Study Guide: Principal Component Analysis and Compressed Sensing

Student Name

October 9, 2024

1 Principal Component Analysis (Section 23.1)

Principal Component Analysis (PCA) is a dimensionality reduction technique that transforms the data into a new coordinate system, where the axes correspond to the directions of maximum variance. PCA is widely used in machine learning for feature extraction, data visualization, and noise reduction.

1.1 Problem Setup

Given a dataset $\{x_1, \ldots, x_n\} \subseteq \mathbb{R}^d$, the goal of PCA is to project the data onto a lower-dimensional space while retaining as much of the variance as possible.

Let $X \in \mathbb{R}^{n \times d}$ be the data matrix, where each row represents an example. The covariance matrix of the data is given by:

$$\Sigma = \frac{1}{n} X^\top X.$$

PCA seeks to find the top k eigenvectors of Σ, which correspond to the directions of maximum variance in the data.

1.2 PCA Algorithm

The steps of the PCA algorithm are as follows:

1. Center the data by subtracting the mean of each feature from the dataset.
2. Compute the covariance matrix Σ.
3. Perform eigendecomposition of Σ to obtain the eigenvectors and eigenvalues.
4. Select the top k eigenvectors corresponding to the largest eigenvalues.
5. Project the data onto the subspace spanned by the top k eigenvectors.

The principal components are the directions of maximum variance, and projecting the data onto these components captures most of the variability in a lower-dimensional space.

1.3 Reconstruction Error

The reconstruction error measures how well the original data is approximated by the projection onto the lower-dimensional space. The total reconstruction error is minimized when the data is projected onto the top k principal components. The error can be computed as the sum of the discarded eigenvalues.

1.4 Applications of PCA

PCA is commonly used for:

- **Data compression**: Reducing the dimensionality of large datasets while preserving most of the relevant information.
- **Noise reduction**: Filtering out noise by keeping only the components that capture the most variance.
- **Visualization**: Projecting high-dimensional data onto 2D or 3D spaces for easy visualization.

2 PCA or Compressed Sensing? (Section 23.4)

Both **Principal Component Analysis (PCA)** and **Compressed Sensing** are techniques used for dimensionality reduction, but they are suitable for different scenarios and have distinct theoretical foundations.

2.1 When to Use PCA?

PCA is optimal when the data lies close to a low-dimensional linear subspace of the original high-dimensional space. It works well in situations where the principal components capture the most significant variance in the data, and noise or less important features are associated with the smaller eigenvalues.

2.2 When to Use Compressed Sensing?

Compressed sensing, on the other hand, is ideal for situations where the data is sparse in some basis. It leverages the fact that the signal of interest can be reconstructed from a small number of linear measurements if it is sparse in a known domain. This is particularly useful in signal processing and image compression, where data often has a sparse representation in a specific basis (e.g., wavelets or Fourier transforms).

2.3 Comparison

- **PCA**: Best suited for data that lies on a low-dimensional linear subspace. The reconstruction quality depends on the variance captured by the principal components.
- **Compressed Sensing**: Effective when the data is sparse in some basis and can be reconstructed from fewer measurements than typically required by Nyquist sampling theory.

2.4 Trade-offs

PCA is simpler and computationally less expensive than compressed sensing. However, compressed sensing can recover sparse signals with fewer measurements than PCA when the sparsity assumption holds. The choice between PCA and compressed sensing depends on the nature of the data and the assumptions that can be made about its structure.

3 Summary

- **PCA** is a powerful technique for dimensionality reduction, used to capture the directions of maximum variance in data. It is optimal when data lies near a low-dimensional linear subspace.
- **Compressed Sensing** is an alternative method for dimensionality reduction, suited for sparse data. It allows reconstruction from a small number of measurements when the data is sparse in a known basis.
- The choice between PCA and compressed sensing depends on the properties of the data, particularly whether it is low-dimensional in the original space or sparse in a known basis.