Matrix VII

REA1121

Mathematics for programming

Outline

- Roadmap
- Computation of eigenvalues
- Principal component analysis
- Singular value decomposition
- Exercises

ROADMAP

Roadmap

Vectors

- Linear combination $c\boldsymbol{v} + d\boldsymbol{w}$
- Dot product $\boldsymbol{v} \cdot \boldsymbol{w}$
- Length
- Unit vector
- Angle

Vector space

- Subspace
- Linear independence
- Basis
- Orthonormal basis
- Dimension
- Column space
- Nullspace
- Rank

Orthogonality

- Projection
- Least-squares approximation
- Pseudoinverse

Eigenvalue and eigenvector

- Eigenvalue Eigenvector
- Diagonalization
- Symmetric matrix
- Similar matrix

COMPUTATION OF EIGENVALUES

Computation of eigenvalues

- Determinant
  - $\det(A - \lambda I) = 0$
- Iterative methods for eigenvalues
  - Power method
  - QR algorithm

Power method

1. Start with any vector $\boldsymbol{u}_0$.
2. Multiply by A to find $\boldsymbol{u}_1 = A\boldsymbol{u}_0$.
3. Multiply by A to find $\boldsymbol{u}_2 = A\boldsymbol{u}_1 = A^2\boldsymbol{u}_0$.
4. If $\boldsymbol{u}_0$ is a combination of the eigenvectors, then A multiplies with each eigenvector $\boldsymbol{x}_i$ by λ_i. $\boldsymbol{u}_0 = c_1\boldsymbol{x}_1 + \cdots + c_n\boldsymbol{x}_n$
5. After k steps, we have $(\lambda_i)^k$:

$$\boldsymbol{u}_k = A^k\boldsymbol{u}_0 = c_1(\lambda_1)^k\boldsymbol{x}_1 + \cdots + c_n(\lambda_n)^k\boldsymbol{x}_n$$

Power method

6. As the power methods continues, the largest eigenvalue begins to dominate.
7. The vector $\boldsymbol{u}_k$ point toward that dominant vector.

Example 1

Find the eigenvector corresponding to the largest eigenvalue of A with the power method.

$$A = \begin{bmatrix} 0.9 & 0.3 \\ 0.1 & 0.7 \end{bmatrix}$$

For reference:

Eigenvalue $\lambda_{max} = 1$

Eigenvector $\boldsymbol{x} = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}$

Example 1

Find the eigenvector corresponding to the largest eigenvalue of A with the power method.

Start with $u_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, u_1 = \begin{bmatrix} 0.9 \\ 0.1 \end{bmatrix}, u_2 = \begin{bmatrix} 0.84 \\ 0.16 \end{bmatrix}$

approaching $u_\infty = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}$

See the attached Matlab code. Try with different u_0 and test how fast the largest eigenvalue dominates.

QR algorithm

1. Let $A_0 = A$, a real matrix
2. At the k^{th} step (starting with $k = 0$), we compute the QR decomposition $A_k = Q_k R_k$
   - Q_k is an orthogonal matrix, i.e., $Q^T = Q^{-1}$
   - R_k is an upper triangular matrix.
3. We then form

$$A_{k+1} = R_k Q_k = Q_k^{-1} Q_k R_k Q_k = Q_k^{-1} A_k Q_k$$
$$= Q_k^T A_k Q_k$$

so all the A_k are similar and hence they have the same eigenvalues.

QR algorithm

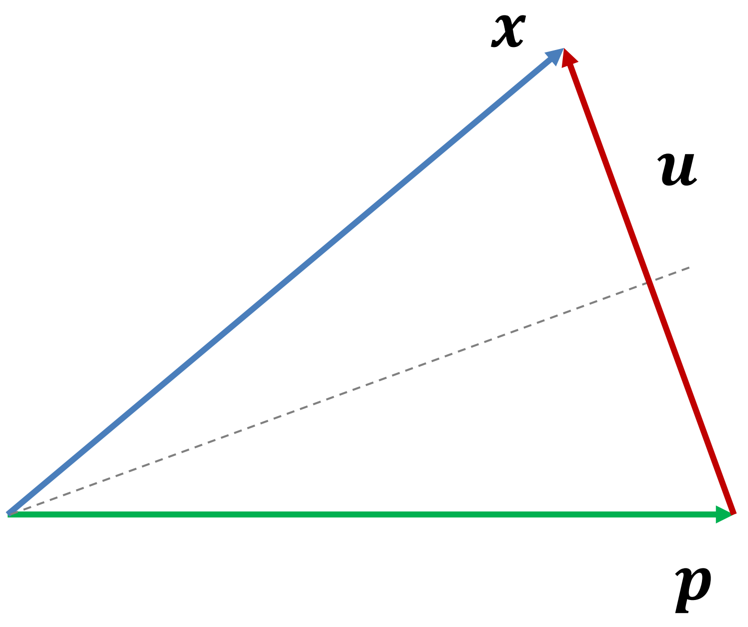
4. Under certain conditions, the matrix A_k converge to a triangular matrix.
5. The eigenvalues of a triangular matrix are listed on the diagonal.

QR decomposition

- A decomposition of a matrix A into a product $A = QR$
  – Q is an orthogonal matrix, i.e., $Q^T = Q^{-1}$
  – R is an upper triangular matrix.
- Computation of the QR decomposition
  – Using Householder reflections

Householder reflection

A Householder reflection is a transformation that takes a vector and reflects it about some plane (or hyperplane).

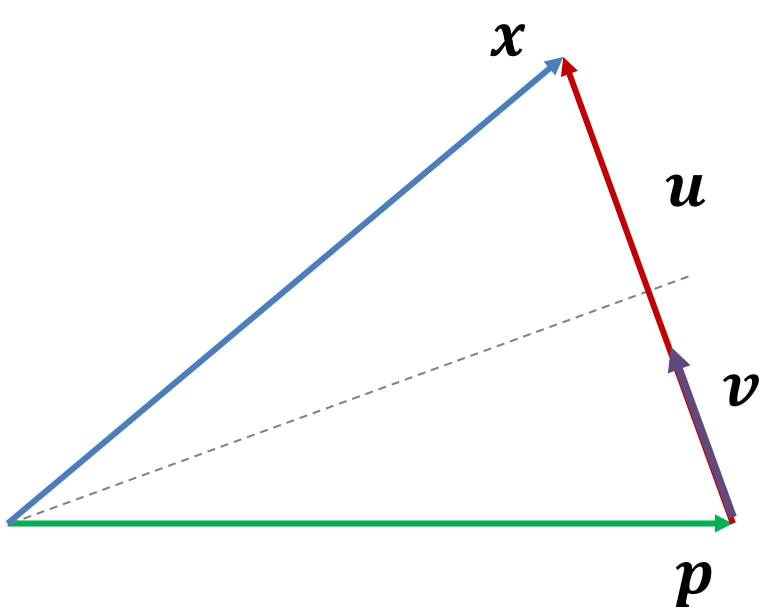


Householder reflection

$$\boldsymbol{v} = \frac{\boldsymbol{u}}{\|\boldsymbol{u}\|}$$

$$\boldsymbol{p} = \boldsymbol{x} - 2(\boldsymbol{x} \cdot \boldsymbol{v})\boldsymbol{v} = \boldsymbol{x} - 2\boldsymbol{v}(\boldsymbol{v}^T \boldsymbol{x})$$

$$P = \boldsymbol{I} - 2\boldsymbol{v}\boldsymbol{v}^T$$



QR decomposition

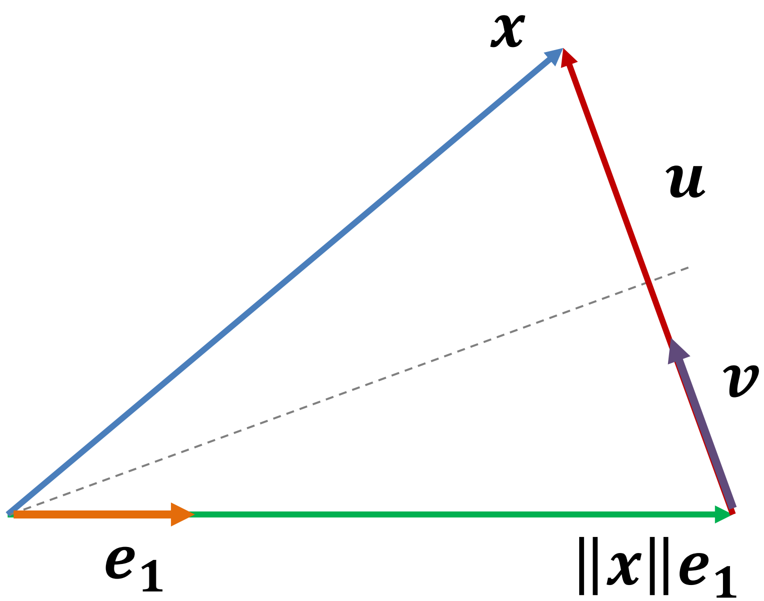
Let $\boldsymbol{x}$ be an arbitrary real m-dimensional column vector of A such that $||\boldsymbol{x}|| = |\alpha|$ for a scalar α.

$\boldsymbol{e}_1 = [1,0,\cdots,0]^T$

$\boldsymbol{u} = \boldsymbol{x} - \alpha\boldsymbol{e}_1$

$\boldsymbol{v} = \dfrac{\boldsymbol{u}}{||\boldsymbol{u}||}$

$Q = \boldsymbol{I} - 2\boldsymbol{v}\boldsymbol{v}^T$

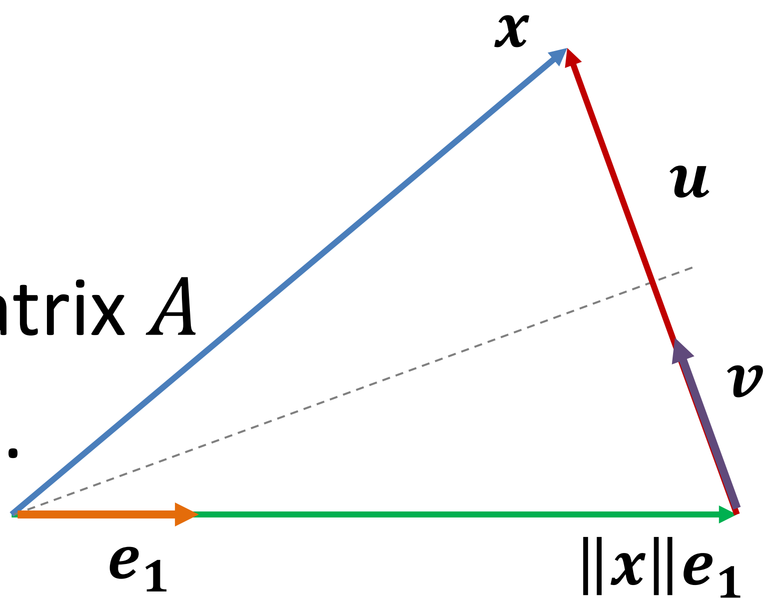


QR decomposition

Q is an m-by-m Householder matrix that is used to reflect $\boldsymbol{x}$ in such a way that all coordinates but one disappear.

$Q\boldsymbol{x} = [\alpha, 0, \cdots, 0]^T$

This is used to gradually transform an m-by-n matrix A to upper triangular form.



QR decomposition

1. Multiply A with the Householder matrix Q_1 we obtain when we choose the first matrix column for $\boldsymbol{x}$. This results in a matrix $Q_1 A$ with zeros in the left column (except for the first row).

$$Q_1 A = \begin{bmatrix} \alpha_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{bmatrix}$$

QR decomposition

2. Repeat step 1 on A' obtained from $Q_1 A$ by deleting the first row and first column, which results in a Householder matrix Q_2'.

$$Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & Q_2' \end{bmatrix}, Q_k = \begin{bmatrix} I_{k-1} & 0 \\ 0 & Q_k' \end{bmatrix}$$

$$Q_2 Q_1 A = \begin{bmatrix} \alpha_1 & * & \cdots * \\ 0 & \alpha_2 & *\cdot* \\ \vdots & \vdots & A' \\ 0 & 0 & \end{bmatrix}$$

QR decomposition

3. After t iterations of this process, $t = min(m-1, n)$

$$R = Q_t \cdots Q_2 Q_1 A = \begin{bmatrix} \alpha_1 & * & \cdots & * \\ 0 & \alpha_1 & * & * \\ \vdots & 0 & \ddots & * \\ \vdots & \vdots & 0 & \alpha_t \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

is a upper triangular matrix.

QR decomposition

4. With $Q = Q_1^T Q_2^T \cdots Q_t^T$,

$$A = QR = Q_1^T Q_2^T \cdots Q_t^T \, Q_t \cdots Q_2 Q_1 A$$

is a *QR decomposition* of A.

Example 2

Calculate the eigenvalues of A with the QR algorithm. (See attached Matlab code.)

$$A = \begin{bmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{bmatrix}$$

Demonstration

- Graphical demonstration of $\boldsymbol{x}$ and $A\boldsymbol{x}$
- Observe the relationship between $\boldsymbol{x}$ and $A\boldsymbol{x}$
- See the attached Matlab code (eigshow.m)
- Make $A\boldsymbol{x}$ parallel to $\boldsymbol{x}$

PRINCIPAL COMPONENT ANALYSIS

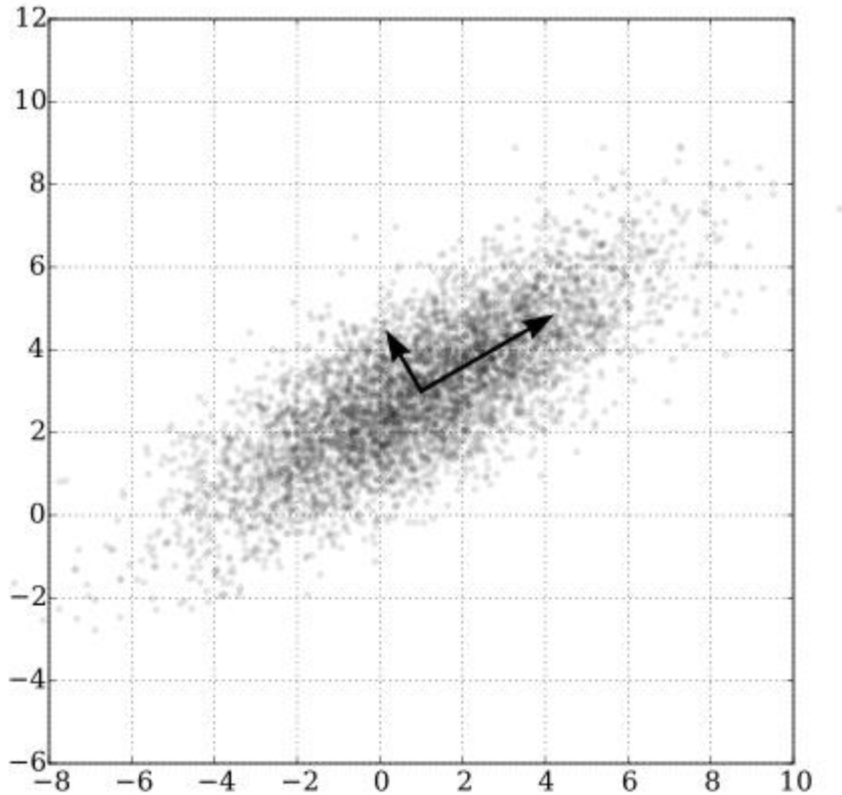
Principal component analysis

- PCA is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called ***principal components***.

Principal component analysis

- This transformation is defined in such a way
  - the first principal component has the largest possible variance
  - each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components.
- The resulting vectors are an uncorrelated orthogonal basis set.

Principal component analysis



Example 3

Computation of PCA

Consider the vectors x_1, x_2, x_3, each of them is a single grouped observation of 2 variables.

$$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \boldsymbol{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \boldsymbol{x}_3 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Example 3

1. Organize the data set so that each row vector of X contains the $p = 2$ variables and each column vector represents $n = 3$ measurements of a single variable.

$$X = \begin{bmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \boldsymbol{x}_3' \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Example 3

2. Calculate the empirical mean for each column vector of X, namely each variable in X

$$E(X) = E\left(\begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

Example 3

3. Calculate the deviations from the mean

$$B = X - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} E(X) = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

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

Example 3

4. Find the symmetric square matrix $C = B^T B$

$$C = \frac{1}{n-1} B^T B = \frac{1}{2} \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}$$

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

Example 3

5. Find the eigenvectors and eigenvalues of the symmetric square matrix C

$$CQ = Q\Lambda$$

$$Q = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

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

Example 3

6. Find the principal components P

$$P = XQ = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ -1 & -1 \end{bmatrix}$$

Example 3

7. Restore the matrix of source data X

$$PQ^T = XQQ^T = XQQ^{-1} = X$$
$$= \frac{1}{2}\begin{bmatrix} 2 & 0 \\ -1 & 1 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}$$

An application of PCA

Apply PCA to recommendation systems (see the attached paper for details)

1. Data formation
2. Covariance Matrix Calculation
3. Calculation of Eigvenvectors/Eigenvalues
4. Formation of a feature vector
5. Construction of the new data set
6. Return to the old data

SINGULAR VALUE DECOMPOSITION

Diagonalization with eigenvectors

$$A\boldsymbol{x} = \lambda\boldsymbol{x}$$
$$AS = S\Lambda$$
$$S^{-1}AS = \Lambda$$
$$A = S\Lambda S^{-1}$$

The eigenvectors in S:

- they are usually not orthogonal
- there are not always enough eigenvectors
- A is required to be square

Diagonalization with singular vectors

$$A\boldsymbol{v} = \sigma \boldsymbol{u}$$

- Two sets of singular vectors:
  - $\boldsymbol{u}$'s are eigenvectors of AA^T
  - $\boldsymbol{v}$'s are eigenvectors of A^TA
- AA^T and A^TA are symmetric matrices, $\boldsymbol{u}$'s and $\boldsymbol{v}$'s can be chosen orthonormal.

Diagonalization with singular vectors

$$A\boldsymbol{v} = \sigma\boldsymbol{u}$$

- Singular vectors $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_r$ are in the row space of A.
- The outputs $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_r$ are in the column space of A.
- The singular values $\sigma_1, \cdots, \sigma_r$ are all positive numbers.

Diagonalization with singular vectors

$$A\boldsymbol{v} = \sigma\boldsymbol{u}$$
$$AV = U\Sigma$$

$$A\begin{bmatrix} v_1 & \cdots & v_r \end{bmatrix} = \begin{bmatrix} u_1 & \cdots & u_r \end{bmatrix}\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$$

A	V	U	Σ
$m \times n$	$n \times r$	$m \times r$	$r \times r$

Diagonalization with singular vectors

$$A\boldsymbol{v} = \sigma\boldsymbol{u}$$

$$AV = U\Sigma$$

$$A\begin{bmatrix} v_1 & \cdots & v_r & \cdots & v_n \end{bmatrix} = \begin{bmatrix} u_1 & \cdots & u_r & \cdots & u_m \end{bmatrix}\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix}$$

A	V	U	Σ
$m \times n$	$n \times n$	$m \times m$	$m \times n$

Singular vector decomposition

$$A\boldsymbol{v} = \sigma\boldsymbol{u}$$
$$AV = U\Sigma$$

A	V	U	Σ
$m \times n$	$n \times n$	$m \times m$	$m \times n$

V is a square orthogonal matrix, with $V^{-1} = V^T$

$$A = U\Sigma V^{-1} = U\Sigma V^T$$
$$A = \boldsymbol{u}_1\sigma_1\boldsymbol{v}_1^T + \cdots + \boldsymbol{u}_r\sigma_r\boldsymbol{v}_r^T$$

Comparison

Eigendecomposition

$$A\boldsymbol{x} = \lambda\boldsymbol{x}$$
$$A = S\Lambda$$
$$A = S\Lambda S^{-1}$$

Singular value decomposition

$$A\boldsymbol{v} = \sigma\boldsymbol{u}$$
$$AV = U\Sigma$$
$$A = U\Sigma V^{-1} = U\Sigma V^{T}$$

SVD

$$AV = U\Sigma$$

$$A\begin{bmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2\end{bmatrix} = \begin{bmatrix}\sigma_1\boldsymbol{u}_1 & \sigma_2\boldsymbol{u}_2\end{bmatrix} = \begin{bmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2\end{bmatrix}\begin{bmatrix}\sigma_1 & \\ & \sigma_2\end{bmatrix}$$

$$A = U\Sigma V^T$$

$$A^T A = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^T U^T U\Sigma V^T$$

$$= V\Sigma^T\Sigma V^T = V\begin{bmatrix}\sigma_1^2 & \\ & \sigma_2^2\end{bmatrix}V^T$$

SVD

$$A^T A = (U\Sigma V^T)^T (U\Sigma V^T) = V\Sigma^T U^T U\Sigma V^T$$
$$= V\Sigma^T \Sigma V^T = V \begin{bmatrix} \sigma_1^2 & \\ & \sigma_2^2 \end{bmatrix} V^T$$

For symmetric matrices, we have

$$A = Q\Lambda Q^{-1} = Q\Lambda Q^T$$

with orthonormal eigenvectors in $S = Q$ and $Q^{-1} = Q^T$

SVD

$$A^T A = (U\Sigma V^T)^T (U\Sigma V^T) = V\Sigma^T U^T U\Sigma V^T$$
$$= V\Sigma^T \Sigma V^T = V \begin{bmatrix} \sigma_1^2 & \\ & \sigma_2^2 \end{bmatrix} V^T$$

- The columns of V are the eigenvectors of $A^T A$
- The diagonal elements of Σ are the eigenvalues of $A^T A$

Example 4

Find the singular value decomposition of the matrix $A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$.

Example 4

Find the singular value decomposition of the matrix $A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$.

Solution Compute $A^T A$ and its eigenvectors. Then make them unit vectors:

$A^T A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$ has unit eigenvectors

$\boldsymbol{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ and $\boldsymbol{v}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

Example 4

The eigenvalues of $A^T A$ are 8 and 2. The $\boldsymbol{v}$'s are perpendicular, because eigenvectors of every symmetric matrix are perpendicular – and $A^T A$ is automatically symmetric.

$A\boldsymbol{v}_1 = \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix}$. The unit vector is $\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \sigma_1 = 2\sqrt{2}$

$A\boldsymbol{v}_2 = \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix}$. The unit vector is $\boldsymbol{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \sigma_2 = \sqrt{2}$

Example 4

$$A = U\Sigma V^T$$

$$\begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Demonstration

- Graphical demonstration of orthogonal vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ and $A\boldsymbol{x}$ and $A\boldsymbol{y}$
- Observe the relationship between $A\boldsymbol{x}$ and $A\boldsymbol{y}$
- See the attached Matlab code (eigshow.m)
- Make $A\boldsymbol{x}$ perpendicular to $A\boldsymbol{y}$

An application of SVD

SVD for image compression (see attached Matlab code)

A grayscale image A can be factorized with SVD:

$$A = U\Sigma V^T$$

A	U	Σ	V
$m \times n$	$m \times m$	$m \times n$	$n \times n$

Reducing the number of singular vectors compresses the image.

An application of SVD

SVD for image compression (see attached Matlab code and try different r)

A grayscale image A can be factorized with SVD:

$$A = U\Sigma V^T$$

Original:	A	U	Σ	V
	$m \times n$	$m \times m$	$m \times n$	$n \times n$
After	A	U	Σ	V
reduction:	$m \times n$	$m \times r$	$r \times r$	$n \times r$

EXERCISES

Problem 1

Rectangular matrix $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$

- Compute PCA and find the vectors and the corresponding principal components
- Compute SVD of A and find the singular vectors and the singular values.
- Verify your results with the attached Matlab code.

Solution 1

PCA

$$P = XQ = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 0 & 2 \\ -1 & 1 \end{bmatrix}$$

$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

Solution 1

SVD

$$A = U\Sigma V^{T}$$

$$U = \begin{bmatrix} -0.4082 & 0.7071 & 0.5774 \\ -0.8165 & 0 & -0.5774 \\ -0.4082 & -0.7071 & 0.5774 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 1.7321 & \\ & 1 \\ & \end{bmatrix}$$

$$V = \begin{bmatrix} -0.7071 & 0.7071 \\ -0.7071 & -0.7071 \end{bmatrix}$$