Lecture 12
Scientific and Data Computing I

Mike Kirby
School of Computing
Lecture

- Wednesday – Principle Component Analysis (PCA)
- **No Office Hours This Week Due to Admin Meeting**
- Monday – Independent Component Analysis (ICA)
- Wednesday – Practicum
Principle Component Analysis (PCA)

Given data $Q$, scale it:

$$\bar{q}_j = \frac{1}{n} \sum_{i=1}^{n} q_{ij}$$

$$p_{ij} = \frac{Q_{ij}}{S_j}$$
Principle Component Analysis (PCA)

• Linear Approximation

\[ p = (p_1, p_2, \ldots, p_m)^T \]

\[ p = \alpha_1 v_1, \]

\[ p = \alpha_1 v_1 + \alpha_2 v_2, \]

\[ p = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3, \]

\[ \vdots \]

\[ p = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_{m-1} v_{m-1}. \]
Principle Component Analysis (PCA)

- Error Function

\[ d_i = \sqrt{p_i \cdot p_i - (p_i \cdot v_1)^2} \]

\[ E_1 = \sum_{i=1}^{n} d_i^2 \]

\[ = \sum_{i=1}^{n} \left[ p_i \cdot p_i - (p_i \cdot v_1)^2 \right] \]
Principle Component Analysis (PCA)

- Error Function (continued)

\[ \bar{d}_i = \| \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 - \mathbf{p}_i \|_2 \]

\[ E_2 = \sum_{i=1}^{n} \left[ \mathbf{p}_i \cdot \mathbf{p}_i - (\mathbf{p}_i \cdot \mathbf{v}_1)^2 - (\mathbf{p}_i \cdot \mathbf{v}_2)^2 \right] \]
Principle Component Analysis (PCA)

- Minimization

\[ F(v_1, \lambda_1) = E_1(v_1) + \lambda_1(\|v_1\|_2^2 - 1) \]

- 2D example

\[ F(v_1, v_2, \lambda_1, \lambda_2) = E_2(v_1, v_2) + \lambda_1(\|v_1\|_2^2 - 1) + \lambda_2(\|v_2\|_2^2 - 1), \]

where \( \lambda_1 \) and \( \lambda_2 \) are the Lagrange multipliers.

The \( v_j \)'s must satisfy \( (P^TP)v_j = \lambda_j v_j \).
Principle Component Analysis (PCA)

- Minimization (continued)

\[ \sum_{i=1}^{n} (p_i \cdot v_j)^2 = \lambda_j \]

\[ E_1 = -\lambda_1 + \sum_{i=1}^{n} (p_i \cdot p_i) \]

\[ E_2 = -\lambda_1 - \lambda_2 + \sum_{i=1}^{n} (p_i \cdot p_i) \]
Principle Component Analysis (PCA)

The SVD has the form $\mathbf{P} = \mathbf{U}\Sigma\mathbf{V}^T$, where $\Sigma$ is a diagonal-like matrix containing the singular values $\sigma_i$, and $\mathbf{V}$ is an orthogonal $m \times m$ matrix. Also, the eigenvalues $\lambda_i$ of $\mathbf{P}^T\mathbf{P}$ and the singular values $\sigma_i$ for $\mathbf{P}$ are connected through the formula $\lambda_i = \sigma_i^2$. Second, the columns of the matrix $\mathbf{V}$ are the eigenvectors for $\mathbf{P}^T\mathbf{P}$. Third, the SVD orders the singular values by size so, $\sigma_1 \geq \sigma_2 \geq \ldots \geq 0$. Based on this, the smallest value of $E_1$ is obtained when $\mathbf{v}_1$ is taken to be the first column of $\mathbf{V}$. 
Principle Component Analysis (PCA)

**Theorem 9.1.** Let \( P \) be a normalized nonzero \( n \times m \) data matrix, with \( m < n \). Also, let the SVD of this matrix be \( P = U\Sigma V^T \). Given \( k \), with \( 1 \leq k < m \), consider a linear fit of the data of the form.

\[
p = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k
\]

The smallest error is obtained when the vectors \( v_1, v_2, \ldots, v_k \) are chosen to be the first \( k \) columns of \( V \).
Principle Component Analysis (PCA)

• Scaling Factors

\[ S_j = \frac{1}{n} \| c_j \|_1, \quad S_j = \frac{1}{\sqrt{n}} \| c_j \|_2, \]

\[ S_j = \| c_j \|_\infty \]

• Autoscaling

\[ S_j = \frac{1}{\sqrt{n - 1}} \| c_j \|_2 \]
Principle Component Analysis (PCA)

• Error

\[ E_k = \sum_{i=1}^{n} p_i \cdot p_i - (\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_k^2) \]

\[ \sigma_{i=1}^{n} p_i \cdot p_i = \sigma_{j=1}^{m} \sigma_j^2 \]
Principle Component Analysis (PCA)

• Error (continued)

\[ \|P\|_F \equiv \sqrt{\sum_{i,j} p_{ij}^2} = \sqrt{\sum_{i=1}^{n} p_i \cdot p_i} \]

\[ R(k) = \frac{\sigma_{k+1}^2 + \sigma_{k+2}^2 + \ldots + \sigma_m^2}{\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_m^2} \]
For Next Time

• Read Chapter 4 and 9 (for volatility)
• Start working on Practicums
• Start working on Homework 3 (due on the following Thursday)