

Neuromatch Academy: Dimensionality Reduction - Summary Sheet ¹

Dimensionality Reduction

Generate Correlated Multivariate Data

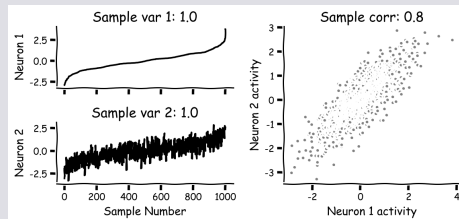
Multivariate data can be visualized as a cloud of points in a high-dimensional vector space. The geometry of this cloud is shaped by the covariance matrix.

The covariance can be found from the equation above:

$$\text{cov}(x_1, x_2) = \rho \sqrt{\sigma_1^2 \sigma_2^2}. \quad (1)$$

The covariance matrix for two dimensions has the following form:

$$\Sigma = \begin{pmatrix} \text{var}(x_1) & \text{cov}(x_1, x_2) \\ \text{cov}(x_1, x_2) & \text{var}(x_2) \end{pmatrix}. \quad (2)$$



Orthonormal Basis

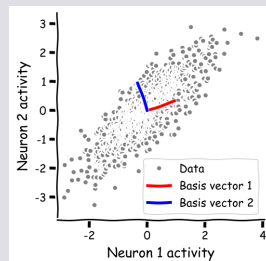
Two vectors are orthonormal if:

1. They are orthogonal (i.e., their dot product is zero):

$$\mathbf{u} \cdot \mathbf{w} = u_1 w_1 + u_2 w_2 = 0 \quad (3)$$

2. They have unit length:

$$\|\mathbf{u}\| = \|\mathbf{w}\| = 1 \quad (4)$$



Principal Component Analysis

Eigenvectors of the Sample Covariance Matrix

PCA represents data in a new orthonormal basis defined by the eigenvectors of the covariance matrix. The bivariate normal data with a specified covariance matrix Σ , whose (i, j) th element is:

$$\Sigma_{ij} = E[x_i x_j] - E[x_i]E[x_j]. \quad (5)$$

We use the sample covariance matrix, $\hat{\Sigma}$, which is calculated directly from the data. The (i, j) th element of the sample covariance matrix is:

$$\hat{\Sigma}_{ij} = \frac{1}{N_{\text{samples}}} \mathbf{x}_i^T \mathbf{x}_j - \bar{\mathbf{x}}_i \bar{\mathbf{x}}_j, \quad (6)$$

where $\mathbf{x}_i = [x_i(1), x_i(2), \dots, x_i(N_{\text{samples}})]^T$ is a column vector representing all measurements of neuron i , and $\bar{\mathbf{x}}_i$ is the mean of neuron i across samples:

$$\bar{\mathbf{x}}_i = \frac{1}{N_{\text{samples}}} \sum_{k=1}^{N_{\text{samples}}} x_i(k). \quad (7)$$

If we assume that the data has already been mean-subtracted, then we can write the sample covariance matrix in a much simpler matrix form:

$$\hat{\Sigma} = \frac{1}{N_{\text{samples}}} \mathbf{X}^T \mathbf{X}. \quad (8)$$

where \mathbf{X} is the full data matrix (each column of \mathbf{X} is a different \mathbf{x}_i).

Principal Component Analysis

PCA by projecting data onto the eigenvectors

To perform PCA project the data onto the eigenvectors of the covariance matrix, i.e.:

$$\mathbf{S} = \mathbf{X}\mathbf{W} \quad (9)$$

where \mathbf{S} is an $N_{\text{samples}} \times N$ matrix representing the projected data (also called *scores*), and \mathbf{W} is an $N \times N$ orthogonal matrix, each of whose columns represents the eigenvectors of the covariance matrix (also called *weights* or *loadings*). Since \mathbf{W} is an orthogonal matrix, $\mathbf{W}^{-1} = \mathbf{W}^T$. So by multiplying by \mathbf{W}^T on each side we can rewrite this equation as

$$\mathbf{X} = \mathbf{S}\mathbf{W}^T. \quad (10)$$

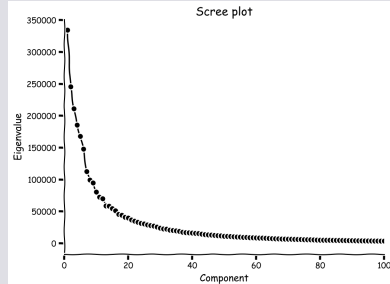
To reconstruct the data from a low-dimensional approximation, we just have to truncate these matrices. Let's denote $\mathbf{S}_{1:K}$ and $\mathbf{W}_{1:K}$ the matrices with only the first K columns of \mathbf{S} and \mathbf{W} , respectively. Then our reconstruction is:

$$\hat{\mathbf{X}} = \mathbf{S}_{1:K} (\mathbf{W}_{1:K})^T. \quad (11)$$

Dimensionality Reduction & Reconstruction

Scree Plot

Each eigenvalue describes the variance of the data projected onto its corresponding eigenvector. This is an important concept because it allows us to rank the PCA basis vectors based on how much variance each one can capture in a scree plot. A scree plot is used to help choose which and how many components to use:

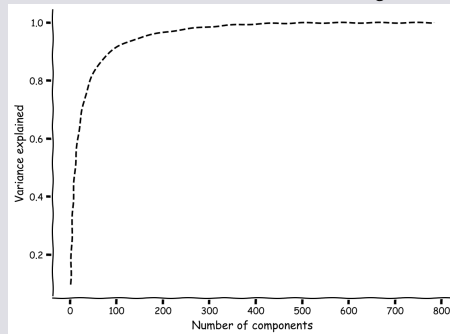


Calculate the variance explained

Another common way to determine the intrinsic dimensionality is by considering the variance explained. This can be examined with a cumulative plot of the fraction of the total variance explained by the top K components, i.e.,

$$\text{var explained} = \frac{\sum_{i=1}^K \lambda_i}{\sum_{i=1}^N \lambda_i} \quad (12)$$

where λ_i is the i^{th} eigenvalue and N is the total number of components (the original number of dimensions in the data). The intrinsic dimensionality is often quantified by the K necessary to explain a large proportion of the total variance of the data (often a defined threshold, e.g., 90%).

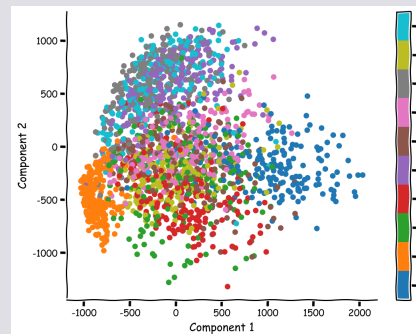


Nonlinear Dimensionality Reduction

Unsupervised Reduction

PCA and t-SNE are unsupervised non-linear dimensionality reduction methods. Nonlinear methods can be more powerful, they can also be sensitive to noise. In contrast, linear methods are useful for their simplicity and robustness. Comparing PCA and t-SNE for data visualization. Using t-SNE, we could visualize clusters in the data corresponding to different digits. While PCA was able to separate some clusters (e.g., 0 vs 1), it performed poorly overall. However, the results of t-SNE can change depending on the choice of perplexity.

Visualize MNIST in 2D using PCA



Visualize MNIST in 2D using t-SNE

