MATHEMATICS FOR COMPUTER VISION  WEEK 3
EIGENDECOMPOSITION AND SVD

Dr Fabio Cuzzolin
MSc in Computer Vision
Oxford Brookes University
Year 2013-14
OUTLINE OF WEEK 3

- Topics: eigendecomposition, eigenvectors, Singular Value Decomposition

- Eigenvectors and eigenvalues
  - Eigenspaces

- Eigendecomposition of a matrix
  - Characteristic polynomial and multiplicity
  - Diagonalisation and Jordan form

- Numerical computation

- Singular Value Decomposition
  - Interpretations and geometry
  - Relation to eigenvalue decomposition

- Applications, PCA
GENERAL THEORY

Week 3 – Eigendecomposition and SVD
EIGENVECTORS AND EIGENVALUES

- eigenvector of a square matrix: a nonzero vector \( \mathbf{v} \) such that

\[
A\mathbf{v} = \lambda \mathbf{v}
\]

for some scalar value \( \lambda \)

- this is called the eigenvalue equation

- the scalar \( \lambda \) is called the eigenvalue of \( A \) corresponding to the eigenvector \( \mathbf{v} \)

- definition can be extended to any operator other than a matrix
GEOMETRIC INTERPRETATION

- in general, when you multiply a vector by a matrix, the resulting vector $A\nu = \nu'$ is not parallel to the original $\nu$
- eigenvectors of $A$ are exactly those vectors that remain parallel, after the linear mapping
- trivial case: every vector is an eigenvector of the identity matrix $I$!
- (for it maps each vector to itself)
NUMERICAL EXAMPLE

- consider the transformation matrix \( A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \),
- the vector \( \mathbf{v} = [4, -4]^T \) is an eigenvector of \( A \) with eigenvalue \( \lambda = 2 \)

\[
Av = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ -4 \end{bmatrix} = \begin{bmatrix} 3 \cdot 4 + 1 \cdot (-4) \\ 1 \cdot 4 + 3 \cdot (-4) \end{bmatrix}
\]

- the vector \( \mathbf{v} = [0, 1]^T \), instead, is not an eigenvector of \( A \)

\[
\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \cdot 0 + 1 \cdot 1 \\ 1 \cdot 0 + 3 \cdot 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}
\]
FROM EIGENVALUES TO EIGENSPACES

- if \( \mathbf{v} \) is an eigenvector, any scalar multiple of \( \mathbf{v} \) is also an eigenvector
- also, if \( \mathbf{u} \) and \( \mathbf{v} \) are both eigenvectors with the same eigenvalue \( \lambda \), then \( \mathbf{u} + \mathbf{v} \) is
- there is a vector space of eigenvectors with the same eigenvalue: eigenspace
- dimension of this eigenspace \( \rightarrow \) multiplicity of \( \lambda \)
- a linear transform (e.g., a matrix) acting on an \( n \)-dimensional space has at most \( n \) distinct eigenvalues
- list of eigenvalues \( \rightarrow \) spectrum
- these definitions can be generalised to infinite-dimensional spaces, e.g. spaces of functions (Week 4)
- we are more interested in the spectra of matrices
MATRX EIGENDECOMPOSITION

Week 3 – Eigendecomposition and SVD
CHARACTERISTIC POLYNOMIAL

- what happens if the operator is a matrix $A$?
- any eigenvector is the solution of the equation
  \[ Av - \lambda v = 0 \]
  or, equivalently, $(A - \lambda I)v = 0$
- this has nonzero solutions iff the determinant is zero
  \[ \det(A - \lambda I) = 0 \]
- the left side is a polynomial function of $\lambda$: the characteristic polynomial
- since this is a polynomial equation in powers of $\lambda$ up to $n$ (size of $A$), again there are at most $n$ eigenvalues
EXAMPLE

- let A be the matrix

\[
A = \begin{bmatrix}
2 & 0 & 0 \\
0 & 3 & 4 \\
0 & 4 & 9
\end{bmatrix}
\]

- the characteristic polynomial of A is

\[
\det(A - \lambda I) = \det \left( \begin{bmatrix}
2 & 0 & 0 \\
0 & 3 & 4 \\
0 & 4 & 9
\end{bmatrix} - \lambda \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix} \right) = \det \begin{bmatrix}
2 - \lambda & 0 & 0 \\
0 & 3 - \lambda & 4 \\
0 & 4 & 9 - \lambda
\end{bmatrix}
\]

- .. which read as

\[
(2 - \lambda)(3 - \lambda)(9 - \lambda) - 16 = -\lambda^3 + 14\lambda^2 - 35\lambda + 22
\]

- solutions (roots) are \( \lambda = 1 \), \( \lambda = 2 \) and \( \lambda = 11 \) corresponding to eigenvectors

\[
[0,2,-1]', [1,0,0]', [0,1,2]'
\]
REAL AND COMPLEX EIGENVALUES

- fundamental theorem of algebra: any polynomial of degree $n$ has exactly $n$ complex roots
- complex number $r + ic$ where $r$ is the real component, $c$ is the imaginary component, and $i$ is the imaginary unit
- can be represented as points of a plane: real number are on the horizontal axis
- so, even when the matrix has all real entries, its eigenvectors are in general complex numbers
- the complex ones are conjugate, e.g. $r + ic$, $r - ic$
ALGEBRAIC AND GEOMETRIC MULTIPLICITY

- **algebraic multiplicity** $\mu(\lambda)$ of an eigenvalue $\lambda$ of $A$: its multiplicity as a root of the characteristic polynomial
- in general **different** from the **geometric** multiplicity $\gamma(\lambda)$ (the dimension of the associated eigenspace)
- indeed, $\gamma(\lambda) \leq \mu(\lambda)$
- example:

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

- characteristic polynomial:

$$\det(A - \lambda I) = \det \begin{bmatrix} 2 - \lambda & 0 & 0 & 0 \\ 1 & 2 - \lambda & 0 & 0 \\ 0 & 1 & 3 - \lambda & 0 \\ 0 & 0 & 1 & 3 - \lambda \end{bmatrix} = (2 - \lambda)^2(3 - \lambda)^2$$

- the algebraic multiplicity of $\lambda = 2$ is 2, but its eigenspace is spanned by the single vector $[0,1,-1,1]'$
- hence, its geometric multiplicity is equal to 1
DIAGONALIZATION AND JORDAN FORM

- when the sum of the algebraic mult. of all eigenvalues is $n$, i.e. $\sum_{\lambda} \mu(\lambda) = n$ then $A$ has a set of $n$ linearly independent eigenvectors

- therefore it can be **diagonalised** by doing

  \[ Q^{-1}AQ = \Lambda \]

  where $Q$ has the eigenvectors as columns, and $\Lambda$ is the diagonal matrix which has in $\Lambda_{ii}$ the eigenvalue assoc. with the i-th column of $Q$

- if $A$ is diagonalizable, $\mathbb{R}^n$ can be decomposed into the (direct) sum of the eigenspaces of $A$ (see SVD later)
JORDAN NORMAL FORM

- for non-diagonalizable matrices, eigenvectors are replaced by \textit{generalised eigenvectors}, and $\Lambda$ by its \textit{Jordan form}
- “pseudo”-diagonal form
- has the following structure

$$
\begin{pmatrix}
\lambda_1 & 1 \\
& \lambda_1 & 1 \\
& & \lambda_1 \\
& & & \lambda_2 \\
& & & & \lambda_2 \\
& & & & & \lambda_3 \\
& & & \ldots & & \ldots \\
& & & & & \lambda_3 \\
& & & & & & \lambda_n \\
& & & & & & & \lambda_n
\end{pmatrix}
$$
FURTHER PROPERTIES

- eigenvalues and vectors are related to the main quantities associated with a matrix
- the trace of $A$ is the sum of all its eigenvalues
  \[ tr(A) = \sum_{i=1}^{n} \lambda_i \]
- the determinant of $A$ is the product of all its eigenvalues
  \[ det(A) = \prod_{i=1}^{n} \lambda_i \]
- the matrix is invertible iff all eigenvalues are nonzero
- the eigenvalues of the inverse are the $1/\lambda_i$
NUMERICAL COMPUTATION

- eigenvalues are the solutions of the characteristic equation
- no explicit algebraic formula exists for n>4
- must be computed by approximated numerical methods
- once they are obtained, eigenvectors are found by solving the original eigenvalue equation $A\mathbf{v} = \lambda \mathbf{v}$
- but this is just a system of linear equations, with known coefficients
A COOL APPLICATION OF EIGENVECTORS!

- The **Schroedinger equation** of quantum mechanics is an eigenvalue equation:
  \[ H\psi = E\psi \]
  where \( H \) is an operator (Hamiltonian), \( \psi \) is the wave function that described the state of a particle, the eigenvalue \( E \) is the energy.
SINGULAR VALUE DECOMPOSITION

Week 3 – Eigendecomposition and SVD

\[ A = U D V^T \]
BRIEF HISTORY

- discovered by differential geometers
  - Bertrami and Jordan, 1873-4
  - Sylvester in 1889
- first proof: Eckart and Young
- practical methods for computing SVD: Golub

19
SINGULAR VALUE DECOMPOSITION

- The singular value decomposition of an $m \times n$ real or complex matrix $M$ is a factorization of the form
  \[ M = U \Sigma V^* \]

- $U$ is a $m \times m$ unitary matrix
- $\Sigma$ is a $m \times n$ rectangular diagonal matrix with non-negative real numbers on the diagonal
- $V^*$ (the conjugate transpose of $V$) is unitary
- Diagonal entries of $\Sigma$ are singular values of $M$
- Singular values are normally listed in descending order
SINGULAR VALUE DECOMPOSITION

- graphical representation of SVD
INTERPRETATIONS

- clearly, it is closely related to eigendecomposition:
  - the columns of $U$ are the eigenvectors of $M M^*$
  - the columns of $V$ are the eigenvectors of $M^* M$
  - the (nonzero) singular values are the square roots of the nonzero eigenvalues of both $M M^*$ and $M^* M$

- since both $U$ and $V^*$ are unitary, their columns form a set of orthonormal vectors
- can be regarded as basis vectors of a vector space
- SVD is not unique (see example)
- has a geometric interpretation in terms of axes of ellipsoids (see PCA in Machine Learning module)
EXAMPLE

- example of Singular Value Decomposition

\[ M = \begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \end{bmatrix} \]

- not unique: this is also ok

\[ U = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \]

\[ \Sigma = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & \sqrt{5} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \]

\[ V^* = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \sqrt{0.2} & 0 & 0 & \sqrt{0.8} & 0 \\ \sqrt{0.4} & 0 & \sqrt{0.5} & 0 & -\sqrt{0.1} \\ -\sqrt{0.4} & 0 & \sqrt{0.5} & \sqrt{0.1} & 0 \end{bmatrix} \]
COMPUTING THE SVD

- a number of different implementations
- first step: reduce $M$ to a bi-diagonal matrix
- second step: computing the SVD of the resulting matrix
- this is done by iterative methods (as in eigenvalue calculation)
  - QR algorithm for computing eigenvalues
  - Jacobi algorithm
- already implemented in Matlab
APPLICATIONS OF SVD

Week 3 – Eigendecomposition and SVD
Applications

- can be used to compute the pseudo-inverse of a matrix
  - pseudo-inverse of $M = U \Sigma V^T$ is $M^* = V \Sigma^* U^T$ where $\Sigma^*$ is obtained by replacing each diagonal entry by its reciprocal, and transposing

- solving homogenous linear equations

- low-rank matrix approximation

- inverse problems in mathematics
  - given the data, infer the parameters of the model describing them; describe most vision problems, in fact

- main applications of interest to us:
  1. Least square optimisation
  2. Principal Component Analysis

- only outlined here, developed later
LEAST SQUARE OPTIMISATION (OUTLINE)

- Standard way of finding approximate solutions for systems that are overdetermined
- too many equations, impossible to satisfy them all
- each equation is met with an error
- “least squares” means minimising the sum of the squares of the errors for all equations
  - typical application: regression,
  - e.g. fitting a curve to data
- solution of $\min \|Ax\|_2$ subject to $\|x\| = 1$ is the singular vector of the smallest singular value
PRINCIPAL COMPONENT ANALYSIS (OUTLINE)

- given a collection of data points $X$
- looks for an orthogonal transformation of the data such that the first dimension (first principal component) has the greatest variance, followed by the second, etc
- mathematically, the PC decomposition is given by $T = XW$ where $W$ is a matrix whose columns are the eigenvectors of $X^TX$, the principal components
- $X^TX$ can be interpreted as the covariance matrix of the data $X$
- relation to SVD:
  - the singular vectors $W$ of $X$ are the eigenvectors of $X^TX$, its singular values are the square roots of the eigenvalues of $X^TX$
  - the score matrix is therefore $T = XW = U \Sigma$
  - computing the SVD is now the standard way of doing PCA
SUMMARY

Week 3 – Eigendecomposition and SVD
SUMMARY OF WEEK 3

- General theory of eigenvectors and eigenvalues, eigenspaces
- Eigendecomposition of matrices
  - Characteristic polynomial, algebraic multiplicity
  - Real and complex eigenvalues
  - Diagonalisation
  - Jordan normal form
- Singular Value Decomposition
- Applications of SVD
  - Least square optimisation
  - Principal Component Analysis