KTH ROYAL INSTITUTE OF TECHNOLOGY

Lecture 2: Eigenvalues, eigenvectors and similarity

Magnus Jansson/Emil Björnson, March 28 2018

"The single most important concept in matrix theory."

German word "eigen" means *proper* or *characteristic*.

Definition

Consider: Square matrix $A \in M_n$
If there exists $x \in \mathbf{C}^n$ ($x \neq 0$) and $\lambda \in \mathbf{C}$ such that

$$Ax = \lambda x$$

then

- λ is an eigenvalue of A
- x is an eigenvector of A associated with λ

More terminology:

- Spectrum of A: Set of all eigenvalues. Notation: $\sigma(A)$.
- Spectral radius of A: $\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$.

Example: Stability of linear systems

Consider a linear discrete time homogenous system:

$$x(n+1) = Ax(n)$$

If the spectral radius of A is greater than 1, and $x(n)$ is not orthogonal to the corresponding eigenvector(s), $x(n+1)$ will grow in the direction of the unstable modes.

Example: Filter output power maximization

Let $y(t)$ be a vector of observations at time t, and let $z(t) = x^T y(t)$ be the output of a filter with filter coefficients in the vector x. The mean output power is

$$\begin{aligned} \mathrm{E}\{z^2(t)\} &= \mathrm{E}\{\left(x^T y(t)\right)\left(y^T(t)x\right)\} \\ &= x^T \mathrm{E}\{y(t)y^T(t)\}x = x^T R_y x \end{aligned}$$

Assume we want to find the filter that maximizes the output power under a normalizing constraint on the filter gain:

$$\max_x \; x^T R_y x \quad \text{subject to } x^T x = 1$$

The solution is given by the unit length eigenvector corresponding to the largest eigenvalue of R_y.

What about zero?

Recall: If exists $x \in \mathbf{C}^n$ ($x \neq 0$) and $\lambda \in \mathbf{C}$ such that

$$Ax = \lambda x$$

then

- λ is an eigenvalue of A
- x is an eigenvector of A associated with λ

Question: Why $x \neq 0$?
Answer: We always have $A0 = \lambda 0$ (uninteresting solution!)

However: $\lambda = 0$ is an important case: $Ax = 0 = 0x$
In fact: $A \in M_n$ is singular iff it exists $x \neq 0$ such that $Ax = 0x = 0$ or, equivalently, iff $0 \in \sigma(A)$

Polynomials of matrices

Consider: Scalar polynomial
$p(t) = a_k t^k + a_{k-1} t^{k-1} + \ldots + a_0$ with $a_i \in \mathbf{C}$.
Define: Matrix polynomial
$p(A) = a_k A^k + a_{k-1} A^{k-1} + \ldots + a_0 I$ for $A \in M_n$. $A^0 = I$ by convention.

Theorem: If λ is an eigenvalue of A and x the associated eigenvector, then

$$p(A)x = p(\lambda)x.$$

Thus, x is also eigenvector of $p(A)$ associated with eigenvalue $p(\lambda)$.

How to find an eigenvalue?

Rewrite eigenvalue definition ($A \in M_n$):

$$\lambda x - Ax = 0 \quad \Leftrightarrow \quad (\lambda I - A)x = 0$$

Observation: Eigenvalues make $(\lambda I - A)$ singular, $\det(\lambda I - A) = 0$.

Definition: Characteristic polynomial is $p_A(t) = \det(tI - A)$.
$p_A(t)$ is polynomial of degree n: Has n solutions/roots to $p_A(t) = 0$.

Conclusions: These roots are the eigenvalues of A.
$A \in M_n$ has n (complex) eigenvalues.
Some eigenvalues may have (algebraic) multiplicity!

Block triangular matrices

If $A \in M_n$ is block upper triangular

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

then $p_A(t) = p_{A_{11}}(t) p_{A_{22}}(t)$. Proof?

Hence, the eigenvalues of A are those of A_{11} together with those of A_{22}.

Special cases: Triangular matrices, diagonal matrices.

Trace and determinant

Definition: Trace is $\mathrm{tr}(A) = \sum_{i=1}^n a_{ii}$
Definition: Determinant is $\det(A)$ = [Laplace expansion in 0.3.1]

Theorem: Expressed using eigenvalues as

$$\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i, \qquad \det(A) = \prod_{i=1}^{n} \lambda_i.$$

Observation: Coefficients in characteristic polynomial

$$p_A(t) = t^n + a_{n-1}t^{n-1} + \ldots + a_0$$

where $a_{n-1} = -\,\mathrm{tr}(A)$, $a_0 = (-1)^n \det(A)$.

Formulas exist for all a_k; see the book.

Similarity

Consider: $A \in M_n$, $B \in M_n$

Definition: B is similar to A if there exists a nonsingular $S \in M_n$ such that

$$A = S^{-1}BS$$

Notation: $B \sim A$

The transformation $A \to S^{-1}AS$ is called a *similarity transformation* by the similarity matrix S.

Equivalence class

Similarity is

- **reflexive** $A \sim A$
- **symmetric** $B \sim A$ implies $A \sim B$
- **transitive** $C \sim B$ and $B \sim A$ imply $C \sim A$

Divides all matrices into (disjoint) equivalence classes:

- Each class has a representative matrix A.
- The class includes all matrices similar to A.

Theorem: If $B \sim A$, then $p_B(t) = p_A(t)$.
B and A have the same eigenvalues (counting multiplicity).

Diagonalizable matrices

Definition:
$A \in M_n$ is diagonalizable if it is similar to a diagonal matrix.

Theorem:
A is diagonalizable iff A has n linearly independent eigenvectors.

Linearly independent eigenvectors

Assume:
$\lambda_1, \lambda_2, \ldots, \lambda_n$ are *distinct* eigenvalues of $A \in M_n$
x_i is eigenvector associated with λ_i

Theorem: $\{x_1, x_2, \ldots, x_n\}$ is a linearly independent set.

Conclusion: If $A \in M_n$ has n distinct eigenvalues, then A is diagonalizable.
(The converse is not true.)

Simultaneous diagonalization

Definition: $A, B \in M_n$ commute if $AB = BA$

Definition: Two diagonalizable matrices $A, B \in M_n$ are *simultaneously diagonalizable* if there exists a single similarity matrix $S \in M_n$ diagonalizing both A and B.

Theorem: A, B commute iff they are simultaneously diagonalizable.

Eigenvalues of products

Assume: $A \in M_{m,n}$ and $B \in M_{n,m}$ with $m \leq n$.

Theorem: $p_{BA}(t) = t^{n-m} p_{AB}(t)$

BA has the same eigenvalues as AB plus $n - m$ additional eigenvalues at zero.

If $m = n$ and A (or B) is nonsingular, then AB is similar to BA.

How to find the eigenvector?

Rewrite eigenvalue definition ($A \in M_n$):

$$\lambda x - Ax = 0 \quad \Leftrightarrow \quad (\lambda I - A)x = 0$$

Observation: x lies in the nullspace of $\lambda I - A$.

Calculate eigenvectors: Solve $(\lambda I - A)x = 0$ for eigenvalue λ.
System of equations: Use Gauss elimination.

Eigenvector is non-unique:

- Any scaling ($x \neq 0$)
- Nullspace can have large dimension.

Eigenspace

The set of all eigenvectors satisfying $Ax = \lambda x$ for a given $\lambda \in \sigma(A)$ is called the eigenspace of A corresponding to λ.

The eigenspace, together with the zero vector, is a subspace of $\mathbf{C}^n$ and it is exactly the nullspace of $\lambda I - A$.

Multiplicity

Algebraic multiplicity: Multiplicity of the corresponding root of the characteristic polynomial.
Geometric multiplicity: Number of linearly independent eigenvectors associated with the eigenvalue.

Theorem: Algebraic multiplicity $\geq$ Geometric multiplicity

Definition:
If strict inequality for some eigenvalue, the matrix is *defective*.

Theorem: A is diagonalizable iff it is not defective.

Left eigenvectors

A nonzero vector $y \in \mathbf{C}^n$ is a left eigenvector of $A \in M_n$ if

$$y^* A = \mu y^*.$$

Observe that $\mu \in \sigma(A)$.

Theorem (Biorthogonality): Let

$$y^* A = \mu y^* \quad \text{and} \quad Ax = \lambda x.$$

Then, if $\mu \neq \lambda$ we have $y^* x = 0$.

Observation:
If A is Hermitian ($A = A^*$), then $x = y$ for same eigenvalue. Biorthogonality then implies that A has n pair-wise orthogonal eigenvectors of (at least if eigenvalues are distinct, more later).

Transpose and conjugate transpose

Transpose:

- A and A^T have same eigenvalues.
- Left/right eigenvectors are interchanged and complex conjugated.

Conjugate transpose:

- Eigenvalues of A^* are complex conjugates of eigenvalues of A.
- Left/right eigenvectors are interchanged.