Measuring predictability with singular vectors

Michael K. Tippett[1] Timothy DelSole[2]

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

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

4/17/07 / Applied Mathematics Colloquium

Main ideas

Applications of the SVD to:

Diagnosing linear dynamics in weather and climate.

SVD used to

- find the errors that grow the most (weather)
- find the signals that grow the most (ENSO)

How can the singular value decomposition be used to:

- Measure predictability
- Identify "predictable components".

Main ideas

Applications of the SVD to:

Diagnosing linear dynamics in weather and climate.

SVD used to

- find the errors that grow the most (weather)
- find the signals that grow the most (ENSO)

How can the singular value decomposition be used to:

- Measure predictability
- Identify "predictable components".

Main ideas

Applications of the SVD to:

Diagnosing linear dynamics in weather and climate.

SVD used to

- find the errors that grow the most (weather)
- find the signals that grow the most (ENSO)

How can the singular value decomposition be used to:

- Measure predictability
- Identify "predictable components".

Main ideas

Applications of the SVD to:

Diagnosing linear dynamics in weather and climate.

SVD used to

- find the errors that grow the most (weather)
- find the signals that grow the most (ENSO)

How can the singular value decomposition be used to:

- Measure predictability
- Identify “predictable components”.

Outline

- The SVD.
- Applications to error and signal growth.
- Why are norms important?
- How to measure predictability.
- How to use the SVD to measure predictability.
- Sea surface temperature example.

The singular value decomposition

$$
\begin{aligned}
A &= U\Sigma V^T \quad m \times n \\
&= [u_1, \dots, u_m]\mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_p)[v_1, \dots, v_n]^T \\
&= \sum_{i=1}^{p} \sigma_i u_i v_i^T
\end{aligned}
$$

U, V are orthogonal.
Σ is diagonal. $\sigma_1 \geq \sigma_2 \geq \dots \sigma_p \geq 0$.
Properties

- A maps a unit sphere to an ellipsoid. $Av_i = \sigma_i u_i$.
- $\|A\| = \max_x \frac{\|Ax\|}{\|x\|} = \sigma_1$
- u_i = eigenvectors of AA^T
  σ_i^2 = eigenvalues of AA^T

The singular value decomposition

$$\begin{aligned} A &= U\Sigma V^T \quad m \times n \\ &= [u_1, \ldots, u_m]\text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_p)[v_1, \ldots, v_n]^T \\ &= \sum_{i=1}^{p} \sigma_i u_i v_i^T \end{aligned}$$

U, V are orthogonal.
Σ is diagonal. $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_p \geq 0$.
Properties

- A maps a unit sphere to an ellipsoid. $Av_i = \sigma_i u_i$.
- $\|A\| = \max_x \frac{\|Ax\|}{\|x\|} = \sigma_1$
- u_i = eigenvectors of AA^T
  σ_i^2 = eigenvalues of AA^T

The singular value decomposition

$$\begin{aligned} A &= U\Sigma V^T \quad m \times n \\ &= [u_1, \dots, u_m]\mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_p)[v_1, \dots, v_n]^T \\ &= \sum_{i=1}^{p} \sigma_i u_i v_i^T \end{aligned}$$

U, V are orthogonal.
Σ is diagonal. $\sigma_1 \geq \sigma_2 \geq \dots \sigma_p \geq 0$.
Properties

- A maps a unit sphere to an ellipsoid. $Av_i = \sigma_i u_i$.
- $\|A\| = \max_x \frac{\|Ax\|}{\|x\|} = \sigma_1$
- u_i = eigenvectors of AA^T
  σ_i^2 = eigenvalues of AA^T

The singular value decomposition

$$
\begin{aligned}
A &= U\Sigma V^T \quad m \times n \\
&= [u_1, \ldots, u_m]\mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_p)[v_1, \ldots, v_n]^T \\
&= \sum_{i=1}^{p} \sigma_i u_i v_i^T
\end{aligned}
$$

U, V are orthogonal.
Σ is diagonal. $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_p \geq 0$.
Properties

- A maps a unit sphere to an ellipsoid. $Av_i = \sigma_i u_i$.
- $\|A\| = \max_x \dfrac{\|Ax\|}{\|x\|} = \sigma_1$
- u_i = eigenvectors of AA^T
  σ_i^2 = eigenvalues of AA^T

Example

$$A = U\Sigma V^T$$

$$A = \begin{bmatrix} \frac{3}{4} & 2.3 \\ 0 & \frac{3}{4} \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 2.52 & 0 \\ 0 & 0.223 \end{bmatrix}$$

$\sigma_1 \geq |\lambda_1(A)|$.
Minimum growth when A is normal

Application: Growth of small errors
(Lorenz, Farrell, Palmer)

Linear dynamics

$$f = A(a + \delta)$$

f = forecast.
a = initial condition (analysis).
δ = perturbation (analysis error).

Which errors are amplified the most?

$$\max_{\delta} \frac{\|A\delta\|}{\|\delta\|} = \sigma_1 \,, \quad Av_1 = \sigma_1 u_1$$

$\sigma_1 \gg 1$ = large amplification of errors
Large errors in u_1 = less predictable?

Application: Growth of small errors
(Lorenz, Farrell, Palmer)

Linear dynamics

$$f = A(a + \delta)$$

f = forecast.
a = initial condition (analysis).
δ = perturbation (analysis error).

Which errors are amplified the most?

$$\max_{\delta} \frac{\|A\delta\|}{\|\delta\|} = \sigma_1 \,, \quad Av_1 = \sigma_1 u_1$$

$\sigma_1 \gg 1$ = large amplification of errors
Large errors in u_1 = less predictable?

Application: Growth of small errors
(Lorenz, Farrell, Palmer)

Linear dynamics

$$f = A(a + \delta)$$

f = forecast.
a = initial condition (analysis).
δ = perturbation (analysis error).

Which errors are amplified the most?

$$\max_{\delta} \frac{\|A\delta\|}{\|\delta\|} = \sigma_1 \,, \quad Av_1 = \sigma_1 u_1$$

$\sigma_1 \gg 1$ = large amplification of errors
Large errors in u_1 = less predictable?

Application: Growth of small errors
(Lorenz, Farrell, Palmer)

Linear dynamics

$$f = A(a + \delta)$$

f = forecast.
a = initial condition (analysis).
δ = perturbation (analysis error).

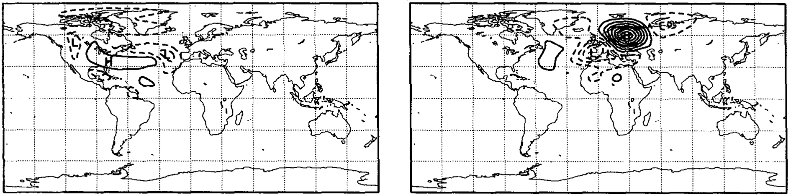
Which errors are amplified the most?

$$\max_{\delta} \frac{\|A\delta\|}{\|\delta\|} = \sigma_1\,, \quad Av_1 = \sigma_1 u_1$$

$\sigma_1 \gg 1$ = large amplification of errors
Large errors in u_1 = less predictable?

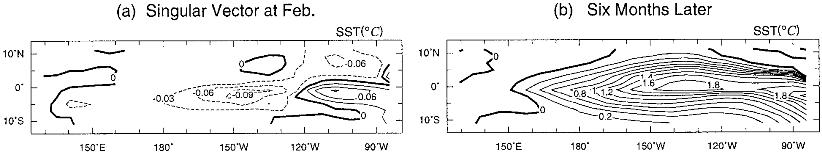
Examples: Perturbation growth

Weather forecasting model (ψ_{200} 3-day 20x)



Buizza and Palmer, 1995 Lanczos method, adjoint

ENSO model



Xue et al., 1997 finite differencing

Application: Signal growth
(Penland and Sardeshmukh)

Linear dynamics + stochastic forcing.

$$f = Ai + \epsilon$$

f = forecast
i = initial condition (no error)
ϵ = stochastic forcing (model error)

Which initial conditions grow “above the noise”?

$$\max_i \frac{\|Ai\|}{\|i\|} = \sigma_1 \,, \quad Av_1 = \sigma_1 u_1$$

$\sigma_1 \gg 1$ = large amplification of initial condition
Large “signal” in u_1 = more predictable?

Application: Signal growth
(Penland and Sardeshmukh)

Linear dynamics + stochastic forcing.

$$f = Ai + \epsilon$$

f = forecast
i = initial condition (no error)
ϵ = stochastic forcing (model error)

Which initial conditions grow "above the noise"?

$$\max_i \frac{\|Ai\|}{\|i\|} = \sigma_1 \,, \quad Av_1 = \sigma_1 u_1$$

$\sigma_1 \gg 1$ = large amplification of initial condition
Large "signal" in u_1 = more predictable?

Application: Signal growth
(Penland and Sardeshmukh)

Linear dynamics + stochastic forcing.

$$f = Ai + \epsilon$$

f = forecast
i = initial condition (no error)
ϵ = stochastic forcing (model error)

Which initial conditions grow “above the noise”?

$$\max_i \frac{\|Ai\|}{\|i\|} = \sigma_1 , \quad Av_1 = \sigma_1 u_1$$

$\sigma_1 \gg 1$ = large amplification of initial condition
Large “signal” in u_1 = more predictable?

Application: Signal growth
(Penland and Sardeshmukh)

Linear dynamics + stochastic forcing.

$$f = Ai + \epsilon$$

f = forecast
i = initial condition (no error)
ϵ = stochastic forcing (model error)

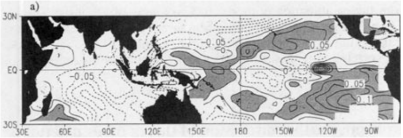
Which initial conditions grow "above the noise"?

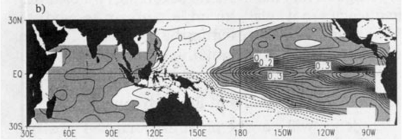
$$\max_i \frac{\|Ai\|}{\|i\|} = \sigma_1\,, \quad Av_1 = \sigma_1 u_1$$

$\sigma_1 \gg 1$ = large amplification of initial condition
Large "signal" in u_1 = more predictable?

Signal growth
ENSO model

initial | final

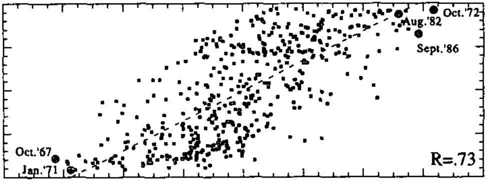




Penland and Sardeshmukh, 1995

Observations projected onto the singular vectors

final



initial

Paradox?

What are the consequences for predictability?

- SVD finds maximum error growth.
  Large singular value = less predictability.
- SVD finds maximums signal growth.
  Large singular value = more predictability?

Solution: “different” SVDs for different problems.

- Is the “error” SVD consistent with the “signal” SVD?
  - Same components, opposite ordering.
- What is the SVD for measuring predictability?
  - How is it related to other predictability analysis methods?

Paradox?

What are the consequences for predictability?

- SVD finds maximum error growth.
  Large singular value = less predictability.
- SVD finds maximums signal growth.
  Large singular value = more predictability?

Solution: “different” SVDs for different problems.

- Is the “error” SVD consistent with the “signal” SVD?
  - Same components, opposite ordering.
- What is the SVD for measuring predictability?
  - How is it related to other predictability analysis methods?

Paradox?

What are the consequences for predictability?

- SVD finds maximum error growth.
  Large singular value = less predictability.
- SVD finds maximums signal growth.
  Large singular value = more predictability?

Solution: “different” SVDs for different problems.

- Is the “error” SVD consistent with the “signal” SVD?
  - Same components, opposite ordering.
- What is the SVD for measuring predictability?
  - How is it related to other predictability analysis methods?

Paradox?

What are the consequences for predictability?

- SVD finds maximum error growth.
  Large singular value = less predictability.
- SVD finds maximums signal growth.
  Large singular value = more predictability?

Solution: “different” SVDs for different problems.

- Is the “error” SVD consistent with the “signal” SVD?
  - Same components, opposite ordering.
- What is the SVD for measuring predictability?
  - How is it related to other predictability analysis methods?

Paradox?

What are the consequences for predictability?

- SVD finds maximum error growth.
  Large singular value = less predictability.
- SVD finds maximums signal growth.
  Large singular value = more predictability?

Solution: “different” SVDs for different problems.

- Is the “error” SVD consistent with the “signal” SVD?
  - Same components, opposite ordering.
- What is the SVD for measuring predictability?
  - How is it related to other predictability analysis methods?

Paradox?

What are the consequences for predictability?

- SVD finds maximum error growth.
  Large singular value = less predictability.
- SVD finds maximums signal growth.
  Large singular value = more predictability?

Solution: “different” SVDs for different problems.

- Is the “error” SVD consistent with the “signal” SVD?
  - Same components, opposite ordering.
- What is the SVD for measuring predictability?
  - How is it related to other predictability analysis methods?

Paradox?

What are the consequences for predictability?

- SVD finds maximum error growth.
  Large singular value = less predictability.
- SVD finds maximums signal growth.
  Large singular value = more predictability?

Solution: “different” SVDs for different problems.

- Is the “error” SVD consistent with the “signal” SVD?
  - Same components, opposite ordering.
- What is the SVD for measuring predictability?
  - How is it related to other predictability analysis methods?

SVD and norms

Usual SVD maximizes

$$\frac{\|Ax\|^2}{\|x\|^2}$$

Cannot apply directly to multivariate problems.

Generalized SVD maximizes

$$\frac{x^T M^T M x}{x^T L^T L x} = \frac{\|Ax\|_M^2}{\|x\|_L^2}$$

$$\text{SVD}(MAL^{-1}) = \text{generalized SVD}(A)$$

$$\begin{aligned}
\max_x \frac{\|MAL^{-1}x\|^2}{\|x\|^2} &= \max_x \frac{x^T L^{-T} A^T M^T M A L^{-1} x}{x^T x} \\
&= \max_y \frac{y^T A^T M^T M A y}{y^T L^T L y} \qquad (y = L^{-1}x) \\
&= \max_y \frac{\|Ay\|_M^2}{\|y\|_L^2}
\end{aligned}$$

SVD and norms

Usual SVD maximizes

$$\frac{\|Ax\|^2}{\|x\|^2}$$

Cannot apply directly to multivariate problems.

Generalized SVD maximizes

$$\frac{x^T M^T M x}{x^T L^T L x} \equiv \frac{\|Ax\|_M^2}{\|x\|_L^2}$$

$$\text{SVD}(MAL^{-1}) = \text{generalized SVD}(A)$$

$$\max_x \frac{\|MAL^{-1}x\|^2}{\|x\|^2} = \max_x \frac{x^T L^{-T} A^T M^T M A L^{-1} x}{x^T x}$$

$$= \max_y \frac{y^T A^T M^T M A y}{y^T L^T L y} \quad (y = L^{-1}x)$$

$$= \max_y \frac{\|Ay\|_M^2}{\|y\|_L^2}$$

SVD and norms

Usual SVD maximizes

$$\frac{\|Ax\|^2}{\|x\|^2}$$

Cannot apply directly to multivariate problems.

Generalized SVD maximizes

$$\frac{x^T M^T M x}{x^T L^T L x} \equiv \frac{\|Ax\|_M^2}{\|x\|_L^2}$$

$$\mathrm{SVD}(MAL^{-1}) = \text{generalized } \mathrm{SVD}(A)$$

$$\max_x \frac{\|MAL^{-1}x\|^2}{\|x\|^2} = \max_x \frac{x^T L^{-T} A^T M^T M A L^{-1} x}{x^T x}$$

$$= \max_y \frac{y^T A^T M^T M A y}{y^T L^T L y} \quad (y = L^{-1}x)$$

$$= \max_y \frac{\|Ay\|_M^2}{\|y\|_L^2}$$

SVD and norms

Usual SVD maximizes

$$\frac{\|Ax\|^2}{\|x\|^2}$$

Cannot apply directly to multivariate problems.

Generalized SVD maximizes

$$\frac{x^T M^T M x}{x^T L^T L x} \equiv \frac{\|Ax\|_M^2}{\|x\|_L^2}$$

$$\begin{aligned}
\text{SVD}(MAL^{-1}) &= \text{generalized SVD}(A) \\
\max_x \frac{\|MAL^{-1}x\|^2}{\|x\|^2} &= \max_x \frac{x^T L^{-T} A^T M^T M A L^{-1} x}{x^T x} \\
&= \max_y \frac{y^T A^T M^T M A y}{y^T L^T L y} \quad (y = L^{-1}x) \\
&= \max_y \frac{\|Ay\|_M^2}{\|y\|_L^2}
\end{aligned}$$

Which norms?

Initial norm

$$f = A(a+\delta)\,, \quad \max_{\delta} \frac{\|A\delta\|_M}{\|\delta\|_L}$$

Lorenz (1965) said choose L so that $L\delta$ is “random in that no direction is …preferred over any other direction.” (Sphere)

Translation: choose L so that $L\delta$ is “white” in space.

Pick L so that

$$\mathrm{Cov}(L\delta) = L\mathrm{Cov}(\delta)L^T = LC_\delta L^T = I$$

$L = C_\delta^{-1/2}$. ($\sim$ standardizing)

$$\sigma_1^2 = \max_{\delta} \frac{\|A\delta\|^2}{\|\delta\|_L^2} = \max_{\delta} \frac{\|A\delta\|^2}{\delta^T C_\delta^{-1}\delta} = \mathrm{SVD}(AC_\delta^{1/2})$$

Singular values measure: $\dfrac{\text{dynamical growth of } \delta}{\text{likelihood of } \delta}$

Which norms?
Initial norm

$$f = A(a+\delta)\,, \quad \max_{\delta} \frac{\|A\delta\|_M}{\|\delta\|_L}$$

Lorenz (1965) said choose L so that $L\delta$ is "random in that no direction is … preferred over any other direction." (Sphere)

Translation: choose L so that $L\delta$ is "white" in space.

Pick L so that

$$\mathrm{Cov}(L\delta) = L\mathrm{Cov}(\delta)L^T = LC_\delta L^T = I$$

$L = C_\delta^{-1/2}$. ($\sim$ standardizing)

$$\sigma_1^2 = \max_{\delta} \frac{\|A\delta\|^2}{\|\delta\|_L^2} = \max_{\delta} \frac{\|A\delta\|^2}{\delta^T C_\delta^{-1}\delta} = \mathrm{SVD}(AC_\delta^{1/2})$$

Singular values measure: $\dfrac{\text{dynamical growth of } \delta}{\text{likelihood of } \delta}$

Which norms?

Initial norm

$$f = A(a + \delta)\,, \quad \max_{\delta} \frac{\|A\delta\|_M}{\|\delta\|_L}$$

Lorenz (1965) said choose L so that $L\delta$ is "random in that no direction is . . . preferred over any other direction." (Sphere)
Translation: choose L so that $L\delta$ is "white" in space.
Pick L so that

$$\mathrm{Cov}(L\delta) = L\mathrm{Cov}(\delta)L^T = LC_\delta L^T = I$$

$L = C_\delta^{-1/2}$. ($\sim$ standardizing)

$$\sigma_1^2 = \max_{\delta} \frac{\|A\delta\|^2}{\|\delta\|_L^2} = \max_{\delta} \frac{\|A\delta\|^2}{\delta^T C_\delta^{-1}\delta} = \mathrm{SVD}(AC_\delta^{1/2})$$

Singular values measure: $\dfrac{\text{dynamical growth of } \delta}{\text{likelihood of } \delta}$

Which norms?
Initial norm

$$f = A(a + \delta)\,, \quad \max_\delta \frac{\|A\delta\|_M}{\|\delta\|_L}$$

Lorenz (1965) said choose L so that $L\delta$ is “random in that no direction is . . . preferred over any other direction.” (Sphere)
Translation: choose L so that $L\delta$ is “white” in space.
Pick L so that

$$\text{Cov}(L\delta) = L\text{Cov}(\delta)L^T = LC_\delta L^T = I$$

$L = C_\delta^{-1/2}$. ($\sim$ standardizing)

$$\sigma_1^2 = \max_\delta \frac{\|A\delta\|^2}{\|\delta\|_L^2} = \max_\delta \frac{\|A\delta\|^2}{\delta^T C_\delta^{-1}\delta} = \text{SVD}(AC_\delta^{1/2})$$

Singular values measure: $\dfrac{\text{dynamical growth of } \delta}{\text{likelihood of } \delta}$

Which norms?

Initial norm

$$f = A(a+\delta)\,, \quad \max_{\delta} \frac{\|A\delta\|_M}{\|\delta\|_L}$$

Lorenz (1965) said choose L so that $L\delta$ is "random in that no direction is ... preferred over any other direction." (Sphere)
Translation: choose L so that $L\delta$ is "white" in space.
Pick L so that

$$\text{Cov}(L\delta) = L\text{Cov}(\delta)L^T = LC_\delta L^T = I$$

$L = C_\delta^{-1/2}$. ($\sim$ standardizing)

$$\sigma_1^2 = \max_{\delta} \frac{\|A\delta\|^2}{\|\delta\|_L^2} = \max_{\delta} \frac{\|A\delta\|^2}{\delta^T C_\delta^{-1}\delta} = \text{SVD}(AC_\delta^{1/2})$$

Singular values measure: $\dfrac{\text{dynamical growth of } \delta}{\text{likelihood of } \delta}$

Which norms?
Initial norm

$$f = A(a + \delta)\,, \quad \max_{\delta} \frac{\|A\delta\|_M}{\|\delta\|_L}$$

Lorenz (1965) said choose L so that $L\delta$ is "random in that no direction is ... preferred over any other direction." (Sphere)
Translation: choose L so that $L\delta$ is "white" in space.
Pick L so that

$$\mathrm{Cov}(L\delta) = L\mathrm{Cov}(\delta)L^T = LC_\delta L^T = I$$

$L = C_\delta^{-1/2}$. ($\sim$ standardizing)

$$\sigma_1^2 = \max_{\delta} \frac{\|A\delta\|^2}{\|\delta\|_L^2} = \max_{\delta} \frac{\|A\delta\|^2}{\delta^T C_\delta^{-1} \delta} = \mathrm{SVD}(AC_\delta^{1/2})$$

Singular values measure: $\dfrac{\text{dynamical growth of } \delta}{\text{likelihood of } \delta}$

Forecast errors

$$f = A(a + \delta)$$

The forecast covariance is

$$\begin{aligned} C_f &= \mathrm{Cov}(f) = \mathrm{Cov}(A\delta) = A\mathrm{Cov}(\delta)A^T = AC_\delta A^T \\ &= (AC_\delta^{1/2})(AC_\delta^{1/2})^T \,. \end{aligned}$$

Eigenvalue decomposition of C_f = SVD($AC_\delta^{1/2}$)
Optimal.
(Houtekamer, Palmer, Tribbia, Ehrendorfer)

Eigenvalues of $C_f = \sigma_i^2$.

Eigenvectors of $C_f = u_i$

Final norm? Energy? Errors in a particular region?

Are there norms so that the SVD measures predictability?

Forecast errors

$$f = A(a + \delta)$$

The forecast covariance is

$$\begin{aligned} C_f &= \mathrm{Cov}(f) = \mathrm{Cov}(A\delta) = A\mathrm{Cov}(\delta)A^T = AC_\delta A^T \\ &= (AC_\delta^{1/2})(AC_\delta^{1/2})^T\,. \end{aligned}$$

Eigenvalue decomposition of C_f = SVD$(AC_\delta^{1/2})$
Optimal.
(Houtekamer, Palmer, Tribbia, Ehrendorfer)

Eigenvalues of $C_f = \sigma_i^2$.

Eigenvectors of $C_f = u_i$

Final norm? Energy? Errors in a particular region?

Are there norms so that the SVD measures predictability?

Forecast errors

$$f = A(a + \delta)$$

The forecast covariance is

$$C_f = \text{Cov}(f) = \text{Cov}(A\delta) = A\text{Cov}(\delta)A^T = AC_\delta A^T$$
$$= (AC_\delta^{1/2})(AC_\delta^{1/2})^T .$$

Eigenvalue decomposition of C_f = SVD$(AC_\delta^{1/2})$
Optimal.
(Houtekamer, Palmer, Tribbia, Ehrendorfer)

Eigenvalues of C_f = σ_i^2.

Eigenvectors of C_f = u_i

Final norm? Energy? Errors in a particular region?

Are there norms so that the SVD measures predictability?

Forecast errors

$$f = A(a + \delta)$$

The forecast covariance is

$$\begin{aligned} C_f &= \mathrm{Cov}(f) = \mathrm{Cov}(A\delta) = A\mathrm{Cov}(\delta)A^T = AC_\delta A^T \\ &= (AC_\delta^{1/2})(AC_\delta^{1/2})^T . \end{aligned}$$

Eigenvalue decomposition of C_f = SVD($AC_\delta^{1/2}$)
Optimal.
(Houtekamer, Palmer, Tribbia, Ehrendorfer)

Eigenvalues of C_f = σ_i^2.

Eigenvectors of C_f = u_i

Final norm? Energy? Errors in a particular region?

Are there norms so that the SVD measures predictability?

Forecast errors

$$f = A(a + \delta)$$

The forecast covariance is

$$\begin{aligned} C_f &= \mathrm{Cov}(f) = \mathrm{Cov}(A\delta) = A\mathrm{Cov}(\delta)A^T = AC_\delta A^T \\ &= (AC_\delta^{1/2})(AC_\delta^{1/2})^T \,. \end{aligned}$$

Eigenvalue decomposition of C_f = SVD$(AC_\delta^{1/2})$
Optimal.
(Houtekamer, Palmer, Tribbia, Ehrendorfer)

Eigenvalues of C_f = σ_i^2.

Eigenvectors of C_f = u_i

Final norm? Energy? Errors in a particular region?

Are there norms so that the SVD measures predictability?

How to measure predictability

Predictability is . . .

- Relative to a baseline
  - usually climatology (historical frequencies).
- Difference between forecast and climatology.
- Univariate predictability $\sim \dfrac{\text{forecast uncertainty}}{\text{climatological uncertainty}}$.
  - relative error not absolute error
- Should be independent of units or variable weighting.
- Pred(A,B) = Pred(A)+Pred(B), if A and B are independent.
- Pred(A,B) = Pred(B) if A=f(B).

How to measure predictability

Predictability is ...

- Relative to a baseline
  - usually climatology (historical frequencies).
- Difference between forecast and climatology.
- Univariate predictability $\sim \dfrac{\text{forecast uncertainty}}{\text{climatological uncertainty}}$.
  - relative error not absolute error
- Should be independent of units or variable weighting.
- Pred(A,B) = Pred(A)+Pred(B), if A and B are independent.
- Pred(A,B) = Pred(B) if A=f(B).

How to measure predictability

Predictability is ...

- Relative to a baseline
  - usually climatology (historical frequencies).
- Difference between forecast and climatology.
- Univariate predictability $\sim \dfrac{\text{forecast uncertainty}}{\text{climatological uncertainty}}$.
  - relative error not absolute error
- Should be independent of units or variable weighting.
- Pred(A,B) = Pred(A)+Pred(B), if A and B are independent.
- Pred(A,B) = Pred(B) if A=f(B).

How to measure predictability

Predictability is ...

- Relative to a baseline
  - usually climatology (historical frequencies).
- Difference between forecast and climatology.
- Univariate predictability $\sim \dfrac{\text{forecast uncertainty}}{\text{climatological uncertainty}}$.
  - relative error not absolute error
- Should be independent of units or variable weighting.
- Pred(A,B) = Pred(A)+Pred(B), if A and B are independent.
- Pred(A,B) = Pred(B) if A=f(B).

How to measure predictability

Predictability is . . .

- Relative to a baseline
  - usually climatology (historical frequencies).
- Difference between forecast and climatology.
- Univariate predictability $\sim \dfrac{\text{forecast uncertainty}}{\text{climatological uncertainty}}$.
  - relative error not absolute error
- Should be independent of units or variable weighting.
- Pred(A,B) = Pred(A)+Pred(B), if A and B are independent.
- Pred(A,B) = Pred(B) if A=f(B).

How to measure predictability

Predictability is ...

- Relative to a baseline
  - usually climatology (historical frequencies).
- Difference between forecast and climatology.
- Univariate predictability $\sim \dfrac{\text{forecast uncertainty}}{\text{climatological uncertainty}}$.
  - relative error not absolute error
- Should be independent of units or variable weighting.
- Pred(A,B) = Pred(A)+Pred(B), if A and B are independent.
- Pred(A,B) = Pred(B) if A=f(B).

How to measure predictability

Predictability is . . .

- Relative to a baseline
  - usually climatology (historical frequencies).
- Difference between forecast and climatology.
- Univariate predictability $\sim \dfrac{\text{forecast uncertainty}}{\text{climatological uncertainty}}$.
  - relative error not absolute error
- Should be independent of units or variable weighting.
- Pred(A,B) = Pred(A)+Pred(B), if A and B are independent.
- Pred(A,B) = Pred(B) if A=f(B).

Measures of predictability

Information theory gives some candidates:

- Mutual information (Leung and North, 1990)
- Predictive information (Schneider and Griffies, 1999)
- Relative entropy (Kleeman, 2002)

On average (over forecasts), they are the same.

Normally distributed variables

Mutual information M

$$M = -\log \frac{\det C_f}{\det C_c}$$

$C_f = \mathrm{Cov}(f)$
$C_c = \mathrm{Cov}(\mathrm{climatology})$

Univariate case:

$$M = \log \frac{\sigma_f^2}{\sigma_c^2} = \log \frac{\text{forecast variance}}{\text{climatological variance}}$$

Predictability is lost when the forecast is no better than a random pick from climatology.

Normally distributed variables

Mutual information M

$$M = -\log \frac{\det C_f}{\det C_c}$$

$C_f = \mathrm{Cov}(f)$
$C_c = \mathrm{Cov}(\text{climatology})$
Univariate case:

$$M = \log \frac{\sigma_f^2}{\sigma_c^2} = \log \frac{\text{forecast variance}}{\text{climatological variance}}$$

Predictability is lost when the forecast is no better than a random pick from climatology.

Normally distributed variables

Mutual information M

$$M = -\log \frac{\det C_f}{\det C_c}$$

$C_f = \mathrm{Cov}(f)$
$C_c = \mathrm{Cov}(\text{climatology})$
Univariate case:

$$M = \log \frac{\sigma_f^2}{\sigma_c^2} = \log \frac{\text{forecast variance}}{\text{climatological variance}}$$

Predictability is lost when the forecast is no better than a random pick from climatology.

Normalized forecast covariance

Multivariate generalization of $\dfrac{\text{forecast variance}}{\text{climatological variance}}$

$$G = C_c^{-1/2} C_f C_c^{-1/2}$$

$$\det G = \det(C_c^{-1/2} C_f C_c^{-1/2}) = \det(C_f C_c^{-1}) = \frac{\det C_f}{\det C_c}$$

$$M = -\log \frac{\det C_f}{\det C_c} = -\log \det G$$

Predictability is measured by the eigenvalues of G.
Eigenvectors of G decompose space according to predictability.

maximize		eigenvectors
variance	principal components (EOFs)	C_c
predictability	predictable components	G

Normalized forecast covariance

Multivariate generalization of $\dfrac{\text{forecast variance}}{\text{climatological variance}}$

$$G = C_c^{-1/2} C_f C_c^{-1/2}$$

$$\det G = \det(C_c^{-1/2} C_f C_c^{-1/2}) = \det(C_f C_c^{-1}) = \frac{\det C_f}{\det C_c}$$

$$M = -\log \frac{\det C_f}{\det C_c} = -\log \det G$$

Predictability is measured by the eigenvalues of G.
Eigenvectors of G decompose space according to predictability.

maximize		eigenvectors
variance	principal components (EOFs)	C_c
predictability	predictable components	G

Normalized forecast covariance

Multivariate generalization of $\dfrac{\text{forecast variance}}{\text{climatological variance}}$

$$G = C_c^{-1/2} C_f C_c^{-1/2}$$

$$\det G = \det(C_c^{-1/2} C_f C_c^{-1/2}) = \det(C_f C_c^{-1}) = \frac{\det C_f}{\det C_c}$$

$$M = -\log \frac{\det C_f}{\det C_c} = -\log \det G$$

Predictability is measured by the eigenvalues of G.

Eigenvectors of G decompose space according to predictability.

maximize		eigenvectors
variance	principal components (EOFs)	C_c
predictability	predictable components	G

Normalized forecast covariance

Multivariate generalization of $\dfrac{\text{forecast variance}}{\text{climatological variance}}$

$$G = C_c^{-1/2} C_f C_c^{-1/2}$$

$$\det G = \det(C_c^{-1/2} C_f C_c^{-1/2}) = \det(C_f C_c^{-1}) = \frac{\det C_f}{\det C_c}$$

$$M = -\log \frac{\det C_f}{\det C_c} = -\log \det G$$

Predictability is measured by the eigenvalues of G.
Eigenvectors of G decompose space according to predictability.

maximize		eigenvectors
variance	principal components (EOFs)	C_c
predictability	predictable components	G

Normalized forecast covariance

Multivariate generalization of $\dfrac{\text{forecast variance}}{\text{climatological variance}}$

$$G = C_c^{-1/2} C_f C_c^{-1/2}$$

$$\det G = \det(C_c^{-1/2} C_f C_c^{-1/2}) = \det(C_f C_c^{-1}) = \frac{\det C_f}{\det C_c}$$

$$M = -\log \frac{\det C_f}{\det C_c} = -\log \det G$$

Predictability is measured by the eigenvalues of G.
Eigenvectors of G decompose space according to predictability.

maximize		eigenvectors
variance	principal components (EOFs)	C_c
predictability	predictable components	G

Linear prediction model

$$f = Ai + \epsilon$$

$$i = a + \delta$$

f = forecast.
i = initial condition (analysis), $\langle i \rangle = a$
δ = initial condition error
ϵ = model error.

How does predictability (G) depend on

- dynamics A?
- initial condition errors δ?
- model errors ϵ?

Linear prediction model

$$f = Ai + \epsilon$$

$$i = a + \delta$$

f = forecast.
i = initial condition (analysis), $\langle i \rangle = a$
δ = initial condition error
ϵ = model error.

How does predictability (G) depend on

- dynamics A?
- initial condition errors δ?
- model errors ϵ?

Linear prediction model

$$f = Ai + \epsilon$$

$$i = a + \delta$$

f = forecast.
i = initial condition (analysis), $\langle i \rangle = a$
δ = initial condition error
ϵ = model error.

How does predictability (G) depend on

- dynamics A?
- initial condition errors δ?
- model errors ϵ?

Linear prediction model

$$f = Ai + \epsilon$$

$$i = a + \delta$$

f = forecast.
i = initial condition (analysis), $\langle i \rangle = a$
δ = initial condition error
ϵ = model error.

How does predictability (G) depend on

- dynamics A?
- initial condition errors δ?
- model errors ϵ?

Linear prediction model

$$f = Ai + \epsilon$$

$$i = a + \delta$$

f = forecast.
i = initial condition (analysis), $\langle i \rangle = a$
δ = initial condition error
ϵ = model error.

How does predictability (G) depend on

- dynamics A?
- initial condition errors δ?
- model errors ϵ?

Linear prediction model

Error dynamics approach

$$f = A(a + \delta) + \epsilon$$

$$\begin{array}{ccccc} C_f & = & A^T C_\delta A & + & C_\epsilon \\ \text{forecast uncertainty} & & \text{uncertainties from IC error} & & \text{model error} \end{array}$$

"Whiten"

$$\begin{array}{ccccc} C_c^{-1/2} C_f C_c^{-1/2} & = & C_c^{-1/2} A C_\delta A^T C_c^{-1/2} & + & C_c^{-1/2} C_\epsilon C_c^{-1/2} \\ G & = & A_{\text{error}} A_{\text{error}}^T & + & C_c^{-1/2} C_\epsilon C_c^{-1/2} \end{array}$$

$A_{\text{error}} \equiv C_c^{-1/2} A C_\delta^{1/2} \leftarrow$ initial and final norms

If there is no stochastic forcing $\epsilon = 0$,

$$G = A_{\text{error}} A_{\text{error}}^T$$

Predictability given by SVD of A_{error}.

Linear prediction model

Error dynamics approach

$$f = A(a + \delta) + \epsilon$$

$$\underset{\text{forecast uncertainty}}{C_f} = \underset{\text{uncertainties from IC error}}{A^T C_\delta A} + \underset{\text{model error}}{C_\epsilon}$$

"Whiten"

$$\begin{aligned} C_c^{-1/2} C_f C_c^{-1/2} &= C_c^{-1/2} A C_\delta A^T C_c^{-1/2} + C_c^{-1/2} C_\epsilon C_c^{-1/2} \\ G &= A_{\text{error}} A_{\text{error}}^T + C_c^{-1/2} C_\epsilon C_c^{-1/2} \end{aligned}$$

$A_{\text{error}} \equiv C_c^{-1/2} A C_\delta^{1/2} \leftarrow$ initial and final norms

If there is no stochastic forcing $\epsilon = 0$,

$$G = A_{\text{error}} A_{\text{error}}^T$$

Predictability given by SVD of A_{error}.

Linear prediction model

Error dynamics approach

$$f = A(a + \delta) + \epsilon$$

$$\begin{array}{ccccc} C_f & = & A^T C_\delta A & + & C_\epsilon \\ \text{forecast uncertainty} & & \text{uncertainties from IC error} & & \text{model error} \end{array}$$

"Whiten"

$$\begin{array}{ccccc} C_c^{-1/2} C_f C_c^{-1/2} & = & C_c^{-1/2} A C_\delta A^T C_c^{-1/2} & + & C_c^{-1/2} C_\epsilon C_c^{-1/2} \\ G & = & A_{\text{error}} A_{\text{error}}^T & + & C_c^{-1/2} C_\epsilon C_c^{-1/2} \end{array}$$

$A_{\text{error}} \equiv C_c^{-1/2} A C_\delta^{1/2} \leftarrow$ initial and final norms

If there is no stochastic forcing $\epsilon = 0$,

$$G = A_{\text{error}} A_{\text{error}}^T$$

Predictability given by SVD of A_{error}.

Linear prediction model

Signal approach

Averaging over forecasts

$$f = A(a + \delta) + \epsilon$$

$$\begin{array}{ccccccc} \text{climatology} & & \text{signal} & & \text{uncertainties from IC error} & & \text{model error} \\ C_c & = & A C_a A^T & + & A C_\delta A^T & + & C_\epsilon \\ & = & & & C_f & & \end{array}$$

Whiten

$$\begin{array}{ccccc} C_c^{-1/2} C_c C_c^{-1/2} & = & C_c^{-1/2} A C_a A^T C_c^{-1/2} & + & C_c^{-1/2} C_f C_c^{-1/2} \\ I & = & A_{\text{signal}} A_{\text{signal}}^T & + & G \end{array}$$

$A_{\text{signal}} \equiv C_c^{-1/2} A C_a^{1/2} \leftarrow$ initial and final norms

$$G = I - A_{\text{signal}} A_{\text{signal}}^T$$

$$\lambda(G) = 1 - \sigma^2(A_{\text{signal}})$$

Predictability determined by SVD of A_{signal}.

If $C_a = C_c$, minimum predictability when A_{signal} is normal.

Linear prediction model

Signal approach

Averaging over forecasts

$$f = A(a + \delta) + \epsilon$$

$$\begin{array}{ccccccc}
\text{climatology} & & \text{signal} & & \text{uncertainties from IC error} & & \text{model error} \\
C_c & = & AC_aA^T & + & AC_\delta A^T & + & C_\epsilon \\
 & = & & & C_f & &
\end{array}$$

Whiten

$$\begin{array}{ccccc}
C_c^{-1/2} C_c C_c^{-1/2} & = & C_c^{-1/2} A C_a A^T C_c^{-1/2} & + & C_c^{-1/2} C_f C_c^{-1/2} \\
I & = & A_{\text{signal}} A_{\text{signal}}^T & + & G
\end{array}$$

$A_{\text{signal}} \equiv C_c^{-1/2} A C_a^{1/2} \leftarrow$ initial and final norms

$$G = I - A_{\text{signal}} A_{\text{signal}}^T$$

$$\lambda(G) = 1 - \sigma^2(A_{\text{signal}})$$

Predictability determined by SVD of A_{signal}.

If $C_a = C_c$, minimum predictability when A_{signal} is normal.

Linear prediction model
Signal approach

Averaging over forecasts

$$f = A(a + \delta) + \epsilon$$

climatology		signal		uncertainties from IC error		model error
C_c	$=$	AC_aA^T	$+$	$AC_\delta A^T$	$+$	C_ϵ
	$=$			C_f		

Whiten

$$C_c^{-1/2} C_c C_c^{-1/2} = C_c^{-1/2} A C_a A^T C_c^{-1/2} + C_c^{-1/2} C_f C_c^{-1/2}$$

$$I = A_{\text{signal}} A_{\text{signal}}^T + G$$

$A_{\text{signal}} \equiv C_c^{-1/2} A C_a^{1/2} \leftarrow$ initial and final norms

$$G = I - A_{\text{signal}} A_{\text{signal}}^T$$

$$\lambda(G) = 1 - \sigma^2(A_{\text{signal}})$$

Predictability determined by SVD of A_{signal}.

If $C_a = C_c$, minimum predictability when A_{signal} is normal.

Linear prediction model

Signal approach

Averaging over forecasts

$$f = A(a + \delta) + \epsilon$$

$$\begin{array}{ccccccc} \text{climatology} & & \text{signal} & & \text{uncertainties from IC error} & & \text{model error} \\ C_c & = & A C_a A^T & + & A C_\delta A^T & + & C_\epsilon \\ & = & & & C_f & & \end{array}$$

Whiten

$$\begin{array}{ccccc} C_c^{-1/2} C_c C_c^{-1/2} & = & C_c^{-1/2} A C_a A^T C_c^{-1/2} & + & C_c^{-1/2} C_f C_c^{-1/2} \\ I & = & A_{\text{signal}} A_{\text{signal}}^T & + & G \end{array}$$

$A_{\text{signal}} \equiv C_c^{-1/2} A C_a^{1/2} \leftarrow$ initial and final norms

$$G = I - A_{\text{signal}} A_{\text{signal}}^T$$

$$\lambda(G) = 1 - \sigma^2(A_{\text{signal}})$$

Predictability determined by SVD of A_{signal}.

If $C_a = C_c$, minimum predictability when A_{signal} is normal.

Linear prediction model

Signal approach

Averaging over forecasts

$$f = A(a + \delta) + \epsilon$$

$$\begin{array}{ccccccc} \text{climatology} & & \text{signal} & & \text{uncertainties from IC error} & & \text{model error} \\ C_c & = & AC_aA^T & + & AC_\delta A^T & + & C_\epsilon \\ & = & & & C_f & & \end{array}$$

Whiten

$$\begin{array}{ccccc} C_c^{-1/2} C_c C_c^{-1/2} & = & C_c^{-1/2} AC_aA^T C_c^{-1/2} & + & C_c^{-1/2} C_f C_c^{-1/2} \\ I & = & A_{\text{signal}} A_{\text{signal}}^T & + & G \end{array}$$

$A_{\text{signal}} \equiv C_c^{-1/2} A C_a^{1/2} \leftarrow$ initial and final norms

$$G = I - A_{\text{signal}} A_{\text{signal}}^T$$

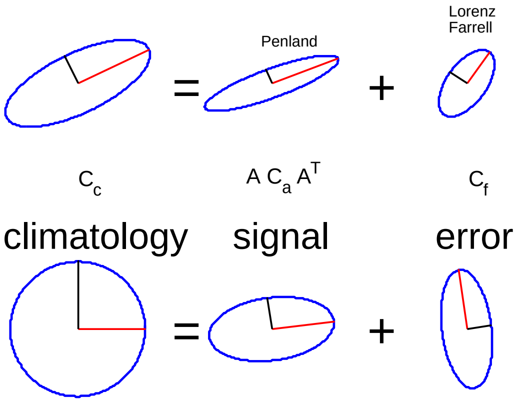
$$\lambda(G) = 1 - \sigma^2(A_{\text{signal}})$$

Predictability determined by SVD of A_{signal}.

If $C_a = C_c$, minimum predictability when A_{signal} is normal.

Whitening transformation

2D picture



Paradox?

No model error

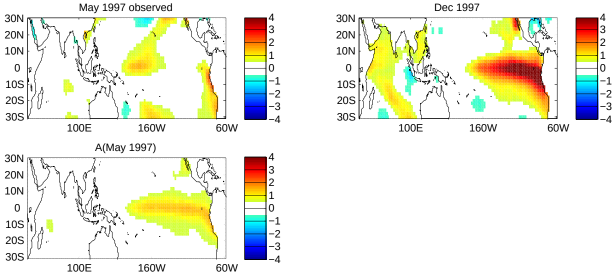
$$G = A_{\text{error}} A_{\text{error}}^T = I - A_{\text{signal}} A_{\text{signal}}^T$$

$$\sigma^2(A_{\text{error}}) = 1 - \sigma^2(A_{\text{signal}})$$

$(\text{rms}^2 = 1 - \text{corr}^2)$

Same singular vectors–opposite ordering.
Minimizing error = maximizing signal.

SST example



A maps

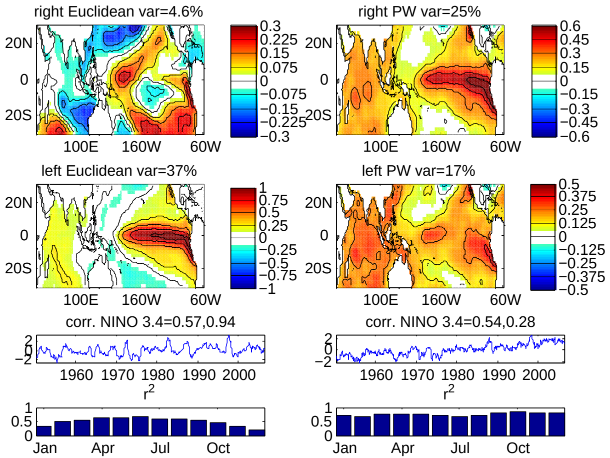
sea surface temperature at time t

to SST temperature at time $t + 7$ months.

A is an empirical model, fit to data.

What are the predictable components?

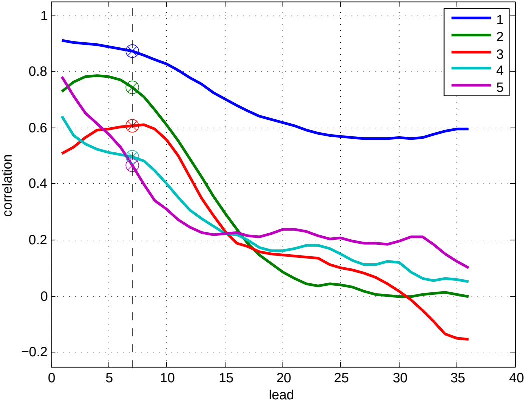
SST example



right Euclidean var=4.6%
right PW var=25%
left Euclidean var=37%
left PW var=17%
corr. NINO 3.4=0.57,0.94
corr. NINO 3.4=0.54,0.28
20N
0
20S
100E
160W
60W
0.3
0.225
0.15
0.075
0
−0.075
−0.15
−0.225
−0.3
0.6
0.45
0.3
0.15
0
−0.15
−0.3
−0.45
−0.6
1
0.75
0.5
0.25
0
−0.25
−0.5
−0.75
−1
0.5
0.375
0.25
0.125
0
−0.125
−0.25
−0.375
−0.5
1960 1970 1980 1990 2000
r2
Jan Apr Jul Oct

SST example

Predictability of 7-month components at other leads.



Summary

- SVD used in weather and climate applications to diagnose
  - error growth
  - signal amplification
- Seeming paradox. Error growth bad – signal growth good.
- Defining predictability , defines SVD initial and final norms
- Predictable component analysis
  - canonical correlation analysis
  - fingerprint methods
  - discriminant analysis
  - S/N EOFs

- Future work
  - Estimating distributions
    - Short record
    - Imperfect models

DelSole, T., and M. K. Tippett, 2007: Measuring Predictability with Singular Vectors. J. Atmos. Sci., submitted.

DelSole, T., and M. K. Tippett, 2007: Predictability: Recent Insights from Information Theory. Rev. Geophys.,

Summary

- SVD used in weather and climate applications to diagnose
  - error growth
  - signal amplification
- Seeming paradox. Error growth bad – signal growth good.
- Defining predictability , defines SVD initial and final norms
- Predictable component analysis
  - canonical correlation analysis
  - fingerprint methods
  - discriminant analysis
  - S/N EOFs
- Future work
  - Estimating distributions
    - Short record
    - Imperfect models

DelSole, T., and M. K. Tippett, 2007: Measuring Predictability with Singular Vectors. J. Atmos. Sci., submitted.

DelSole, T., and M. K. Tippett, 2007: Predictability: Recent Insights from Information Theory. Rev. Geophys.,