Principal component analysis II: the top k directions

Sanjoy Dasgupta

University of California, San Diego

Topics we'll cover

1. Projecting onto multiple directions
2. Principal component analysis
3. Reconstruction from projections

Projection onto multiple directions

Projecting $x \in \mathbb{R}^d$ into the k-dimensional subspace defined by vectors $u_1, \ldots, u_k \in \mathbb{R}^d$.

This is easiest when the u_i's are **orthonormal**:

- They have length one.
- They are at right angles to each other: $u_i \cdot u_j = 0$ when $i \neq j$

The projection is a k-dimensional vector:

$$(x \cdot u_1, x \cdot u_2, \ldots, x \cdot u_k) \;=\; \underbrace{\begin{pmatrix} \longleftarrow & u_1 & \longrightarrow \\ \longleftarrow & u_2 & \longrightarrow \\ & \vdots & \\ \longleftarrow & u_k & \longrightarrow \end{pmatrix}}_{\text{call this } U^T} \begin{pmatrix} \uparrow \\ x \\ \downarrow \end{pmatrix}$$

U is the $d \times k$ matrix with columns $u_1, \ldots, u_k$.

Projection onto multiple directions: example

E.g. project data in $\mathbb{R}^4$ onto the first two coordinates.

Take vectors $u_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ (notice: orthonormal)

Write $U^T = \begin{pmatrix} \longleftarrow & u_1 & \longrightarrow \\ \longleftarrow & u_2 & \longrightarrow \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$

The projection of $x \in \mathbb{R}^4$ is $U^T x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

The best k-dimensional projection

Let Σ be the $d \times d$ covariance matrix of X.
In $O(d^3)$ time, we can compute its **eigendecomposition**, consisting of

- real **eigenvalues** $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$
- corresponding **eigenvectors** $u_1, \ldots, u_d \in \mathbb{R}^d$ that are orthonormal (unit length and at right angles to each other)

Fact: Suppose we want to map data $X \in \mathbb{R}^d$ to just k dimensions, while capturing as much of the variance of X as possible. The best choice of projection is:

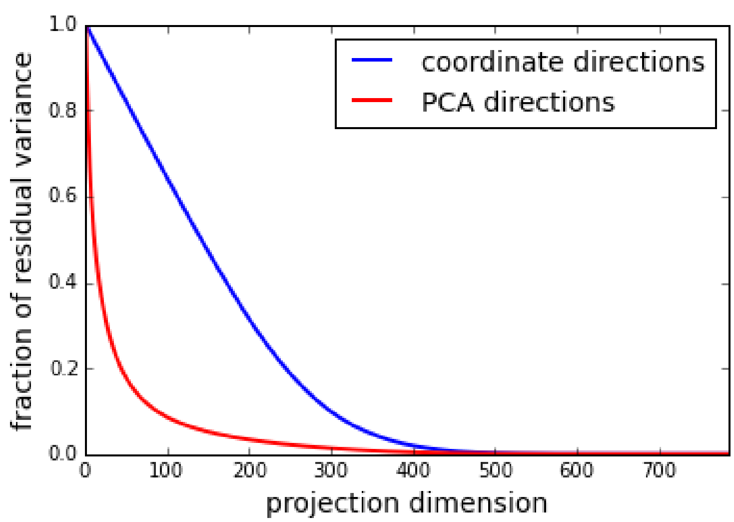
$$x \mapsto (u_1 \cdot x, u_2 \cdot x, \ldots, u_k \cdot x),$$

where u_i are the eigenvectors described above.

This projection is called **principal component analysis** (PCA).

Example: MNIST

Contrast coordinate projections with PCA:



Applying PCA to MNIST: examples

Reconstruct this original image from its PCA projection to k dimensions.

$k = 200$

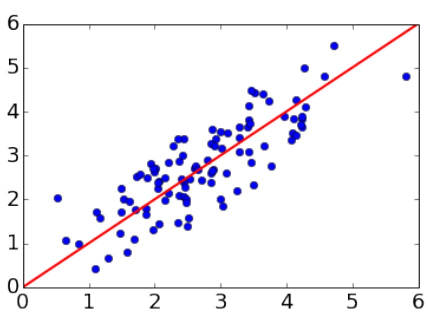
$k = 150$

$k = 100$

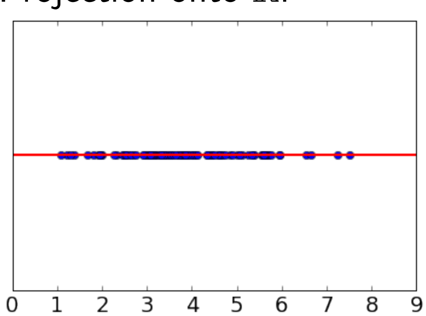
$k = 50$

How do we get these **reconstructions**?

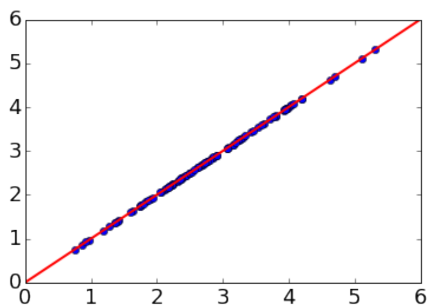
Reconstruction from a 1-d projection



Projection onto $\mathbb{R}$:



Reconstruction in $\mathbb{R}^2$:



Reconstruction from projection onto multiple directions

Projecting into the k-dimensional subspace defined by **orthonormal** $u_1, \ldots, u_k \in \mathbb{R}^d$.

The projection of x is a k-dimensional vector:

$$(x \cdot u_1, x \cdot u_2, \ldots, x \cdot u_k) \;=\; \underbrace{\begin{pmatrix} \longleftarrow & u_1 & \longrightarrow \\ \longleftarrow & u_2 & \longrightarrow \\ & \vdots & \\ \longleftarrow & u_k & \longrightarrow \end{pmatrix}}_{\text{call this } U^T} \begin{pmatrix} \uparrow \\ x \\ \downarrow \end{pmatrix}$$

The reconstruction from this projection is:

$$(x \cdot u_1)u_1 + (x \cdot u_2)u_2 + \cdots + (x \cdot u_k)u_k = UU^T x.$$

MNIST: image reconstruction

Reconstruct this original image x from its PCA projection to k dimensions.

$k = 200$

$k = 150$

$k = 100$

$k = 50$

Reconstruction $UU^T x$, where U's columns are top k eigenvectors of Σ.