenvalues of $\langle xx^T \rangle^{-1} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle$ = CCA

Multivariate

Claim: multivariate = replace variances by covariances.

CCA

$x \to \langle xx^T \rangle^{-1/2} x = Lx,\ y \to \langle yy^T \rangle^{-1/2} y = My$

(replace data by PCs, "whitening", unit variance, uncorrelated)

Proof 1: new norm SVD $\to$ CCA eigenvalue equations.

$$\max_v \frac{\|Av\|_y^2}{\|v\|_x^2} = \max_v \frac{\|MAv\|^2}{\|Lv\|^2} = \max_v \frac{\|\langle yy^T \rangle^{-1/2} Av\|^2}{\|\langle yy^T \rangle^{-1/2} v\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} A \langle xx^T \rangle^{1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'\|^2}{\|v'\|^2}$$

$$= \max_{v'} \frac{v'^T \langle xx^T \rangle^{-1/2} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'}{v'^T v'}$$

Eigenvalues of $\langle xx^T \rangle^{-1} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle$ = CCA

Multivariate

Claim: multivariate = replace variances by covariances.

CCA

$x \to \langle xx^T \rangle^{-1/2} x = Lx,\ y \to \langle yy^T \rangle^{-1/2} y = My$

(replace data by PCs, “whitening”, unit variance, uncorrelated)

Proof 1: new norm SVD $\to$ CCA eigenvalue equations.

$$
\begin{aligned}
\max_v \frac{\|Av\|_y^2}{\|v\|_x^2} &= \max_v \frac{\|MAv\|^2}{\|Lv\|^2} = \max_v \frac{\|\langle yy^T \rangle^{-1/2} Av\|^2}{\|\langle yy^T \rangle^{-1/2} v\|^2} \\
&= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} A \langle xx^T \rangle^{1/2} v'\|^2}{\|v'\|^2} \\
&= \max_{v'} \frac{\|\langle yy^T \rangle^{-1/2} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'\|^2}{\|v'\|^2} \\
&= \max_{v'} \frac{v'^T \langle xx^T \rangle^{-1/2} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle \langle xx^T \rangle^{-1/2} v'}{v'^T v'}
\end{aligned}
$$

Eigenvalues of $\langle xx^T \rangle^{-1} \langle xy^T \rangle \langle yy^T \rangle^{-1} \langle yx^T \rangle$ = CCA

Proof 2: SVD of transformed regression MAL^{-1} maximizes correlation.

$$A = \langle yx^T \rangle \langle xx^T \rangle^{-1}, M = \langle yy^T \rangle^{-1/2}, L = \langle xx^T \rangle^{-1/2}$$

$$\begin{aligned}
\max_{u,v} \frac{u^T MAL^{-1} v}{\|u\|\|v\|} &= \max_{u,v} \frac{\boxed{u^T \langle yy^T \rangle^{-1/2}} \langle yx^T \rangle \boxed{\langle xx^T \rangle^{-1/2} v}}{\|u\|\|v\|} \\
&= \max_{u',v'} \frac{u'^T \langle yx^T \rangle v'}{\sqrt{u'^T \langle yy^T \rangle u'} \sqrt{v'^T \langle xx^T \rangle v'}} \\
& \qquad u' = \langle yy^T \rangle^{-1/2} u, v' = \langle xx^T \rangle^{-1/2} v \\
&= \max_{u',v'} \frac{\mathrm{cov}(u'^T y, v'^T x)}{\sqrt{\mathrm{var}(u'^T y)} \sqrt{\mathrm{var}(v'^T x)}} \\
&= \max_{u',v'} \text{correlation } (u'^T y, v'^T x)
\end{aligned}$$

Finds projections of the data that maximize correlation.

Proof 2: SVD of transformed regression MAL^{-1} maximizes correlation.

$$A = \langle yx^T \rangle \langle xx^T \rangle^{-1}, M = \langle yy^T \rangle^{-1/2}, L = \langle xx^T \rangle^{-1/2}$$

$$\begin{aligned}
\max_{u,v} \frac{u^T MAL^{-1} v}{\|u\|\|v\|} &= \max_{u,v} \frac{\boxed{u^T \langle yy^T \rangle^{-1/2}} \langle yx^T \rangle \boxed{\langle xx^T \rangle^{-1/2} v}}{\|u\|\|v\|} \\
&= \max_{u',v'} \frac{u'^T \langle yx^T \rangle v'}{\sqrt{u'^T \langle yy^T \rangle u'}\sqrt{v'^T \langle xx^T \rangle v'}} \\
& \qquad\qquad {\scriptstyle u' = \langle yy^T \rangle^{-1/2} u, v' = \langle xx^T \rangle^{-1/2} v} \\
&= \max_{u',v'} \frac{\mathrm{cov}(u'^T y, v'^T x)}{\sqrt{\mathrm{var}(u'^T y)}\sqrt{\mathrm{var}(v'^T x)}} \\
&= \max_{u',v'} \text{ correlation } (u'^T y, v'^T x)
\end{aligned}$$

Finds projections of the data that maximize correlation.

Proof 2: SVD of transformed regression MAL^{-1} maximizes correlation.

$$A = \langle yx^T \rangle \langle xx^T \rangle^{-1}, M = \langle yy^T \rangle^{-1/2}, L = \langle xx^T \rangle^{-1/2}$$

$$\max_{u,v} \frac{u^T MAL^{-1} v}{\|u\|\|v\|} = \max_{u,v} \frac{\boxed{u^T \langle yy^T \rangle^{-1/2}} \langle yx^T \rangle \boxed{\langle xx^T \rangle^{-1/2} v}}{\|u\|\|v\|}$$

$$= \max_{u',v'} \frac{u'^T \langle yx^T \rangle v'}{\sqrt{u'^T \langle yy^T \rangle u'} \sqrt{v'^T \langle xx^T \rangle v'}}$$

$$u' = \langle yy^T \rangle^{-1/2} u, v' = \langle xx^T \rangle^{-1/2} v$$

$$= \max_{u',v'} \frac{\text{cov}(u'^T y, v'^T x)}{\sqrt{\text{var}(u'^T y)} \sqrt{\text{var}(v'^T x)}}$$

$$= \max_{u',v'} \text{correlation}\,(u'^T y, v'^T x)$$

Finds projections of the data that maximize correlation.

Proof 2: SVD of transformed regression MAL^{-1} maximizes correlation.

$$A = \langle yx^T \rangle \langle xx^T \rangle^{-1}, M = \langle yy^T \rangle^{-1/2}, L = \langle xx^T \rangle^{-1/2}$$

$$\max_{u,v} \frac{u^T MAL^{-1} v}{\|u\| \|v\|} = \max_{u,v} \frac{\boxed{u^T \langle yy^T \rangle^{-1/2}} \langle yx^T \rangle \boxed{\langle xx^T \rangle^{-1/2} v}}{\|u\| \|v\|}$$

$$= \max_{u',v'} \frac{u'^T \langle yx^T \rangle v'}{\sqrt{u'^T \langle yy^T \rangle u'} \sqrt{v'^T \langle xx^T \rangle v'}}$$

$$u' = \langle yy^T \rangle^{-1/2} u, v' = \langle xx^T \rangle^{-1/2} v$$

$$= \max_{u',v'} \frac{\text{cov}(u'^T y, v'^T x)}{\sqrt{\text{var}(u'^T y)} \sqrt{\text{var}(v'^T x)}}$$

$$= \max_{u',v'} \text{correlation}\, (u'^T y, v'^T x)$$

Finds projections of the data that maximize correlation.

Proof 2: SVD of transformed regression MAL^{-1} maximizes correlation.

$$A = \langle yx^T \rangle \langle xx^T \rangle^{-1}, M = \langle yy^T \rangle^{-1/2}, L = \langle xx^T \rangle^{-1/2}$$

$$\max_{u,v} \frac{u^T MAL^{-1} v}{\|u\|\|v\|} = \max_{u,v} \frac{\boxed{u^T \langle yy^T \rangle^{-1/2}} \langle yx^T \rangle \boxed{\langle xx^T \rangle^{-1/2} v}}{\|u\|\|v\|}$$

$$= \max_{u',v'} \frac{u'^T \langle yx^T \rangle v'}{\sqrt{u'^T \langle yy^T \rangle u'} \sqrt{v'^T \langle xx^T \rangle v'}}$$

$$u' = \langle yy^T \rangle^{-1/2} u, v' = \langle xx^T \rangle^{-1/2} v$$

$$= \max_{u',v'} \frac{\text{cov}(u'^T y, v'^T x)}{\sqrt{\text{var}(u'^T y)} \sqrt{\text{var}(v'^T x)}}$$

$$= \max_{u',v'} \text{correlation}\,(u'^T y, v'^T x)$$

Finds projections of the data that maximize correlation.

Proof 2: SVD of transformed regression MAL^{-1} maximizes correlation.

$$A = \langle yx^T \rangle \langle xx^T \rangle^{-1}, M = \langle yy^T \rangle^{-1/2}, L = \langle xx^T \rangle^{-1/2}$$

$$\begin{aligned}
\max_{u,v} \frac{u^T MAL^{-1} v}{\|u\|\|v\|} &= \max_{u,v} \frac{\boxed{u^T \langle yy^T \rangle^{-1/2}} \langle yx^T \rangle \boxed{\langle xx^T \rangle^{-1/2} v}}{\|u\|\|v\|} \\
&= \max_{u',v'} \frac{u'^T \langle yx^T \rangle v'}{\sqrt{u'^T \langle yy^T \rangle u'} \sqrt{v'^T \langle xx^T \rangle v'}} \\
&\qquad\qquad u' = \langle yy^T \rangle^{-1/2} u, v' = \langle xx^T \rangle^{-1/2} v \\
&= \max_{u',v'} \frac{\operatorname{cov}(u'^T y, v'^T x)}{\sqrt{\operatorname{var}(u'^T y)} \sqrt{\operatorname{var}(v'^T x)}} \\
&= \max_{u',v'} \text{ correlation } (u'^T y, v'^T x)
\end{aligned}$$

Finds projections of the data that maximize correlation.

Proof 2: SVD of transformed regression MAL^{-1} maximizes correlation.

$$A = \langle yx^T \rangle \langle xx^T \rangle^{-1}, M = \langle yy^T \rangle^{-1/2}, L = \langle xx^T \rangle^{-1/2}$$

$$\max_{u,v} \frac{u^T MAL^{-1} v}{\|u\|\|v\|} = \max_{u,v} \frac{\boxed{u^T \langle yy^T \rangle^{-1/2}} \langle yx^T \rangle \boxed{\langle xx^T \rangle^{-1/2} v}}{\|u\|\|v\|}$$

$$= \max_{u',v'} \frac{u'^T \langle yx^T \rangle v'}{\sqrt{u'^T \langle yy^T \rangle u'}\sqrt{v'^T \langle xx^T \rangle v'}}$$

$$u' = \langle yy^T \rangle^{-1/2} u, v' = \langle xx^T \rangle^{-1/2} v$$

$$= \max_{u',v'} \frac{\operatorname{cov}(u'^T y, v'^T x)}{\sqrt{\operatorname{var}(u'^T y)}\sqrt{\operatorname{var}(v'^T x)}}$$

$$= \max_{u',v'} \text{ correlation } (u'^T y, v'^T x)$$

Finds projections of the data that maximize correlation.

Summary

- Common coupled pattern methods are SVDs of the matrix of regression coefficients.
- Different pattern methods correspond to SVD with different norms – measure different properties of the regression.
- If a complete set of patterns are used, all methods give the same prediction model.
- One method can be transformed to another by transforming the data.

$$
\begin{aligned}
\mathrm{MCA}[X, Y] &= \mathrm{SVD}\left[\langle yx^T \rangle\right] \\
\mathrm{CCA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x,\ \langle yy^T \rangle^{-1/2} y\right] \\
\mathrm{RDA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x,\ y\right] \\
\mathrm{PPA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x,\ (\mathrm{Diag}\langle yy^T \rangle)^{-1/2} y\right]
\end{aligned}
$$

Summary

- Common coupled pattern methods are SVDs of the matrix of regression coefficients.
- Different pattern methods correspond to SVD with different norms – measure different properties of the regression.
- If a complete set of patterns are used, all methods give the same prediction model.
- One method can be transformed to another by transforming the data.

$$\mathrm{MCA}[X, Y] = \mathrm{SVD}\left[\langle yx^T \rangle\right]$$

$$\mathrm{CCA}[X, Y] = \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x, \langle yy^T \rangle^{-1/2} y\right]$$

$$\mathrm{RDA}[X, Y] = \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x, y\right]$$

$$\mathrm{PPA}[X, Y] = \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x, (\mathrm{Diag}\langle yy^T \rangle)^{-1/2} y\right]$$

Summary

- Common coupled pattern methods are SVDs of the matrix of regression coefficients.
- Different pattern methods correspond to SVD with different norms – measure different properties of the regression.
- If a complete set of patterns are used, all methods give the same prediction model.
- One method can be transformed to another by transforming the data.

$$\begin{aligned}
\mathrm{MCA}[X, Y] &= \mathrm{SVD}\left[\langle yx^T \rangle\right] \\
\mathrm{CCA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x, \langle yy^T \rangle^{-1/2} y\right] \\
\mathrm{RDA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x, y\right] \\
\mathrm{PPA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x, (\mathrm{Diag}\langle yy^T \rangle)^{-1/2} y\right]
\end{aligned}$$

Summary

- Common coupled pattern methods are SVDs of the matrix of regression coefficients.
- Different pattern methods correspond to SVD with different norms – measure different properties of the regression.
- If a complete set of patterns are used, all methods give the same prediction model.
- One method can be transformed to another by transforming the data.

$$\begin{aligned}
\mathrm{MCA}[X, Y] &= \mathrm{SVD}\left[\langle yx^T\rangle\right] \\
\mathrm{CCA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T\rangle^{-1/2}x,\ \langle yy^T\rangle^{-1/2}y\right] \\
\mathrm{RDA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T\rangle^{-1/2}x,\ y\right] \\
\mathrm{PPA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T\rangle^{-1/2}x,\ (\mathrm{Diag}\langle yy^T\rangle)^{-1/2}y\right]
\end{aligned}$$

Summary

- Common coupled pattern methods are SVDs of the matrix of regression coefficients.
- Different pattern methods correspond to SVD with different norms – measure different properties of the regression.
- If a complete set of patterns are used, all methods give the same prediction model.
- One method can be transformed to another by transforming the data.

$$
\begin{aligned}
\mathrm{MCA}[X, Y] &= \mathrm{SVD}\left[\langle yx^T \rangle\right] \\
\mathrm{CCA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x, \langle yy^T \rangle^{-1/2} y\right] \\
\mathrm{RDA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x, y\right] \\
\mathrm{PPA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x, (\mathrm{Diag}\langle yy^T \rangle)^{-1/2} y\right]
\end{aligned}
$$

Summary

- Common coupled pattern methods are SVDs of the matrix of regression coefficients.
- Different pattern methods correspond to SVD with different norms – measure different properties of the regression.
- If a complete set of patterns are used, all methods give the same prediction model.
- One method can be transformed to another by transforming the data.

$$
\begin{aligned}
\mathrm{MCA}[X, Y] &= \mathrm{SVD}\left[\langle yx^T\rangle\right] \\
\mathrm{CCA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T\rangle^{-1/2}x,\ \langle yy^T\rangle^{-1/2}y\right] \\
\mathrm{RDA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T\rangle^{-1/2}x,\ y\right] \\
\mathrm{PPA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T\rangle^{-1/2}x,\ (\mathrm{Diag}\langle yy^T\rangle)^{-1/2}y\right]
\end{aligned}
$$

Summary

- Common coupled pattern methods are SVDs of the matrix of regression coefficients.
- Different pattern methods correspond to SVD with different norms – measure different properties of the regression.
- If a complete set of patterns are used, all methods give the same prediction model.
- One method can be transformed to another by transforming the data.

$$
\begin{aligned}
\mathrm{MCA}[X, Y] &= \mathrm{SVD}\left[\langle yx^T \rangle\right] \\
\mathrm{CCA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x,\ \langle yy^T \rangle^{-1/2} y\right] \\
\mathrm{RDA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x,\ y\right] \\
\mathrm{PPA}[X, Y] &= \mathrm{MCA}\left[\langle xx^T \rangle^{-1/2} x,\ (\mathrm{Diag}\langle yy^T \rangle)^{-1/2} y\right]
\end{aligned}
$$

Summary

- Common coupled pattern methods are SVDs of the matrix of regression coefficients.
- Different pattern methods correspond to SVD with different norms – measure different properties of the regression.
- If a complete set of patterns are used, all methods give the same prediction model.
- One method can be transformed to another by transforming the data.

$$
\begin{aligned}
\text{MCA}[X, Y] &= \text{SVD}\left[\langle yx^T \rangle\right] \\
\text{CCA}[X, Y] &= \text{MCA}\left[\langle xx^T \rangle^{-1/2} x,\ \langle yy^T \rangle^{-1/2} y\right] \\
\text{RDA}[X, Y] &= \text{MCA}\left[\langle xx^T \rangle^{-1/2} x,\ y\right] \\
\text{PPA}[X, Y] &= \text{MCA}\left[\langle xx^T \rangle^{-1/2} x,\ (\text{Diag}\langle yy^T \rangle)^{-1/2} y\right]
\end{aligned}
$$