Geometrical Embeddings in Machine Learning – Preliminary Definitions

R. Basili

Course on *Web Mining and Retrieval*
a.a. 2012-13

January 7, 2013

Change of Basis

Change of Basis

Given two alternative basis $B = \{\underline{b}_1, ..., \underline{b}_n\}$ and $B' = \{\underline{b}'_1, ..., \underline{b}'_n\}$, such that the square matrix $\mathbf{C} = (c_i k)$ describe the change of the basis, i.e.

$$\underline{b}'_k = c_{1k}\underline{b}_1 + c_{2k}\underline{b}_2 + ...c_{nk}\underline{b}_n \qquad \forall k = 1, ..., n$$

Matrix and Change of Basis

Matrix and Change of Basis

The effect of the matrix **C** on a generic vector $\underline{x}$ allows to compute the change of basis according only to the involved basis B and B'. For every $\underline{x} = \sum_{k=1}^{n} x_k \underline{b}_k$ such that in the new basis B', $\underline{x}$ can be expressed by $\underline{x} = \sum_{k=1}^{n} x'_k \underline{b}'_k$, then it follows that:

$$\underline{x} = \sum_{k=1}^{n} x'_k \underline{b}'_k = \sum_{k} x'_k \left(\sum_{i} c_{ik} \underline{b}_i \right) = \sum_{i,k=1}^{n} x'_k c_{ik} \underline{b}_i$$

from which it follows that:

$$x_i = \sum_{k=1}^{n} x'_k c_{ik} \qquad \forall i = 1, \ldots, n$$

Matrix and Change of Basis

Matrix and Change of Basis

The effect of the matrix $\mathbf{C}$ on a generic vector $\underline{x}$ allows to compute the change of basis according only to the involved basis B and B'. For every $\underline{x} = \sum_{k=1}^{n} x_k \underline{b}_k$ such that in the new basis B', $\underline{x}$ can be expressed by $\underline{x} = \sum_{k=1}^{n} x'_k \underline{b}'_k$, then it follows that:

$$\underline{x} = \sum_{k=1}^{n} x'_k \underline{b}'_k = \sum_{k} x'_k \left(\sum_{i} c_{ik} \underline{b}_i \right) = \sum_{i,k=1}^{n} x'_k c_{ik} \underline{b}_i$$

from which it follows that:

$$x_i = \sum_{k=1}^{n} x'_k c_{ik} \qquad \forall i = 1, \ldots, n$$

The above condition suggests that $\mathbf{C}$ is sufficient to describe any change of basis through the matrix vector mutliplication operations:

$$\underline{x} = \mathbf{C} \underline{x}'$$

Matrix and Change of Basis

Matrix and Change of Basis

The effect of the matrix $\mathbf{C}$ on a matrix $\mathbf{A}$ can be seen by studying the case where $\underline{x}, \underline{y}$ are the expression of two vectors in a base B while their counterpart on B' are $\underline{x}', \underline{y}'$, respectively. Now if $\mathbf{A}$ and $\mathbf{B}$ are such that $\underline{y} = \mathbf{A}\underline{x}$ and $\underline{y}' = \mathbf{B}\underline{x}'$, then it follows that:

$$\underline{y} = \mathbf{C}\underline{y}' = \mathbf{A}\underline{x} = \mathbf{A}(\mathbf{C}\underline{x}') = \mathbf{AC}\underline{x}'$$
$$\text{(this means that)}$$
$$\underline{y}' = \mathbf{C}^{-1}\mathbf{AC}\underline{x}'$$

from which it follows that:

$$\mathbf{B} = \mathbf{C}^{-1}\mathbf{AC}$$

The transformation of basis $\mathbf{C}$ is a *similarity transformation* and matrices $\mathbf{A}$ and $\mathbf{C}$ are said *similar*.

Matrix and Change of Basis

Matrix and Change of Basis

The effect of the matrix $\mathbf{C}$ on a matrix $\mathbf{A}$ can be seen by studying the case where $\underline{x}, \underline{y}$ are the expression of two vectors in a base B while their counterpart on B' are $\underline{x}', \underline{y}'$, respectively. Now if $\mathbf{A}$ and $\mathbf{B}$ are such that $\underline{y} = \mathbf{A}\underline{x}$ and $\underline{y}' = \mathbf{B}\underline{x}'$, then it follows that:

$$\underline{y} = \mathbf{C}\underline{y}' = \mathbf{A}\underline{x} = \mathbf{A}(\mathbf{C}\underline{x}') = \mathbf{AC}\underline{x}'$$
$$\text{(this means that)}$$
$$\underline{y}' = \mathbf{C}^{-1}\mathbf{AC}\underline{x}'$$

from which it follows that:

$$\mathbf{B} = \mathbf{C}^{-1}\mathbf{AC}$$

The transformation of basis $\mathbf{C}$ is a *similarity transformation* and matrices $\mathbf{A}$ and $\mathbf{C}$ are said *similar*.

Eigenvalues and eigenvectors

Eigenvectors

An eigenvector $\underline{x}$ for a matrix $\mathbf{A}$ is a non-zero vector for which a scalar $\lambda \in \Re$ exists such that

$$\mathbf{A}\underline{x} = \lambda \underline{x}$$

The value of the scalar λ is called eigenvalue of $\mathbf{A}$ associated to $\underline{x}$, and correspond to the scaling factor along the direction of $\underline{x}$.

Example

$$\mathbf{A} = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \quad \text{and} \quad \underline{x} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \end{pmatrix} = 2 \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

$\underline{x}$ is an eigenvector of $\mathbf{A}$ and $\lambda = 2$ is its eigenvalue.

Eigenvalues, eigenvectors and some properties

Eigenvalues, eigenvectors: Some Consequences

When a matrix $\mathbf{A}$ has an eigenvector $\underline{x}$ it must satisfy the following condition:

$$\mathbf{A}\underline{x} = \lambda \underline{x}$$

We can rewrite the condition $\mathbf{A}\underline{x} = \lambda \underline{x}$ as

$$(\mathbf{A} - \lambda \mathbf{I}\underline{x}) = \underline{0}$$

where $\mathbf{I}$ is the Identity matrix.

In order for a non-zero vector $\underline{x}$ to satisfy this equation, $\mathbf{A} - \lambda \mathbf{I}$ *must not be invertible*(see next slide).
The consequence is that the determinant of $\mathbf{A} - \lambda \mathbf{I}$ must equal 0. This function is $p(\lambda) = det(\mathbf{A} - \lambda \mathbf{I})$, called the *characteristic polynomial* of $\mathbf{A}$.
The eigenvalues of $\mathbf{A}$ are simply the roots of the characteristic polynomial of $\mathbf{A}$.

Eigenvalues, eigenvectors and some properties: Proof

$\mathbf{A} - \lambda\mathbf{I}$ *must not be invertible: Why?*

$\mathbf{A} - \lambda\mathbf{I}$ *must not be invertible*, as otherwise, if $\mathbf{A} - \lambda\mathbf{I}$ has an inverse, and

$$\begin{aligned}(\mathbf{A} - \lambda\mathbf{I})^{-1}(\mathbf{A} - \lambda\mathbf{I})\underline{x} &= (\mathbf{A} - \lambda\mathbf{I})^{-1}\underline{0x} \\ \mathbf{I}\underline{x} &= \underline{0}.\end{aligned}$$

the zero vector is derived. This is not admissibile as, by definition, $\underline{x} \neq \underline{0}$.

Eigenvalues and eigenvectors

An example: computing eigenvalues

Let $\mathbf{A} = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix}$. Then

$p(\lambda) = (2-\lambda)(-1-\lambda) - (-4)(-1) = \lambda^2 - \lambda - 6 = (\lambda - 3)(\lambda + 2)$

The eigenvectors are then the solution of the linear equation system given by $(\mathbf{A} - \lambda\mathbf{I})\underline{x} = \underline{0}$.

Given the first eigenvalue $\lambda_1 = 3$, $(\mathbf{A} - 3\mathbf{I})\underline{x} = \underline{0}$ gives the following system:

$$\begin{cases} -x_1 - 4x_2 & = & 0 \\ -x_1 - 4x_2 & = & 0 \end{cases}$$

This suggests that all vectors of the form $\alpha\underline{x}_1$ are eigenvectors with $\underline{x}_1^T = (-4, 1)$. The span of the vector $(-4, 1)^T$ is the **eigenspace** corresponding to $\lambda_1 = 3$. Correspondingly, the span of the vector $\underline{x}_2 = (1, 1)^T$ corresponds to the eigenspace of $\lambda_2 = -2$.

Notice that $\underline{x}_1$ and $\underline{x}_2$ are linearly independent, so they can form a basis.

Eigenvalues and eigenvectors

Eigenvectors of Symmetric matrices

A symmetric non singular real-valued matrix $\mathbf{A}$ is such that $\mathbf{A} = \mathbf{A}^T$, and on two dimensions, this means that :

$$\begin{array}{ll} i) & a_{11}, a_{22} \neq 0 \\ ii) & a_{12} = a_{21} = a \end{array}$$

In order for $\mathbf{A}$ to have two real eigenvalues the following must hold:

$$\begin{aligned} p(\lambda) &= (a_{11} - \lambda)(a_{22} - \lambda) - a^2 = \\ &= \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a^2 = 0 \end{aligned}$$

from which eigenvalues are distinct **iff**:

$$(a_{11} - a_{22})^2 + 4a^2 \geq 0$$

The above inequality is always satisfied, with the 0 case only when $\mathbf{A} = \mathbf{I}$.

Eigenvalues and eigenvectors

Eigenvectors and orthogonality

Whenever a matrix $\mathbf{A}$ has n distinct eigenvectors $\underline{x}_i$ with all real-valued and distinct eigenvalues λ_i, it is called *non-degenerate*.

Eigenvalues and eigenvectors

Eigenvectors and orthogonality

Whenever a matrix $\mathbf{A}$ has n distinct eigenvectors $\underline{x}_i$ with all real-valued and distinct eigenvalues λ_i, it is called *non-degenerate*.
A non degenerate matrix $\mathbf{A}$ has all the eigenvectors mutually orthogonal.

In fact, given two any eigenvectors $\underline{x}_1 \neq \underline{x}_2$, with $\mathbf{A}\underline{x}_i = \lambda_i \underline{x}_i \qquad (i = 1, 2)$, it follows that

$$\lambda_1(\underline{x}_1, \underline{x}2) = (\mathbf{A}\underline{x}_1, \underline{x}_2) = (\underline{x}_1, \mathbf{A}\underline{x}2) = \lambda_2(\underline{x}_1, \underline{x}_2)$$

from which it follows that $\qquad (\lambda_1 - \lambda_2)(\underline{x}_1, \underline{x}_2) = 0$

However as $\lambda_1 \neq \lambda_2$, and $\underline{x}_1$, $\underline{x}_2$ were arbitrarily chosen, the result is that

$$\forall i, j = 1, \ldots, n \qquad (\underline{x}_i, \underline{x}_j) = \begin{cases} \|\underline{x}_i\|^2 & i = j \\ 0 & i \neq j \end{cases}$$

Spectral Theorem

Spectral theorem

For every self-adjoint matrix $\mathbf{A}$ on a finite dimensional inner product space V_n, there correspond real valued numbers $\alpha_1, \ldots, \alpha_r$, and orthonormal projections $\mathbf{E}_1, \ldots, \mathbf{E}_r$,with $r \leq n$, such that:

- (1) all α_l are *pairwise distinct*
- (2) all $\mathbf{E}_j$ are not null (i.e. $\forall j, \mathbf{E}_j \neq \mathbf{0}$)
- (3) $\sum_{j=1}^{r} \mathbf{E}_j = \mathbf{I}$
- (4) $\mathbf{A} = \sum_{j=1}^{r} \alpha_j \mathbf{E}_j$

Notice that the set of self-adjoint matrices whenever the underlying field is the set of real numbers consists of the set of symmetric matrices. The spectral theorem suggests that a possible basis where to diagonalize them is always available through their eigenvectors.
Applications: document similarity matrices where $a_{ij} = sim(d_i, d_j)$.

Eigen/Diagonal Decomposition

Spectral theorem and non degenerate matrices

The spectral theorem over the set of symmetric matrices imply a special kind of decomposition such that the eigenvectors corresponds to a new (orthogonal) basis and the eigen values are the factors of a transformation able to reconstruct the original matrix.

Eigen/Diagonal Decomposition

Spectral theorem and non degenerate matrices

The spectral theorem over the set of symmetric matrices imply a special kind of decomposition such that the eigenvectors corresponds to a new (orthogonal) basis and the eigen values are the factors of a transformation able to reconstruct the original matrix.

This is called Eigen decomposition of a non-degenerate matrix.

Eigen/Diagonal Decomposition

Spectral theorem and Eigen Decomposition

Let $\mathbf{S}$ be a square matrix with m linearly independent eigenvectors (a '**non-degenerate**' matrix).
Theorem: it exists an eigen decomposition

$$\mathbf{S} = \mathbf{U}\Lambda\mathbf{U}^{-1}$$

such that (cf. matrix diagonalization theorem)

- Columns of $\mathbf{U}$ are eigenvectors of $\mathbf{S}$
- Λ is a diagonal $m \times m$ matrix whose diagonal elements are the m eigenvalues of S, with $\lambda_i \geq \lambda_{i+1}, \qquad \forall i = 1, \ldots, m-1$

Eigen/Diagonal Decomposition: an example

Eigen Decomposition

Given the matrix **U** where the columns correpond to m eigenvectors $v_1, \ldots, v_m$ of **S**, i.e. Let $\mathbf{U} = \begin{pmatrix} \underline{v}_1 & \ldots & \underline{v}_m \end{pmatrix}$ then it follows that:

$$\mathbf{SU} = \mathbf{S}\begin{pmatrix} \underline{v}_1 & \ldots & \underline{v}_m \end{pmatrix} = \begin{pmatrix} \lambda_1 \underline{v}_1 & \ldots & \lambda_m \underline{v}_m \end{pmatrix} =$$

$$\begin{pmatrix} \underline{v}_1 & \ldots & \underline{v}_m \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ldots & \\ & & \lambda_m \end{pmatrix} = \mathbf{U\Lambda}$$

Decomposition

Thus: $\mathbf{SU} = \mathbf{U\Lambda}$ and $\mathbf{S} = \mathbf{U\Lambda U}^{-1}$

Eigen Decomposition

An example

Let $\mathbf{S} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. The eigenvectors are the solutions of $(\mathbf{A} - \lambda \mathbf{I})\underline{x} = \underline{0}$.
Given the first eigenvalue $\lambda_1 = 1$, all vectors of the form $\alpha \underline{v}_1$ are eigenvectors with $\underline{v}_1^T = (1, -1)$. Correspondingly, a second eigenvector $\underline{v}_2 = (1, 1)^T$ corresponds to $\lambda_2 = 3$.
It follows that:

$$\mathbf{U} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \text{ and } \mathbf{U}^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

So that:

$$\mathbf{S} = \mathbf{U}\Lambda\mathbf{U}^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Symmetric Eigen Decomposition

Normalized eigenvectors

Notice that if we take normalized eigenvectors $\underline{v}_1$ and $\underline{v}_2$, i.e. respectively by

$$\underline{v}_1 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \text{ and } \underline{v}_2 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

(with the normalized $\mathbf{U}$ written as $\mathbf{Q}$)

Symmetric Eigen decomposition

Then

$$\mathbf{S} = \mathbf{Q}\Lambda\mathbf{Q}^{-1} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

with $\mathbf{Q}$ *orthogonal*, i.e. such that $\mathbf{Q}^T = \mathbf{Q}^{-1}$ and

$$\mathbf{S} = \mathbf{Q}\Lambda\mathbf{Q}^T$$

Symmetric Eigen Decomposition

Theorem

If $\mathbf{S}$ is a $m \times m$ real-valued symmetric matrix, then it exists a unique **eigen decomposition**

$$\mathbf{S} = \mathbf{Q}\Lambda\mathbf{Q}^T$$

where $\mathbf{Q}$ is orthogonal, i.e.

Symmetric Eigen Decomposition

Theorem

If $\mathbf{S}$ is a $m \times m$ real-valued symmetric matrix, then it exists a unique **eigen decomposition**

$$\mathbf{S} = \mathbf{Q}\Lambda\mathbf{Q}^T$$

where $\mathbf{Q}$ is orthogonal, i.e.

- $\mathbf{Q}^{-1} = \mathbf{Q}^T$
- The columns of $\mathbf{Q}$ are normalized eigenvectors
- All the columns in $\mathbf{Q}$ are orthogonal.

Notice that everything is real.

Symmetric Eigen Decomposition and SVD

Non square matrices

Symmetric Eigen Decomposition and SVD

Non square matrices

When $\mathbf{A}$ is a real-valued $m \times n$ matrix, its co-variance matrix $\mathbf{A}\mathbf{A}^T$ (as well as $\mathbf{A}^T\mathbf{A}$) is a symmetric matrix. By applying the symmetric eigen decomposition to $\mathbf{W} = \mathbf{A}^T\mathbf{A}$ we get:

$$\exists! \ \mathbf{Q} \text{ such that } \mathbf{W} = \mathbf{Q}\Lambda\mathbf{Q}^T$$

where $\lambda_1, \ldots, \lambda_r$ values corresponds to the different eigenvalues of the $\mathbf{W}$ matrix, r is the range and $\mathbf{Q}$ is orthogonal.

Symmetric Eigen Decomposition and SVD

Non square matrices

When $\mathbf{A}$ is a real-valued $m \times n$ matrix, its co-variance matrix $\mathbf{A}\mathbf{A}^T$ (as well as $\mathbf{A}^T\mathbf{A}$) is a symmetric matrix. By applying the symmetric eigen decomposition to $\mathbf{W} = \mathbf{A}^T\mathbf{A}$ we get:

$$\exists! \ \mathbf{Q} \text{ such that } \mathbf{W} = \mathbf{Q}\Lambda\mathbf{Q}^T$$

where $\lambda_1, \ldots, \lambda_r$ values corresponds to the different eigenvalues of the $\mathbf{W}$ matrix, r is the range and $\mathbf{Q}$ is orthogonal.
Notice that: a similar $\mathbf{Q}'$ exists for $\mathbf{A}^T\mathbf{A}$.

Symmetric Eigen Decomposition and SVD

Non square matrices

When $\mathbf{A}$ is a real-valued $m \times n$ matrix, its co-variance matrix $\mathbf{A}\mathbf{A}^T$ (as well as $\mathbf{A}^T\mathbf{A}$) is a symmetric matrix. By applying the symmetric eigen decomposition to $\mathbf{W} = \mathbf{A}^T\mathbf{A}$ we get:

$$\exists!\ \mathbf{Q} \text{ such that } \mathbf{W} = \mathbf{Q}\Lambda\mathbf{Q}^T$$

where $\lambda_1, \ldots, \lambda_r$ values corresponds to the different eigenvalues of the $\mathbf{W}$ matrix, r is the range and $\mathbf{Q}$ is orthogonal.
Notice that: a similar $\mathbf{Q}'$ exists for $\mathbf{A}^T\mathbf{A}$.
We can keep as $\mathbf{U}$ (in SVD) the matrix $\mathbf{Q}$ and as $\mathbf{V}$ (in SVD) the matrix $\mathbf{Q}'$.
They are the **(left and right) normalized eigenvector matrices of $\mathbf{A}\mathbf{A}^T$** and $\mathbf{A}^T\mathbf{A}$, respectively.

Singular Value Decomposition

SVD

For an $m \times n$ matrix $\mathbf{A}$ of rank r there exists a factorization called **Singular Value Decomposition** (SVD) as follows:

$$\mathbf{A} = \mathbf{U}_{(m\times r)}\Sigma_{(r\times r)}\mathbf{V}^T_{(n\times r)}$$

where

- $\mathbf{U}$ expresses the normalized eigevectors of $\mathbf{A}\mathbf{A}^T$
- $\mathbf{V}$ expresses the normalized eigevectors of $\mathbf{A}^T\mathbf{A}$
- Σ is a diagonal matrix whose non-zero elements $\sigma_i \quad (i = 1, \ldots r)$ are called *singular values* and are defined as

$$\sigma_i = \sqrt{\lambda_i}$$

with λ_i as eigenvalues of $\mathbf{A}\mathbf{A}^T$ (and $\mathbf{A}^T\mathbf{A}$)

Singular Value Decomposition

An example with $m = 3$, $n = 2$

Let

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Then the corresponding SVD is:

$\mathbf{A} = \mathbf{U\Sigma V} =$

$$= \begin{pmatrix} 0 & 2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3} \end{pmatrix} \begin{pmatrix} \mathbf{1} & 0 \\ 0 & \sqrt{\mathbf{3}} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

(OSS: The singular values are arranged in decreasing order. Moreover, the diagonal elements of Λ are filled with 0 for the $\sigma_i \quad \forall i > r$

Singular Values and Eigenvectors

Eigenvalues of the Covariance Matrix

In the above example, where

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$$

it is useful to compute the covariance matrix $\mathbf{A}^T\mathbf{A}$, i.e.

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

Singular Values and Eigenvectors

Eigenvalues of the Covariance Matrix

In the above example, where

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$$

it is useful to compute the covariance matrix $\mathbf{A}^T\mathbf{A}$, i.e.

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

whose eigenvalues λ_i are: **3** and **1**.

Covariance Matrix and singular values

Notice how the singular values of $\mathbf{A}$ σ_i are exactly the square roots of eigenvalues for $\mathbf{A}^T\mathbf{A}$, i.e. $\sigma_i = \sqrt{\lambda_i} \qquad i = 1, \ldots, r$

Singular Values and Eigenvectors

Eigenvalues of the Covariance Matrix

In the above example, given the covariance matrix $\mathbf{A}^T\mathbf{A}$, i.e.

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

Singular Values and Eigenvectors

Eigenvalues of the Covariance Matrix

In the above example, given the covariance matrix $\mathbf{A}^T\mathbf{A}$, i.e.

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

and its eigenvalues $\lambda_1 = 3$ and $\lambda_2 = 1$, the corresponding eigenspaces correspond to the rotations and stretching of the original (2) dimensions along the direction of maximal variance.

Eigenspace

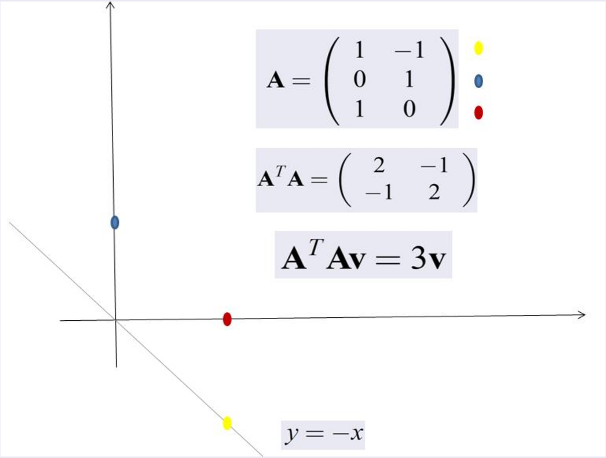
For the eigenvalue $\lambda_1 = 3$, the corresponding direction is the one for which $\mathbf{A}^T\mathbf{A}\mathbf{v} = \lambda_1\mathbf{v}$, i.e. $\mathbf{A}^T\mathbf{A}\mathbf{v} = 3\mathbf{v}$:

$$\begin{cases} 2x_1 - x_2 & = & 3x_1 \\ -x_1 + 2x_2 & = & 3x_2 \end{cases}$$

whose solution $x_2 = -x_1$ corresponds to: $y = -x$.

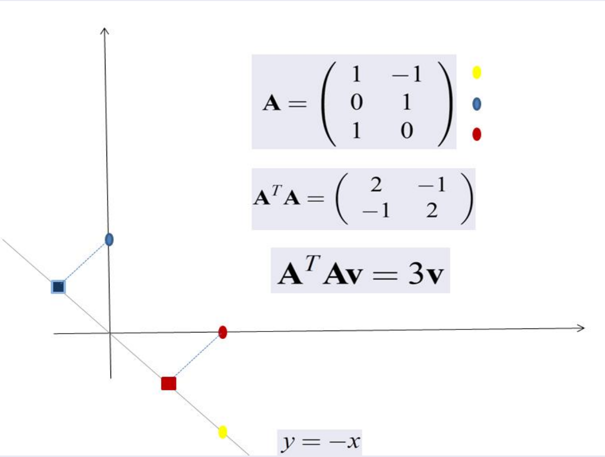
Singular Values and Eigenvectors

The direction of Maximal Covariance



Singular Values and Eigenvectors

Covariance and Dimensionality Reduction



References

Vectors, Operations, Norms and Distances

K. Van Rijesbergen, The Geometry of Information Retrieval, CUP Press, 2004. Chapter 4.

SVD and the fundamental theorem of linear algebra

G. Golub & C. van Loan (1996): Matrix computations. Third edition. London: The Johns Hopkins University Press.