TYPES OF PCA

PCA: Dimensionality reduction and factor identification.

Kernel PCA: Capturing non-linear relationships.

Incremental PCA: Real-time data processing and model updating.

Sparse PCA: Feature selection and interoperability.

Robust PCA: Handling data with outliers.

ICA: Separating independent sources and noise reduction.

QUANT INSIDER

Principal Component Analysis (PCA)

Principal Component Analysis (PCA) transforms data into a new coordinate system such that the greatest variances by any projection of the data come to lie on the first coordinate (the first principal component), the second greatest variance on the second coordinate, and so on. Mathematically, PCA involves:

1. **Standardizing the data** if variables are measured on different scales.
2. **Computing the covariance matrix** $\mathbf{C} = \frac{1}{n-1}\mathbf{X}^T\mathbf{X}$, where $\mathbf{X}$ is the data matrix.
3. **Calculating eigenvalues and eigenvectors** of the covariance matrix.
4. **Sorting eigenvectors** by descending eigenvalues and forming a feature vector $\mathbf{P}$.
5. **Transforming the data** to the new basis: $\mathbf{X}' = \mathbf{X}\mathbf{P}$.

Incremental PCA

Incremental PCA (IPCA) processes data in mini-batches, updating the principal components incrementally. It is particularly suited for large datasets that cannot fit into memory.

1. **Initialize** mean vector $\mathbf{m}_0$ and covariance matrix $\mathbf{C}_0$.
2. For each mini-batch $\mathbf{X}_k$:
   - **Update the mean:** $\mathbf{m}_k = \mathbf{m}_{k-1} + \frac{1}{k}(\bar{\mathbf{X}}_k - \mathbf{m}_{k-1})$
   - **Update the covariance matrix:** $\mathbf{C}_k = \mathbf{C}_{k-1} + \frac{1}{k}(\mathbf{X}_k^T\mathbf{X}_k - \mathbf{C}_{k-1})$
3. **Compute eigenvalues and eigenvectors** of the final covariance matrix.

Independent Component Analysis (ICA)

Independent Component Analysis (ICA) decomposes a multivariate signal into additive, independent components. Given $\mathbf{X} = \mathbf{A}\mathbf{S}$, where $\mathbf{S}$ are the independent sources and $\mathbf{A}$ is the mixing matrix, ICA aims to find $\mathbf{W}$ such that:

$$\mathbf{S} = \mathbf{W}\mathbf{X}$$

where $\mathbf{W}$ is the unmixing matrix. This is often done by maximizing non-Gaussianity or mutual information.

Kernel PCA

Kernel PCA (KPCA) extends PCA to non-linear data by using kernel methods. KPCA first maps the data into a higher-dimensional space $\Phi : \mathbf{x} \mapsto \phi(\mathbf{x})$ and then applies PCA in this space. The kernel trick allows us to compute the dot products in ϕ-space without explicitly mapping $\mathbf{x}$ to $\phi(\mathbf{x})$.

1. **Compute the kernel matrix** $\mathbf{K}$ where $\mathbf{K}_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$.
2. **Center the kernel matrix:** $\mathbf{K} = \mathbf{K} - \mathbf{1}_n\mathbf{K} - \mathbf{K}\mathbf{1}_n + \mathbf{1}_n\mathbf{K}\mathbf{1}_n$, where $\mathbf{1}_n$ is a matrix of ones.
3. **Compute eigenvalues and eigenvectors** of the centered kernel matrix.
4. **Transform the data** using the top eigenvectors.

Robust PCA

Robust PCA (RPCA) separates data into a low-rank matrix $\mathbf{L}$ and a sparse matrix $\mathbf{S}$:

$$\min_{\mathbf{L},\mathbf{S}} \|\mathbf{L}\|_* + \lambda\|\mathbf{S}\|_1 \quad \text{subject to} \quad \mathbf{X} = \mathbf{L} + \mathbf{S}$$

where $\|\mathbf{L}\|_*$ is the nuclear norm (sum of singular values) of $\mathbf{L}$, and $\|\mathbf{S}\|_1$ is the L_1 norm of $\mathbf{S}$.

Sparse PCA

Sparse PCA introduces sparsity into the principal components by adding a sparsity-inducing penalty to the PCA optimization problem. It involves solving:

$$\min_{\mathbf{P},\mathbf{Z}} \|\mathbf{X} - \mathbf{XPZ}^T\|_F^2 + \alpha\|\mathbf{P}\|_1 \quad \text{subject to} \quad \mathbf{Z}^T\mathbf{Z} = \mathbf{I}$$

where $\|\mathbf{P}\|_1$ is the L_1 norm promoting sparsity.