Data Mining: Linear Algebra (Recap)

Tom Heskes

Linear Algebra

- vectors and matrices
- operations on these
- SVD and PCA

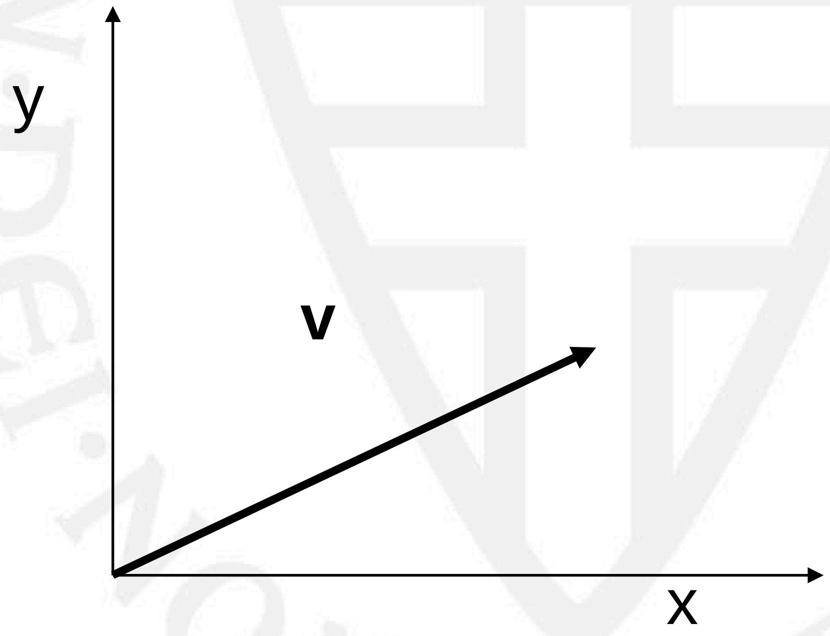
- Note: see Appendix A and B of TSK

Vectors

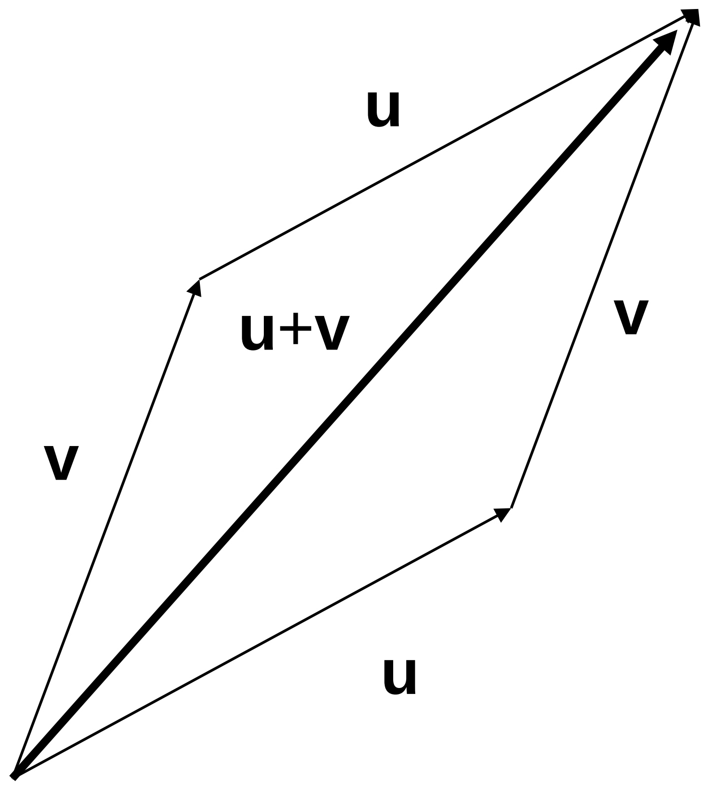
- A vector has a length and a direction
- Interpretation: a line in N dimensional space
- For example, 2 steps in x-direction and 1 step in y-direction

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

$$\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$



Vectors: Addition

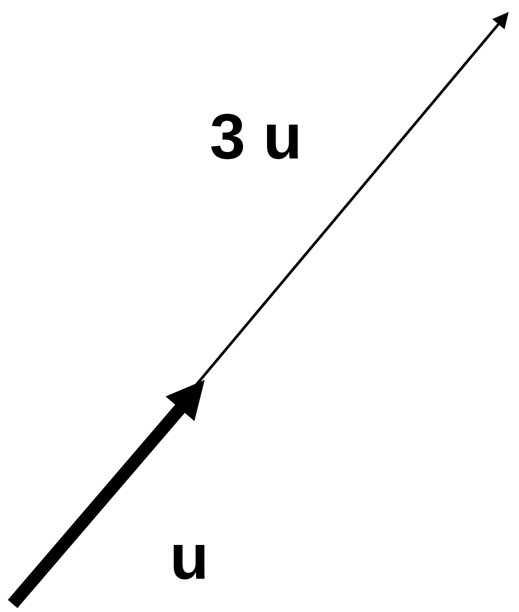


- Head-to-tail method
- Note: interpretation always relative to “tail”

Properties of Vector Addition

- Commutativity: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
- Associativity: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
- Identity element: $\mathbf{u} + \mathbf{0} = \mathbf{u}$
- Additive inverse: $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$

Vectors: Scalar Multiplication



- same direction
- only changes length

Multiplying a Vector by a Scalar

- Associativity: $a\,(b\,\mathbf{u}) = (a\,b)\,\mathbf{u}$
- Distributivity (I): $(a + b)\,\mathbf{u} = a\,\mathbf{u} + b\,\mathbf{u}$
- Distributivity (II): $a\,(\mathbf{u} + \mathbf{v}) = a\,\mathbf{u} + a\,\mathbf{v}$
- Scalar identity: $1\,\mathbf{u} = \mathbf{u}$

Vector Space

- Set of vectors with associated set of scalars (real numbers)
- Closed under addition and scalar multiplication
- Any vector can be represented as a linear combination of a basis $\{\mathbf{u}_1,\ldots,\ \mathbf{u}_n\}$, i.e., we can always find $\{a_1,\ldots,a_n\}$ such that

$$\mathbf{v} = \sum_{i=1}^{n} a_i \mathbf{u}_i$$

Basis Vectors

- Basis vectors “span” the vector space
- Dimension of vector space: minimum number of basis vectors needed
- Typically unit length and orthogonal (more later)
- For example, in two-dimensional Euclidean space:

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Dot Product

- Dot product or scalar product or inner product:

$$\mathbf{u}\cdot\mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n = \sum_{i=1}^{n} u_iv_i$$

- $\begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -3 \\ 4 \end{bmatrix} = 2\times(-3) + 1\times 4 = -2$

$$\begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} = (-2)\times 2 + 1\times 1 + 3\times(-1) = -6$$

Orthogonal Vectors

- Two vectors **u** and **v** are said to be orthogonal if

$$\mathbf{u} \cdot \mathbf{v} = 0$$

- For example,

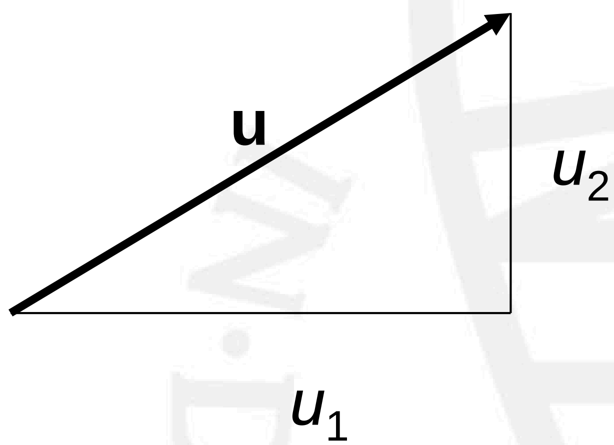
$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -2 \\ -1 \end{bmatrix} = 1 \times 3 + 2 \times (-2) + (-1)(-1) = 3 - 4 + 1 = 0$$

- Euclidean space: orthogonal = perpendicular

Length and Angles

- Length of a vector: $\|\mathbf{u}\| = \sqrt{\mathbf{u}\cdot\mathbf{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2} = \sqrt{\sum_{i=1}^{n} u_i^2}$

- Generalization of Pythogoras:



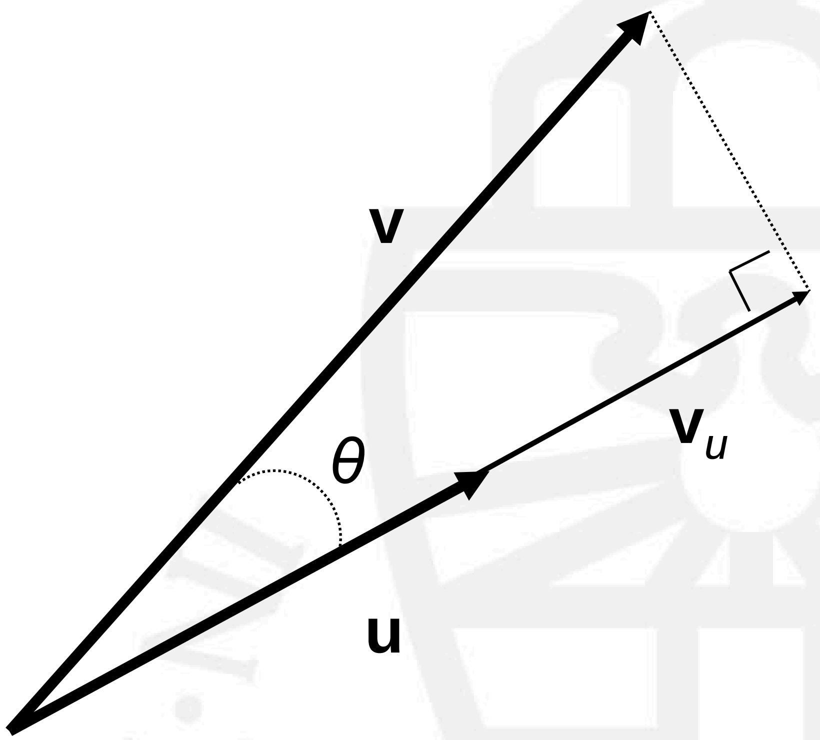
- Dot product in terms of angles:

$$\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\|\cos(\theta)$$

Projection

- Projection of **v** on **u**:

$$\mathbf{v}_u = \frac{(\mathbf{u}\cdot\mathbf{v})\,\mathbf{u}}{\|\mathbf{u}\|^2} = \|\mathbf{v}\|\cos(\theta)\,\frac{\mathbf{u}}{\|\mathbf{u}\|}$$



$$\cos(\theta) = \frac{\|\mathbf{v}_u\|}{\|\mathbf{v}\|}$$

- Example:

$$\mathbf{v} = \begin{bmatrix}2\\1\end{bmatrix} \qquad \mathbf{u} = \begin{bmatrix}1\\0\end{bmatrix} \qquad \mathbf{v}_u = ?\begin{bmatrix}2\\0\end{bmatrix}$$

Linear (In)dependence

- A set of vectors is linearly {independent / dependent} if {no / at least one} vector in the set can be written as a linear combination of the other vectors in the set.

- $\left\{\begin{bmatrix}2\\1\end{bmatrix},\begin{bmatrix}0\\3\end{bmatrix}\right\}?$ **independent**

  $\left\{\begin{bmatrix}2\\1\end{bmatrix},\begin{bmatrix}0\\3\end{bmatrix},\begin{bmatrix}1\\-4\end{bmatrix}\right\}?$ **dependent**

  $\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}?$ **independent**

  $\left\{\begin{bmatrix}2\\1\end{bmatrix},\begin{bmatrix}-4\\-2\end{bmatrix}\right\}?$ **dependent**

Vectors and Data analysis

- Often used in computing similarities and distances between objects
- For example, cosine similarity

$$\cos(\mathbf{u},\mathbf{v}) = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$$

- Euclidean distance

$$dist(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\| = \sqrt{(\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v})}$$

Matrices

- A matrix is a set of elements, organized into rows and columns

$$\begin{array}{cc} & \text{rows} \\ \text{columns} & \begin{bmatrix} a & b \\ c & d \end{bmatrix} \end{array}$$

- "*n* by *m*" matrix: *n* rows and *m* columns

Matrix as a Set of Vectors

- First subscript indexes rows
- Second subscript indexes columns

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{1*} \\ \mathbf{x}_{2*} \\ \vdots \\ \mathbf{x}_{n*} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{*1} & \mathbf{x}_{*2} & \cdots & \mathbf{x}_{*n} \end{bmatrix}$$

- Column of row vectors or row of column vectors

The Transpose of a Matrix

$$\mathbf{X}^T = \begin{bmatrix} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \end{bmatrix}^T$$

$$= \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1m} & x_{2m} & \cdots & x_{nm} \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

$$\mathbf{A}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 14 & 2 \\ 3 & 14 \end{bmatrix}$$

$$\mathbf{B}^T = \begin{bmatrix} 14 & 3 \\ 2 & 14 \end{bmatrix}$$

Note that $(\mathbf{X}^T)^T = \mathbf{X}$

Properties of Matrix Addition

- Commutative: $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
- Associative: $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$
- Identity: $\mathbf{A} + \mathbf{0} = \mathbf{A}$
- Additive inverse: $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$

Multiplying a Matrix by a Scalar

- Associativity: $b\,(c\,\mathbf{A}) = (b\,c)\,\mathbf{A}$
- Distributivity (I): $(b + c)\,\mathbf{A} = b\,\mathbf{A} + c\,\mathbf{A}$
- Distributivity (II): $c\,(\mathbf{A} + \mathbf{B}) = c\,\mathbf{A} + c\,\mathbf{B}$
- Scalar identity: $1\,\mathbf{A} = \mathbf{A}$

Matrix Multiplication

- If ${}_m\mathbf{C}_p = {}_m\mathbf{A}_n\,{}_n\mathbf{B}_p$

  then $c_{ij} = \mathbf{a}_{i*}^T \cdot \mathbf{b}_{*j} = \sum_{k=1}^{n} a_{ik} b_{kj}$

$$\begin{bmatrix} \Rightarrow & \Rightarrow & \cdots & \Rightarrow \end{bmatrix} \begin{bmatrix} \Downarrow \\ \Downarrow \\ \cdots \\ \Downarrow \end{bmatrix}$$

- For example,

$$\mathbf{C} = \begin{bmatrix} -1 & 2 & 3 \\ 2 & -2 & 1 \end{bmatrix} \begin{bmatrix} 5 & 4 \\ 0 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} (-1)(5)+(2)(0)+(3)(-1) & (-1)(4)+(2)(2)+(3)(1) \\ (2)(5)+(-2)(0)+(1)(-1) & (2)(4)+(-2)(2)+(1)(1) \end{bmatrix} = \begin{bmatrix} -8 & 3 \\ 9 & 5 \end{bmatrix}$$

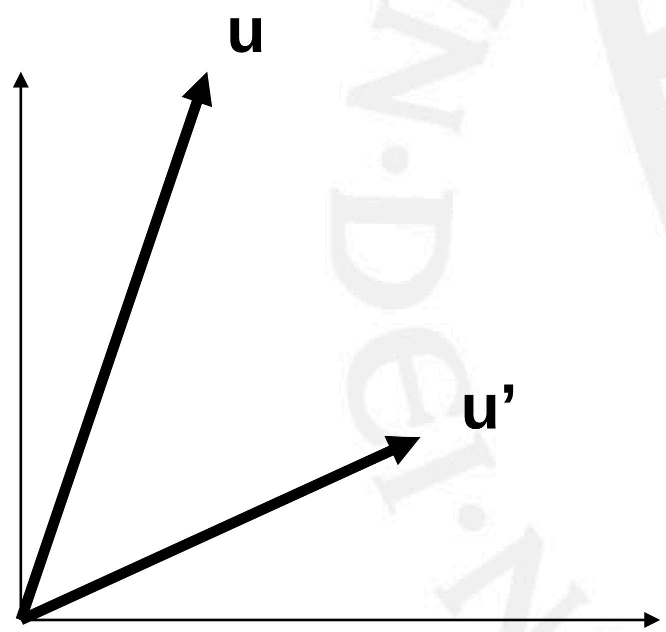
Multiplying a Matrix by a Matrix

- Associativity: $\mathbf{A}\,(\mathbf{B}\,\mathbf{C}) = (\mathbf{A}\,\mathbf{B})\,\mathbf{C}$
- Distributivity: $(\mathbf{B} + \mathbf{C})\,\mathbf{A} = \mathbf{B}\,\mathbf{A} + \mathbf{C}\,\mathbf{A}$
  and
  $\mathbf{A}\,(\mathbf{B} + \mathbf{C}) = \mathbf{A}\,\mathbf{B} + \mathbf{A}\,\mathbf{C}$
- Identity matrix: $\mathbf{I}\,\mathbf{A} = \mathbf{A} = \mathbf{A}\,\mathbf{I}$
- *No commutativity:* $\mathbf{A}\,\mathbf{B} \neq \mathbf{B}\,\mathbf{A}$

Vector Interpretation

- Think of a matrix as a transformation on a line or set of lines

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} u_1' \\ u_2' \end{bmatrix}$$



Scaling Matrix

- Same direction, only changing length:

$$\mathbf{C} = \begin{bmatrix} c & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c \end{bmatrix}$$

- Easy to check: **C u** = c **u** for any **u**

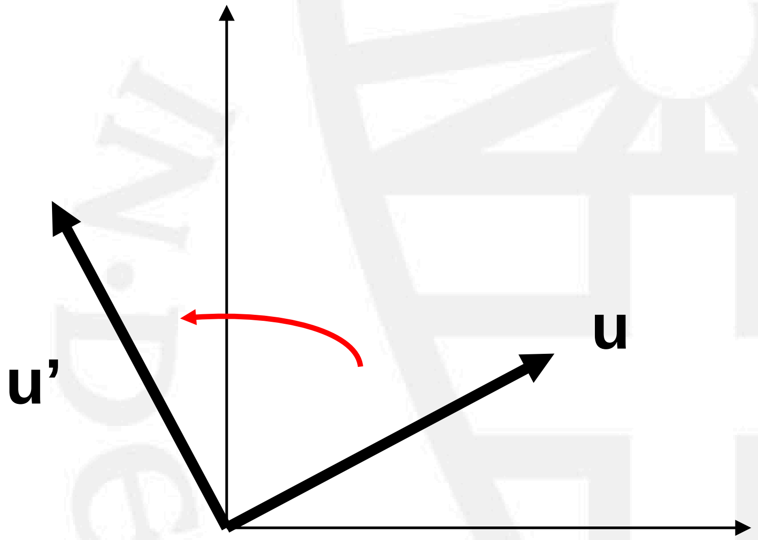
Rotation Matrix

- Rotate, leave length intact
- Two-dimensional rotation around angle θ

$$\mathbf{R} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

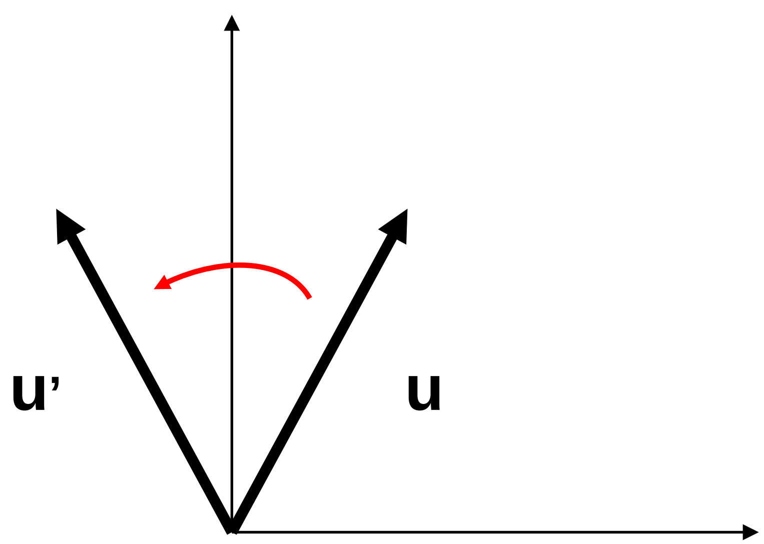
- For example, $\theta = 90^o$:

$$\mathbf{u}' = \begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -u_2 \\ u_1 \end{bmatrix}$$



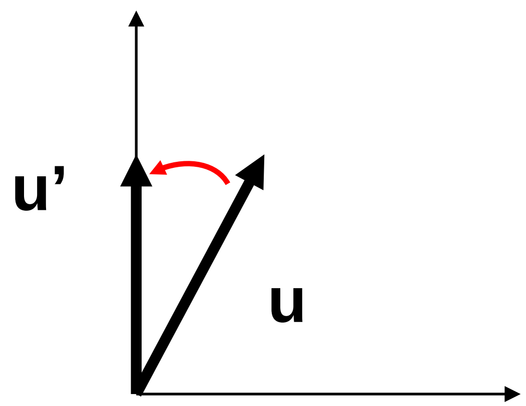
Reflection Matrix

- Reflect along one of the coordinate axes
- Example:



$$\mathbf{u}' = \begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \begin{bmatrix} -u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

Projection Matrix

- Project onto lower-dimensional subspace
- 

$$\mathbf{u}' = \begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \begin{bmatrix} 0 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

- General case, projection onto vector **u**:

$$\mathbf{P} = \frac{\mathbf{u}\,\mathbf{u}^T}{\|\mathbf{u}\|^2} \qquad \text{see} \qquad \mathbf{P}\,\mathbf{v} = \frac{\mathbf{u}\left(\mathbf{u}^T\mathbf{v}\right)}{\|\mathbf{u}\|^2} = \frac{\left(\mathbf{u}\cdot\mathbf{v}\right)\mathbf{u}}{\|\mathbf{u}\|^2}$$

Linear transformations

- Matrices are linear transformations:

$$\mathbf{A}\ (\mathbf{u} + \mathbf{v}) = \mathbf{A}\ \mathbf{u} + \mathbf{A}\ \mathbf{v}$$

- *Column space*: the set of all column vectors. Any result can be written as the linear combination of those vectors:

$$\mathbf{v} = \mathbf{A}\mathbf{u} = \sum_{j=1}^{n} u_j \mathbf{a}_{*j}$$

- Similarly, *row space*, set of all row vectors:

$$\mathbf{v} = \mathbf{u}\ \mathbf{A} = \sum_{i=1}^{m} u_i \mathbf{a}_{i*}$$

Rank of a Matrix

- Dim of {column/row} space {d_{col} /d_{row}}: number of linearly independent {columns/rows} of **A**
- It can be shown that $d_{col} = d_{row}$, which is called the rank of the matrix

$$\begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}?$$

2

$$\begin{bmatrix} 2 & 0 & 1 \\ 1 & 3 & -4 \end{bmatrix}?$$

2

$$\begin{bmatrix} 2 & 6 & -4 \\ 1 & 3 & -2 \end{bmatrix}?$$

1

$$\begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}?$$

1

Matrix Inverse

- Only for square matrices with rank equal to the number of rows (and thus columns)
- Rotation matrices are invertible, projection matrices not
- The inverse of a matrix is such that:
  $\mathbf{AA^{-1} = I = A^{-1}A}$ with $\mathbf{I}$ the identity matrix
- Inversion is a bit tricky: $\mathbf{(AB)^{-1} = B^{-1}A^{-1}}$

Examples

- Diagonal matrix (but only then...)

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}^{-1} \overset{?}{=} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

- Rotation matrix

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}^{-1} \overset{?}{=} \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

since

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} = \begin{bmatrix} \cos^2(\theta) + \sin^2(\theta) & 0 \\ 0 & \cos^2(\theta) + \sin^2(\theta) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Eigenvalues and Eigenvectors

- **u** is an eigenvector of matrix **A** with corresponding eigenvalue λ if **A u** = λ **u**
- **u** should be such that its direction does not change, only the length
- Example, the matrix

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

has 2 eigenvalue/vector combi's:

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 1\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Eigenvalue Decomposition (1)

- If **A** is an n by n matrix with n independent (orthogonal) eigenvectors $\mathbf{u}_1$ through $\mathbf{u}_n$ and corresponding nonnegative eigenvalues λ_1 through λ_n, then it can be decomposed as

$$\mathbf{A} = \mathbf{U}\,\mathbf{\Lambda}\,\mathbf{U}^{\mathrm{T}} = \sum_{\mathrm{i}=1}^{\mathrm{n}} \lambda_{\mathrm{i}}\,\mathbf{u}_{\mathrm{i}}\,\mathbf{u}_{\mathrm{i}}^{\mathrm{T}}$$

with $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_n]$ and $\mathbf{\Lambda}$ a diagonal matrix with elements λ_1 through λ_n

Eigenvalue Decomposition (2)

- Check:

$$\mathbf{A}\mathbf{u}_j = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^T \mathbf{u}_j = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \delta_{ij} = \lambda_j \mathbf{u}_j$$

- Example:

eigenvectors

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

eigenvalues

Singular Value Decomposition

- Generalization of eigenvalue decomposition to non-square matrices
- **A** is *m* by *n*, can be expressed as

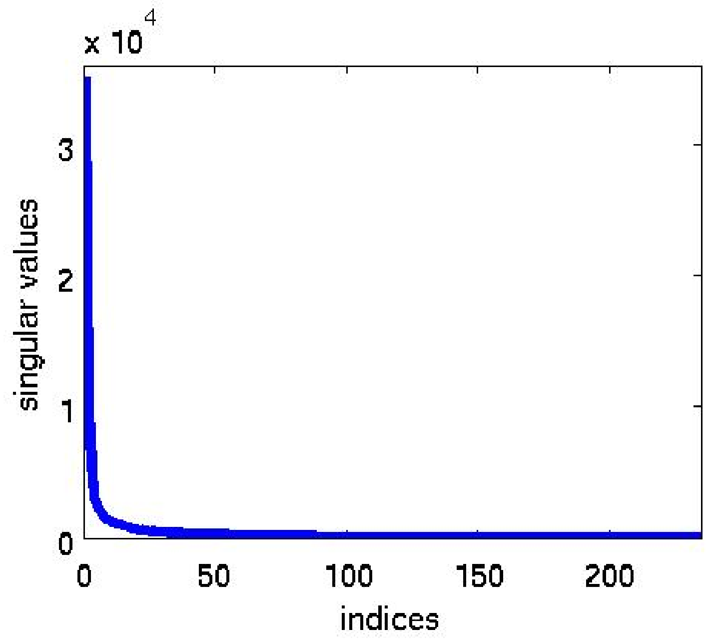
$$\mathbf{A} = \mathbf{U}\,\mathbf{\Sigma}\,\mathbf{V}^{\mathrm{T}} = \sum_{i=1}^{n} \sigma_i\,\mathbf{u}_i\,\mathbf{v}_i^{\mathrm{T}}$$

- **U** is *m* by *m*, **V** is *n* by *n*, **Σ** is a diagonal matrix with non-negative entries; **U** and **V** are orthonormal, i.e., $\mathbf{U}\,\mathbf{U}^T = \mathbf{I}$ and $\mathbf{V}\,\mathbf{V}^T = \mathbf{I}$
- Relation to eigenvalue decomposition: $\mathbf{u}_i$'s (left singular values) are eigenvectors of $\mathbf{A}\mathbf{A}^T$, $\mathbf{v}_i$'s (right singular values) of $\mathbf{A}^T\mathbf{A}$
- Dimensionality reduction: set small singular values to zero

SVD for Dimensionality Reduction

- Treat picture as a matrix **A** of gray values
- Apply SVD: $\mathbf{A} = \mathbf{U}\,\mathbf{\Sigma}\,\mathbf{V}^T$
- Keep only 20 singular values: set $\underline{\sigma}_i = 0$ for $i > 20$
- Transform back to $\underline{\mathbf{A}} = \mathbf{U}\,\underline{\mathbf{\Sigma}}\,\mathbf{V}^T$

original



compressed

Basic Statistics

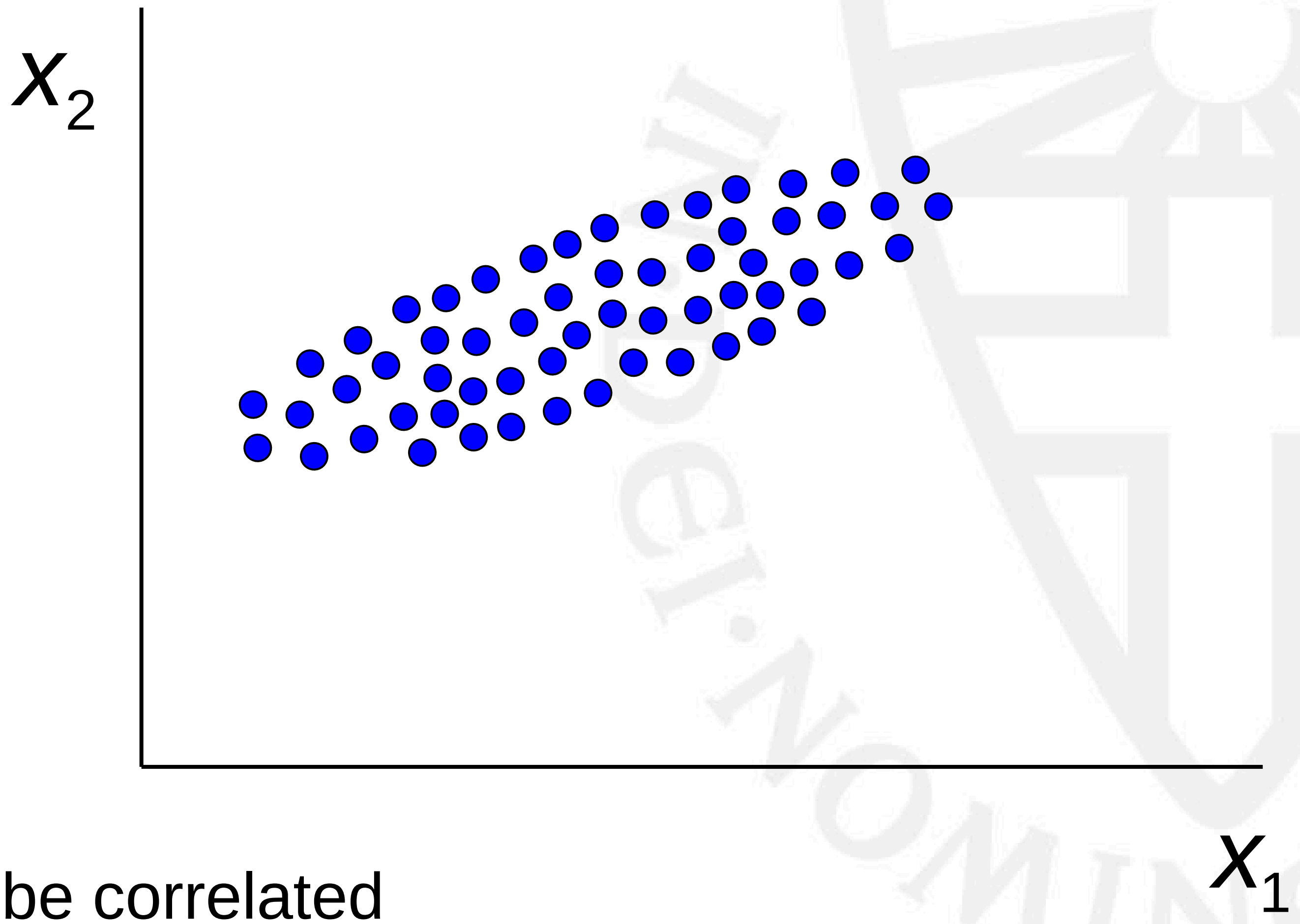
- Given a data matrix **X** with *m* objects in *n* dims
- Mean $\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 & \cdots & \bar{x}_n \end{bmatrix}$ with $\bar{x}_i = \frac{1}{m}\sum_{k=1}^{m} x_{ki}$
- Variance $\mathrm{Var}(x_i) = \frac{1}{m}\sum_{k=1}^{m}(x_{ki} - \bar{x}_i)^2$
- Covariance

$$\mathrm{Cov}(x_i, x_j) = \frac{1}{m}\sum_{k=1}^{m}(x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)$$

- Covariance matrix **Σ** with $\sigma_{ij} = \mathrm{Cov}(x_i, x_j)$

Plot

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} \\ \vdots & \vdots \\ x_{m1} & x_{m2} \end{bmatrix}$$



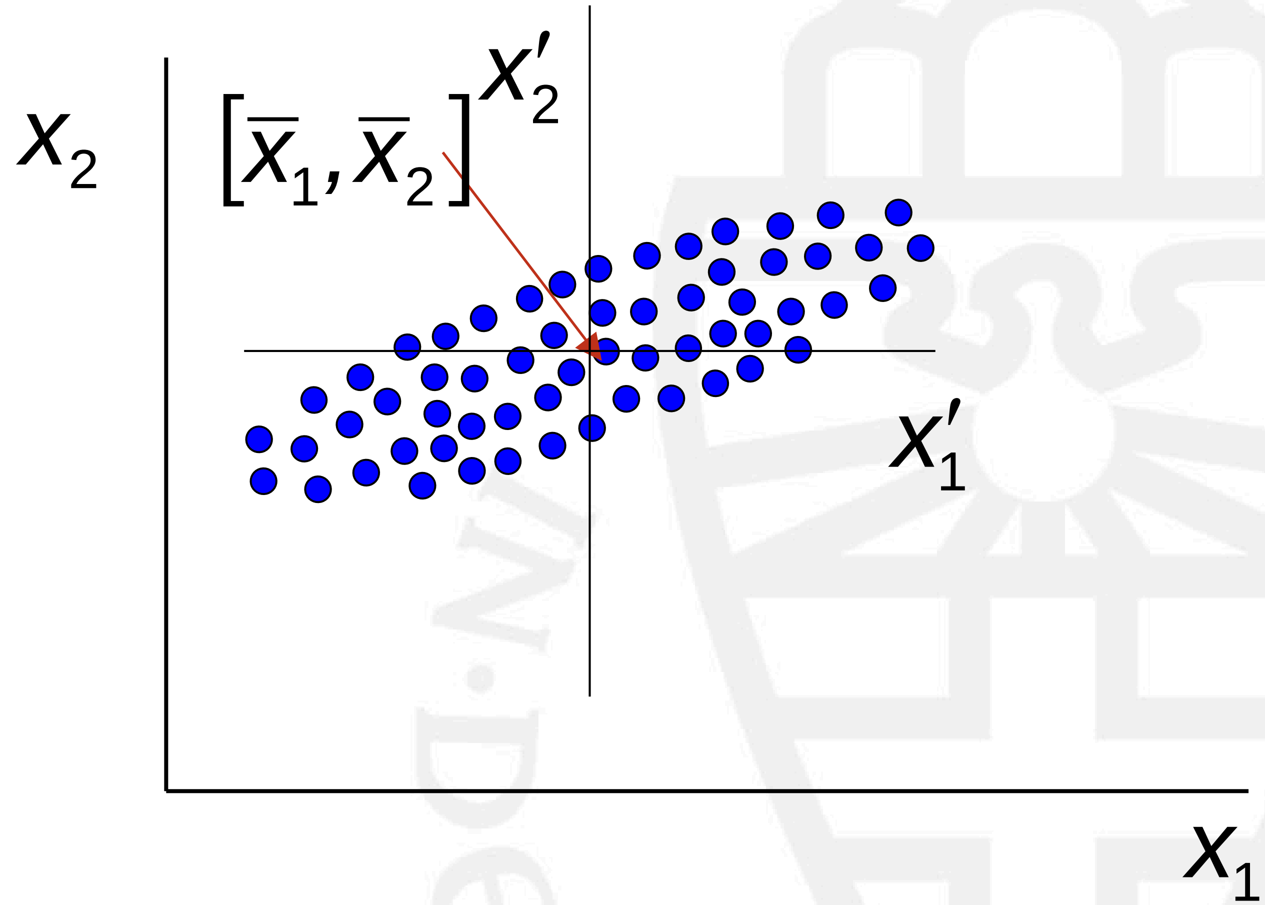
Data appears to be correlated

Compute the Covariance Matrix

- Move to a data centered coordinate system

- Calculate the covariance matrix:

$$\begin{bmatrix} \frac{1}{m}\sum_{k=1}^{m}(x'_{k1})^2 & \frac{1}{m}\sum_{k=1}^{m}x'_{k1}\,x'_{k2} \\ \frac{1}{m}\sum_{k=1}^{m}x'_{k1}\,x'_{k2} & \frac{1}{m}\sum_{k=1}^{m}(x'_{k2})^2 \end{bmatrix}$$



- Diagonal terms are the variances in the two directions

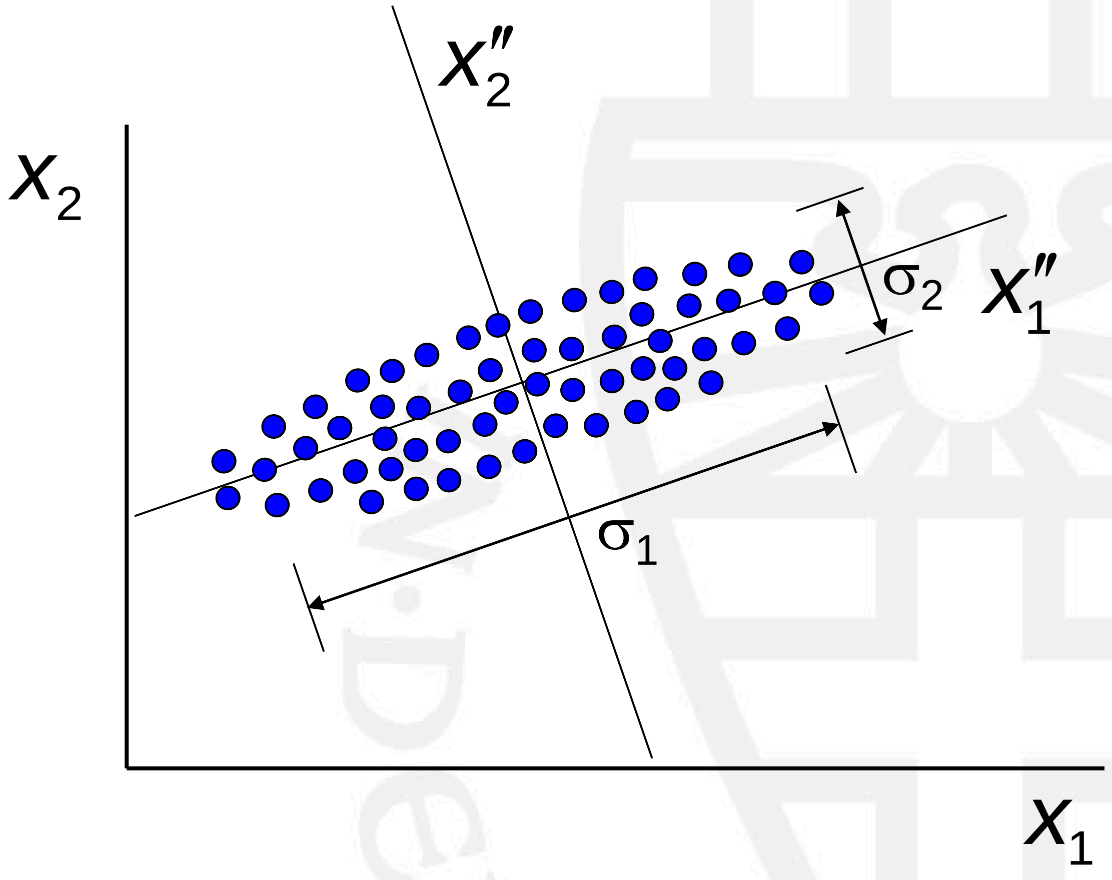
- Off-diagonal terms are the covariances /cross-correlations

Rotate to Remove the Correlations

- Now rotate to decorrelate

- Covariance matrix in the (x_1'', x_2'') coordinate system

$$\begin{bmatrix} \frac{1}{m}\sum_{k=1}^{m}(x_{k1}'')^2 & 0 \\ 0 & \frac{1}{m}\sum_{k=1}^{m}(x_{k2}'')^2 \end{bmatrix}$$



Principal Component Analysis

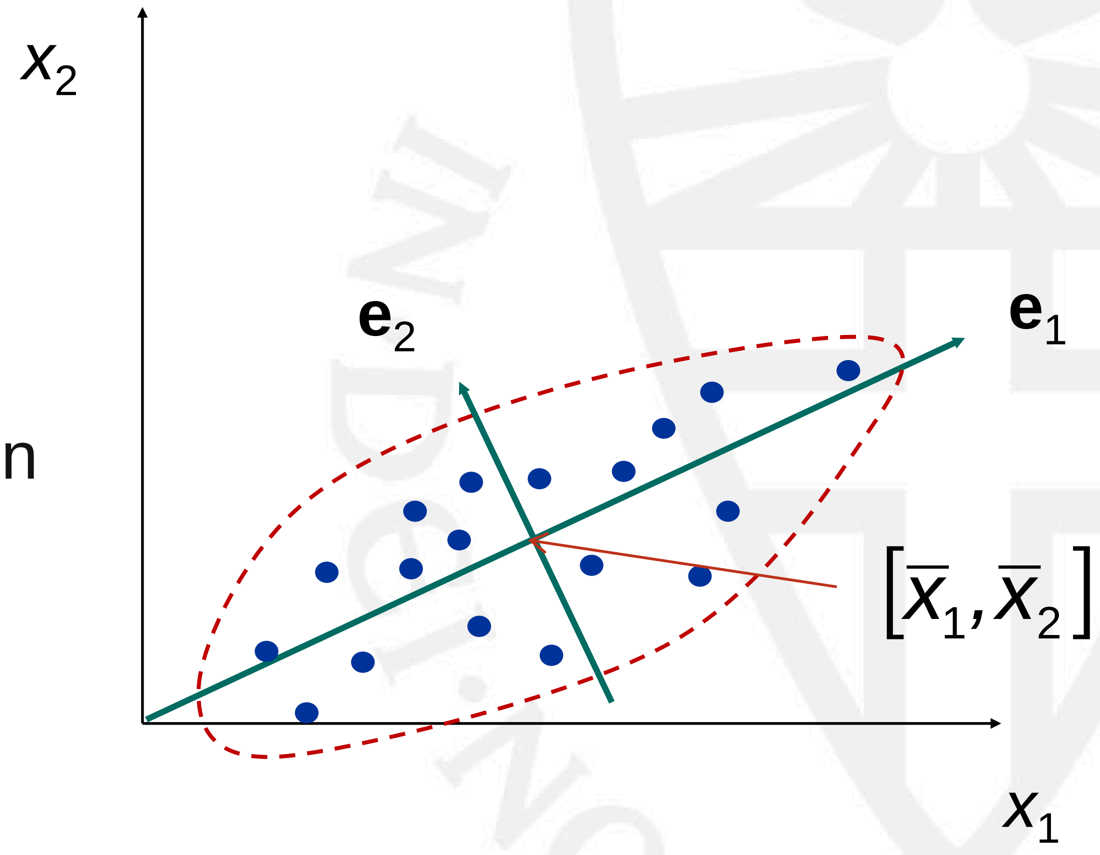
- The direction with the largest variation makes the first principal component
- The direction orthogonal to the principal direction with then the largest variation makes the second principal component, and so on
- These happen to correspond to eigenvectors of the covariance matrix, ordered by their corresponding eigenvalues
- Equivalent: singular value decomposition on mean-centered data

Dimensionality Reduction

- Purpose:
  - Avoid curse of dimensionality
  - Reduce amount of time and memory required by data mining algorithms
  - Allow data to be more easily visualized
  - May help to eliminate irrelevant features or reduce noise
- Techniques
  - Principal Component Analysis
  - Singular Value Decomposition
  - Others: supervised and non-linear techniques

Dimensionality Reduction: PCA (1)

- Goal is to find a projection that captures the largest amount of variation in data
- Find the eigenvectors of the covariance matrix
- The eigenvectors define the new space
- Example: largest variation in the direction of $\mathbf{e}_1$, the first principal component
- Next principal component $\mathbf{e}_2$ perpendicular to $\mathbf{e}_1$

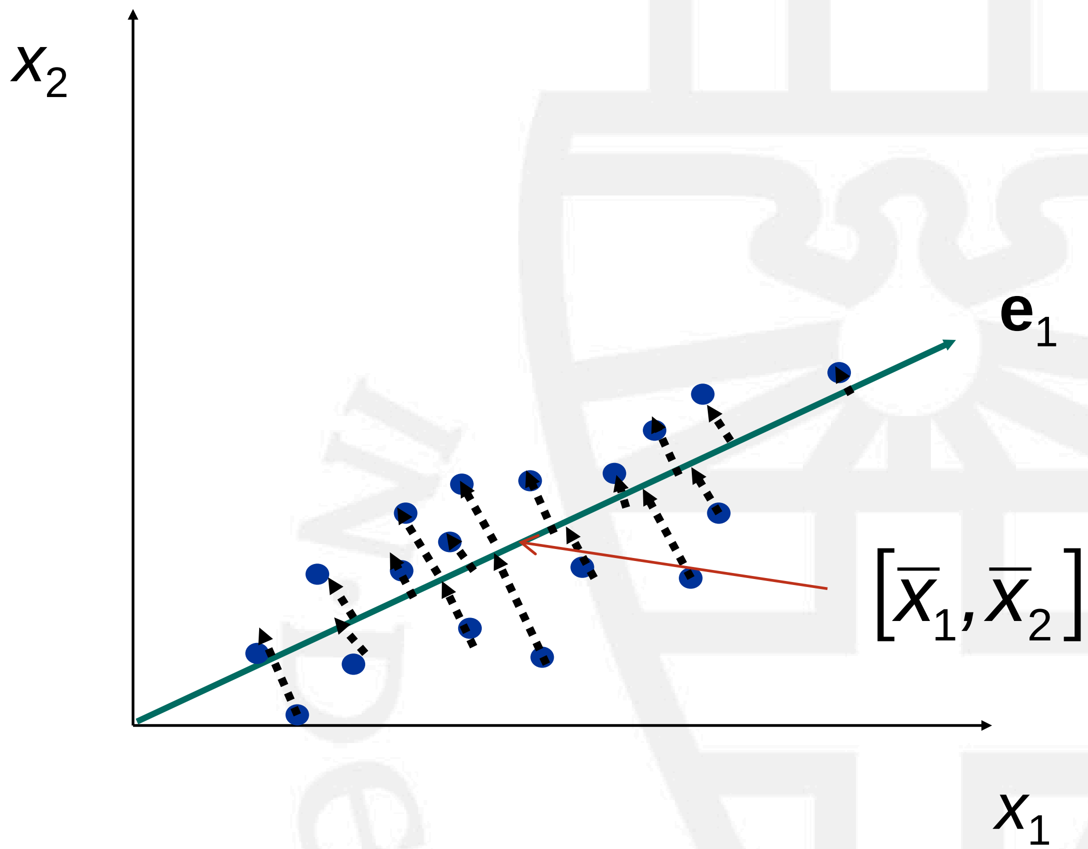


Dimensionality Reduction: PCA (2)

- Project the data points onto the first few principal components
- Example: project onto $\mathbf{e}_1$, i.e., ignore variations in the direction of $\mathbf{e}_2$
- In math, projection matrix: $\mathbf{P} = \dfrac{\mathbf{e}_1 \mathbf{e}_1^T}{\|\mathbf{e}_1\|^2}$
- Contribution in direction of $\mathbf{e}_1$:

$$\mathbf{Px} = \frac{\mathbf{e}_1\left(\mathbf{e}_1^T\mathbf{x}\right)}{\|\mathbf{e}_1\|^2} = \frac{\left(\mathbf{e}_1 \bullet \mathbf{x}\right)\mathbf{e}_1}{\|\mathbf{e}_1\|^2}$$

- Reconstruction: $\tilde{\mathbf{x}} = \overline{\mathbf{x}} + \mathbf{Px} = \overline{\mathbf{x}} + \dfrac{\left(\mathbf{e}_1 \bullet \mathbf{x}\right)\mathbf{e}_1}{\|\mathbf{e}_1\|^2}$



Eigenfaces (1)

- Eigenfaces are the eigenvectors of the covariance matrix of the statistical distribution of the vector space of human faces
- Eigenfaces are the ‘standardized face ingredients’ derived from the statistical analysis of many pictures of human faces
- A human face may be considered to be a combination of these standard faces

Eigenfaces (2)

To generate a set of eigenfaces**:**

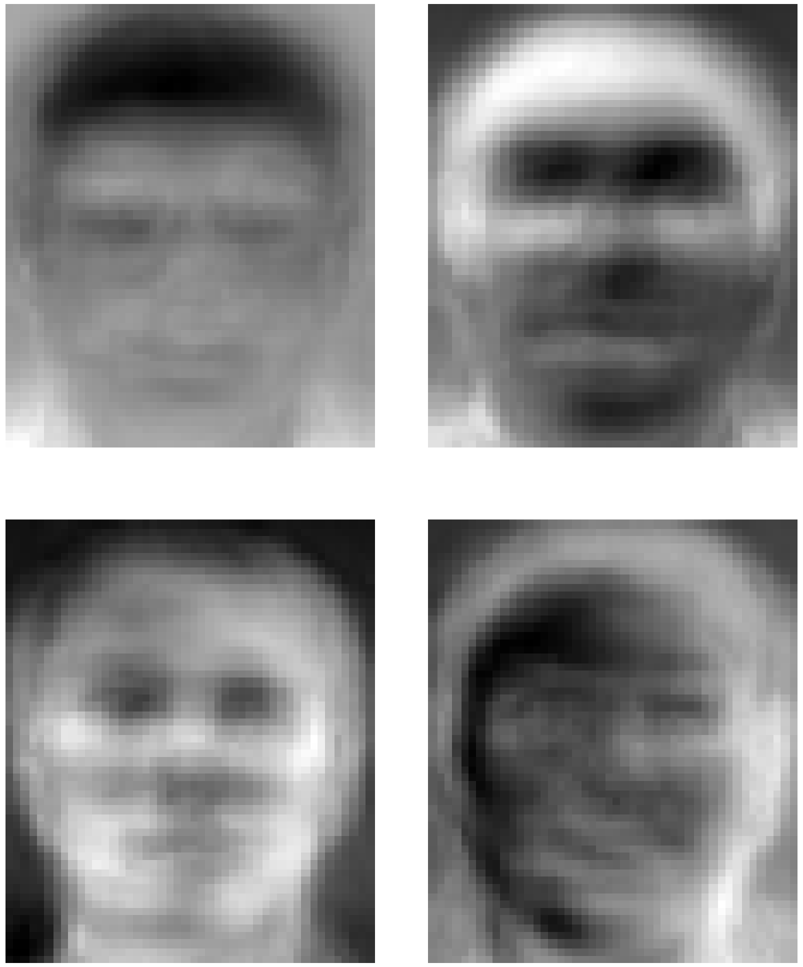
1. Large set of digitized images of human faces is taken under the same lighting conditions

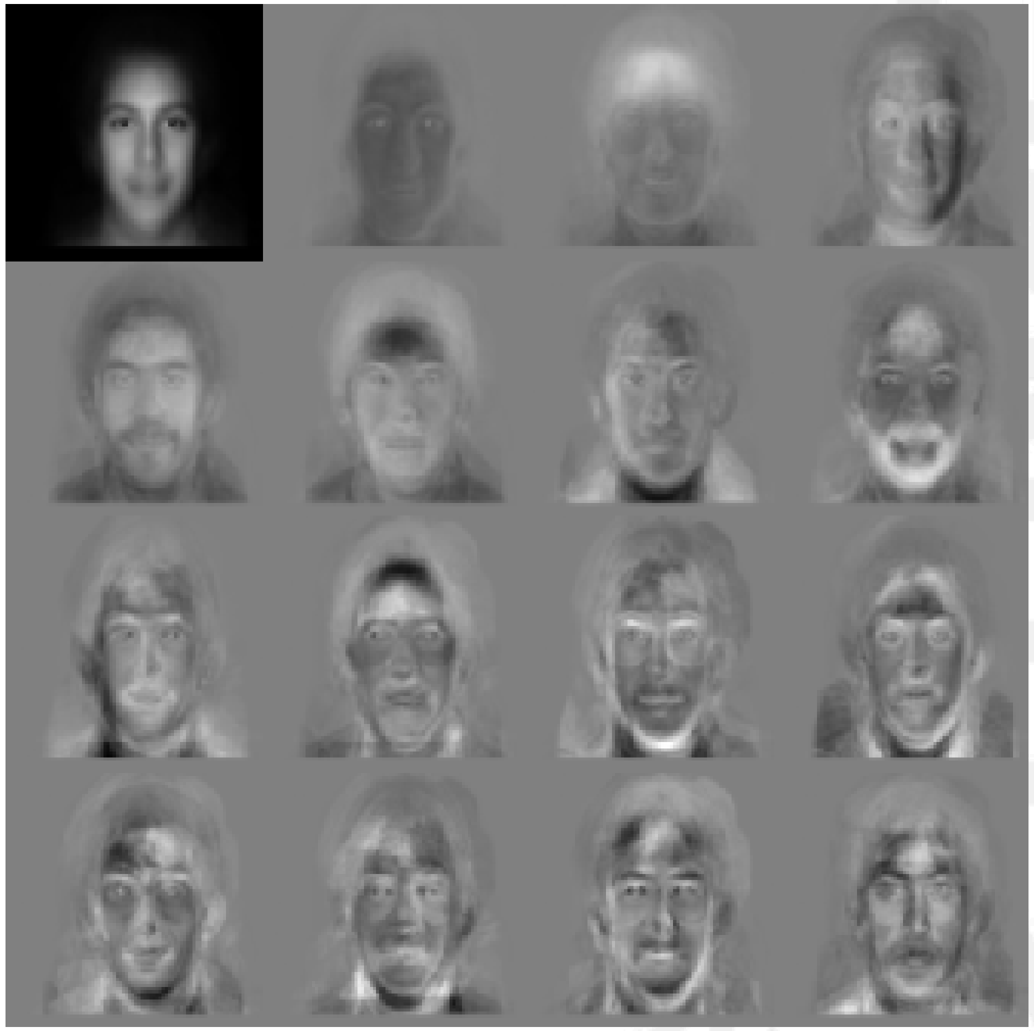
2. The images are normalized to line up the eyes and mouths

3. The eigenvectors of the covariance matrix of the statistical distribution of face image vectors are then extracted

4. These eigenvectors are called eigenfaces

Eigenfaces (3)





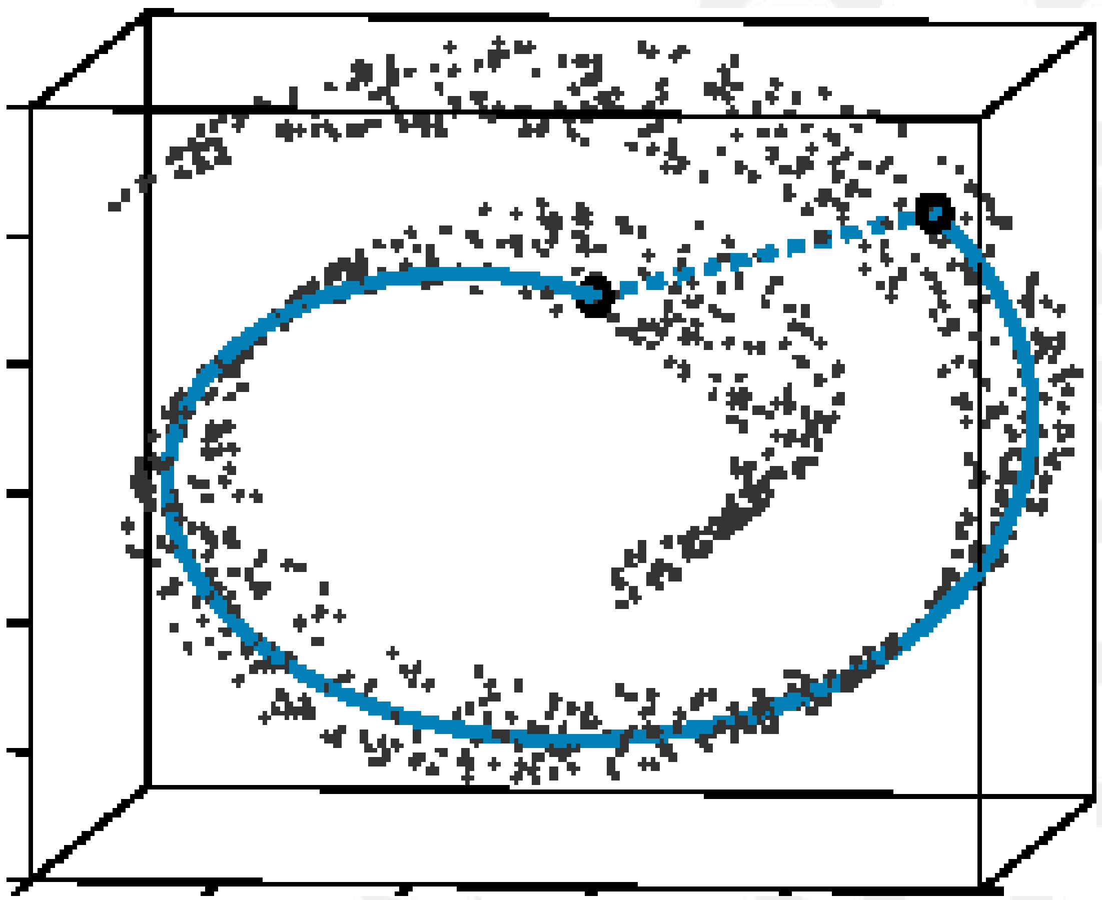
http://en.wikipedia.org/wiki/Eigenface

Eigenfaces (4)

- When properly weighted, eigenfaces can be summed together to create an approximate gray-scale rendering of a human face
- Remarkably few eigenvector terms are needed to give a fair likeness of most people's faces
- Hence eigenfaces provide a means of applying data compression to faces for identification purposes

Dimensionality Reduction: ISOMAP

- Construct a neighbourhood graph
- For each pair of points in the graph, compute the shortest path distances – geodesic distances
- Also works for non-linear manifolds



Tenenbaum, de Silva, Langford (2000)