18-660: Numerical Methods for Engineering Design and Optimization

Xin Li
Department of ECE
Carnegie Mellon University
Pittsburgh, PA 15213
Overview

- Principal Component Analysis (PCA)
  - Correlation decomposition
  - Dimension reduction
Monte Carlo Analysis

- Monte Carlo analysis for $f(X)$
  - Randomly select $M$ samples for $X$
  - Evaluate function $f(X)$ at each sampling point
  - Estimate distribution of $f$ using these $M$ samples

We assume that random samples can be easily created from a random number generator
A random number generator creates a pseudo-random sequence for which the period is extremely large:
- MATLAB function “randn(•)”: period is $\sim 2^{64}$
- MATLAB function “rand(•)”: period is $\sim 2^{1492}$

All samples in $\{x^{(1)}, x^{(2)}, \ldots\}$ are “almost” independent.
Monte Carlo Analysis

Example: sample independent random variables $x$ and $y$

- Generate random sequence $\{x^{(1)},y^{(1)},x^{(2)},y^{(2)},\ldots\}$
- Create sampling pair $\{(x^{(1)},y^{(1)}),(x^{(2)},y^{(2)}),\ldots\}$
- $x^{(i)}$ and $y^{(i)}$ in each pair are independent

However, how can we sample correlated random variables?
Monte Carlo Analysis

- Correlated random variables cannot be directly sampled by a random number generator

- We can decompose correlated random variables to a set of independent variables, if they are jointly Normal
  - Focus of this lecture

- Other techniques also exist to sample correlated variables
  - Details can be found in many text books on Monte Carlo analysis

Fishman, A First Course In Monte Carlo, 2006
Correlation Decomposition

- Key idea: given the correlated random variables \( \{x_1, x_2, \ldots\} \), find a linear transform \( Y = P \cdot X \) such that \( \{y_1, y_2, \ldots\} \) are independent.
  - Only applicable to **jointly Normal** random variables for which \( \{y_1, y_2, \ldots\} \) just need to be uncorrelated.
  - Otherwise, if the random variables are not jointly Normal, such a linear transform may not exist.

\[
\begin{bmatrix}
  y_1 \\
  y_2 \\
  y_3 \\
\end{bmatrix} = \begin{bmatrix}
  \phantom{x} & \phantom{x} & \phantom{x} \\
  \phantom{x} & \phantom{x} & \phantom{x} \\
  \phantom{x} & \phantom{x} & \phantom{x}
\end{bmatrix} \cdot \begin{bmatrix}
  x_1 \\
  x_2 \\
  x_3 \\
\end{bmatrix}
\]

\( P \)
Principal Component Analysis (PCA)

- Given a set of jointly Normal random variables

\[
X = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^T
\]

- Assume that all \(x_i\)'s have zero mean

- Covariance matrix is

\[
E[X \cdot X^T] = E\begin{bmatrix}
x_1^2 & x_1x_2 & x_1x_3 \\
x_1x_2 & x_2^2 & x_2x_3 \\
x_1x_3 & x_2x_3 & x_3^2
\end{bmatrix}
\]

The covariance matrix has many important properties, e.g., it is symmetric
Principal Component Analysis (PCA)

- Covariance matrix is positive semi-definite

- A symmetric matrix $A$ is called positive semi-definite if

$$Q^T AQ \geq 0$$

for any real-valued vector $Q$

Why is a covariance matrix positive semi-definite?
Principal Component Analysis (PCA)

- Assume that $X = [x_1 \ x_2 \ ... \ x_N]^T$ are $N$ random variables with zero mean.

$$y = Q^T X$$

$y$ is a scalar and $Q^T X$ is any real-value vector.

$$E[y^2] = E[(Q^T X) \cdot (X^T Q)] = Q^T \cdot E[XX^T] \cdot Q \geq 0$$

$y \quad y^T = y$
Principal Component Analysis (PCA)

To remove correlation, we decompose the covariance matrix by eigenvalues & eigenvectors

\[ A = E[X \cdot X^T] \quad AV_i = V_i \cdot \lambda_i \]

Normalized eigenvectors: \( ||V_i||_2 = 1 \)

\[ A \cdot V = V \cdot \Sigma \]

\[ V = [V_1 \quad V_2 \quad \ldots] \]

\[ \Sigma = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{bmatrix} \]

Eigenvalues
The eigen-decomposition of a covariance matrix $A$ has a number of important properties:

- $A$ is symmetric $\rightarrow$ all eigenvalues are real
- $A$ is symmetric $\rightarrow$ all eigenvectors are real and orthogonal

$$A \cdot V = V \cdot \Sigma$$

$$V^T V = I$$

Identity matrix
Principal Component Analysis (PCA)

- The eigen-decomposition of a covariance matrix $A$ has a number of important properties
  - $A$ is positive semi-definite $\iff$ all eigenvalues are non-negative

\[
A \cdot V = V \cdot \Sigma \quad V^T V = I
\]

\[
A \cdot V \cdot V^{-1} = V \cdot \Sigma \cdot V^{-1}
\]

\[
A = V \cdot \Sigma \cdot V^{-1}
\]

\[
A = V \cdot \Sigma \cdot V^T
\]

$\Sigma = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{bmatrix}$

Eigenvalues
Principal Component Analysis (PCA)

- Define new random variables Y (principal components)
  \[ Y = \Sigma^{-0.5} \cdot V^T \cdot X \]
  \[ X = V \cdot \Sigma^{0.5} \cdot Y \]

- All principal components (also called principal factors) are jointly Normal
  - They are linear combination of jointly Normal random variables

- We will theoretically prove that all principal components are independent and standard Normal
Principal Component Analysis (PCA)

- All principal components have zero mean

\[ Y = \Sigma^{-0.5} \cdot V^T \cdot X \]

\[ E[Y] = \Sigma^{-0.5} \cdot V^T \cdot E[X] \]

\[ E[X] = 0 \quad \text{All random variables in X have zero mean} \]

\[ E[Y] = 0 \]
Principal Component Analysis (PCA)

- All principal components are independent and standard Normal

\[ Y = \Sigma^{-0.5} \cdot V^T \cdot X \]

\[ E[Y \cdot Y^T] = E[\Sigma^{-0.5} \cdot V^T \cdot X \cdot X^T \cdot V \cdot \Sigma^{-0.5}] \]

\[ E[Y \cdot Y^T] = \Sigma^{-0.5} \cdot V^T \cdot E[X \cdot X^T] \cdot V \cdot \Sigma^{-0.5} \]
Principal Component Analysis (PCA)

\[ E[Y \cdot Y^T] = \Sigma^{-0.5} \cdot V^T \cdot E[X \cdot X^T] \cdot V \cdot \Sigma^{-0.5} \]

\[ E[X \cdot X^T] = V \cdot \Sigma \cdot V^T \]

\[ E[Y \cdot Y^T] = \Sigma^{-0.5} \cdot V^T \cdot V \cdot \Sigma \cdot V^T \cdot V \cdot \Sigma^{-0.5} \]

\[ V^TV = I \]

\[ E[Y \cdot Y^T] = \Sigma^{-0.5} \cdot \Sigma \cdot \Sigma^{-0.5} = I \quad \text{Unit variance and uncorrelated} \]

“Uncorrelated” = “independent” for jointly Normal random variables
Principal Component Analysis (PCA)

- Example: $x_1$ and $x_2$ are zero mean and jointly Normal

\[
E[X \cdot X^T] = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}
\]

\[
\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}
\]

Eigen decomposition
Example (continued):

\[ \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \]

\[ Y = \Sigma^{-0.5} \cdot V^T \cdot X = \begin{bmatrix} 1 & 0 \\ 0 & 1/3 \end{bmatrix} \cdot \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ \sqrt{2} & \sqrt{2} \end{bmatrix} \cdot X \]

\[ Y = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \\ 3\sqrt{2} & 3\sqrt{2} \end{bmatrix} \cdot X \]
Example (continued):

\[
Y = \begin{bmatrix}
-\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\
\frac{1}{3\sqrt{2}} & \frac{1}{3\sqrt{2}}
\end{bmatrix} \cdot X
\]

\[
E[Y \cdot Y^T] = \begin{bmatrix}
-\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\
\frac{1}{3\sqrt{2}} & \frac{1}{3\sqrt{2}}
\end{bmatrix} \cdot E[X \cdot X^T] \cdot \begin{bmatrix}
-\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\
\frac{1}{3\sqrt{2}} & \frac{1}{3\sqrt{2}}
\end{bmatrix}^T
\]
Principal Component Analysis (PCA)

- Example (continued):

\[
E[Y \cdot Y^T] = \begin{bmatrix}
-\frac{1}{\sqrt{2}} & 1 \\
1 & \frac{1}{\sqrt{2}} \\
\frac{1}{3\sqrt{2}} & \frac{1}{3\sqrt{2}}
\end{bmatrix} \cdot E[X \cdot X^T] \cdot \begin{bmatrix}
-\frac{1}{\sqrt{2}} & 1 \\
1 & \frac{1}{\sqrt{2}} \\
\frac{1}{3\sqrt{2}} & \frac{1}{3\sqrt{2}}
\end{bmatrix}^T
E[X \cdot X^T] = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}
\]

\[
E[Y \cdot Y^T] = \begin{bmatrix}
-\frac{1}{\sqrt{2}} & 1 \\
1 & \frac{1}{\sqrt{2}} \\
\frac{1}{3\sqrt{2}} & \frac{1}{3\sqrt{2}}
\end{bmatrix} \cdot \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \cdot \begin{bmatrix}
-\frac{1}{\sqrt{2}} & 1 \\
1 & \frac{1}{\sqrt{2}} \\
\frac{1}{3\sqrt{2}} & \frac{1}{3\sqrt{2}}
\end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\]

All principal components in Y are independent and standard Normal
The decomposition for independence is not unique

Define

\[ Z = U \cdot Y \]

U is an orthogonal matrix, i.e., \( U^T U = I \)

\[
E[Z \cdot Z^T] = E[U \cdot Y \cdot Y^T \cdot U^T] = U \cdot E[Y \cdot Y^T] \cdot U^T = U \cdot U^T = I
\]

All random variables in Z are also independent and standard Normal
Example: $x_1, x_2$ and $x_3$ are zero mean and jointly Normal

$$E[X \cdot X^T] = \begin{bmatrix} 5 & 4 & 3 \\ 4 & 5 & 3 \\ 3 & 3 & 2 \end{bmatrix}$$

Eigen decomposition

$$\Sigma = \begin{bmatrix} 11 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} 0.6396 & 0.7071 & 0.3015 \\ 0.6396 & -0.7071 & 0.3015 \\ 0.4264 & 0 & -0.9045 \end{bmatrix}$$

One of the eigenvalues is 0
Example (continued):

- In this case, the 3x3 covariance matrix has a rank of 2
- Only 2 independent principal components (Y) are required to **EXACTLY** represent the 3-dimensional random space

\[
X = V \cdot \Sigma^{0.5} \cdot Y = \begin{bmatrix}
0.6396 & 0.7071 & 0.3015 \\
0.6396 & -0.7071 & 0.3015 \\
0.4264 & 0 & -0.9045
\end{bmatrix} \begin{bmatrix}
\sqrt{11} & 0 & 0 \\
0 & \sqrt{1} & 0 \\
0 & 0 & 0
\end{bmatrix} \cdot Y
\]

Only \(y_1\) and \(y_2\) are required

\[
X = \begin{bmatrix}
2.1213 & 0.7071 & 0 \\
2.1213 & -0.7071 & 0 \\
1.4142 & 0 & 0
\end{bmatrix} \begin{bmatrix}
y_1 \\
y_2 \\
y_3
\end{bmatrix}
\]

\(y_3\) does not affect \(X\)
Dimension Reduction by PCA

- In general, if some of the eigenvalues are small, they can be ignored to reduce the random space dimension
  - Allows us to use a compact set of independent principal components to approximate the original high-dimensional space
  - E.g., only two random variables $y_1$ and $y_2$ are required to represent the variations of $x_1$, $x_2$ and $x_3$ in the previous example

- PCA is useful to reduce problem size in many applications
  - But applicable to jointly Normal variables only
Summary

- Principal component analysis (PCA)
  - Correlation decomposition
  - Dimension reduction