EQ2820 / EM3220 Matrix Algebra

An accelerated program course / PhD course

Course organizer: Magnus Jansson (magnus.jansson@ee.kth.se, 08-790 8443)

Lectures by: Magnus Jansson, and Mats Bengtsson

Teaching Assistant: Efthymios Tsakonas

Course homepages:

`https://www.kth.se/social/course/EQ2820/`

`https://www.kth.se/social/group/em3220-matrix-algebr/`

Outline

- Organization
- Schedule
- Contents
- Chapter 0: Review and Miscellanea

Course organization

- Main course literature: "Matrix Analysis" by R.A. Horn and C. R. Johnson.
  We will also use a few chapters from "Topics in Matrix analysis" by the same authors (+ additional material in the form of lecture slides).
- Format: Lectures and homework on a weekly basis. Course contents can be learnt by cooperative discussions, but homework problems should be solved individually and handed in in due time for grading. Please recall the KTH rules for examination.
- For PhD students we will in addition require:
  - peer grading of homework
  - presentation of selected topics.

Requirements

- Individual solutions to homework problems, and active participation. Preliminary grading for the masters level course will be: E=60%, D=65% , C=70%, B=80%, A=90% of max score.
- For PhD students we require at least 80% plus the additional tasks.
- Number of credits:
  Master students: 7.5 ECTS (graded by F,E,D,C,B,A)
  PhD students: 10 ECTS (Pass/Fail)

Preliminary Schedule

Lect	Date	Time	Room
1	26/3	10-12	Q31
2	2/4	10-12	Q34
3	9/4	10-12	Q34
4	16/4	10-12	Q33
5	24/4	10-12	
6	30/4	15-17	
7	7/5	10-12	
8	12/5	10-12	
9	19/5	10-12	

For PhD students there will be some additional lectures with presentations.

Course Outline (preliminary)

1. Ch 0: Review : vector spaces, inner product, determinants, rank
2. Ch. 1: Eigenvalues, eigenvectors, similarity characteristic polynomial
3. Ch. 2-3: Unitary equivalence, QR-factorization, canonical forms, polynomials and matrices
4. Ch. 4: Hermitian and symmetric matrices, variational characterization of eigenvalues, simultaneous diagonalization
5. Ch. 5: Norms for vectors and matrices
6. Ch. 7: Positive definite matrices, singular value decomposition
7. Ch. 6,8: Location and perturbation of eigenvalues nonnegative matrices, positive matrices, stochastic matrices
8. (HJ "Topics ..." + add): Field of values, stable matrices, Lyapunovs theorem
9. (HJ "Topics ..." + add): Matrix equations and the Kronecker product, vectorization, Khatri-Rao product, differentiation

Vector spaces

A set V is a vector space over a field $\mathbf{F}$ (for example, the field of real $\mathbf{R}$ or of complex numbers $\mathbf{C}$) if, given

- an operation *vector addition* defined in V, denoted $v + w$ (where $v, w \in V$), and
- an operation *scalar multiplication* in V, denoted $a * v$ (where $v \in V$ and $a \in \mathbf{F}$),

the following ten properties hold for all $a, b \in \mathbf{F}$ and u, v, and $w \in V$:

Vector spaces cont'd

1. $v + w$ belongs to V. (Closure of V under vector addition.)
2. $u + (v + w) = (u + v) + w$. (Associativity of vector addition in V.)
3. There exists a neutral element 0 in V, such that for all elements v in V, $v + 0 = v$. (Existence of an additive identity element in V.)
4. For all v in V, there exists an element w in V, such that $v + w = 0$. (Existence of additive inverses in V.)
5. $v + w = w + v$. (Commutativity of vector addition in V.)

Vector spaces cont'd

6. $a * v$ belongs to V. (Closure of V under scalar multiplication.)
7. $a * (b * v) = (ab) * v$. (Associativity of scalar multiplication in V.)
8. If 1 denotes the multiplicative identity of the field $\mathbf{F}$, then $1 * v = v$. (Neutrality of one.)
9. $a * (v + w) = a * v + a * w$. (Distributivity with respect to vector addition.)
10. $(a + b) * v = a * v + b * v$. (Distributivity with respect to field addition.)

Vector spaces, cont'd

The concept of a vector space is entirely abstract. To determine if a set V is a vector space, one only has to specify the set V, a field $\mathbf{F}$, and define vector addition and scalar multiplication in V. Then, if V satisfies the above ten properties, it is a vector space over the field $\mathbf{F}$.

The members of a vector space are called vectors.

Examples

We will typically encounter vector spaces formed by n-tuples of scalars from $\mathbf{F}$ denoted $\mathbf{F}^n$. (E.g., $\mathbf{R}^n$ and $\mathbf{C}^n$.)

Note however that vector spaces are also generated by, e.g.,
(i) polynomials with coefficients from $\mathbf{F}$
(ii) or functions over an interval $[a, b] \subset \mathbf{R}$.

Some other examples are:

- $\mathbf{C}$ is a vector space over $\mathbf{R}$
- $\mathbf{R}$ is a vector space over the rational numbers

Subspaces and Span

A subspace of a vector space V is a subset of V that is by itself a vector space.

Examples: $\{[\alpha\ 2\alpha]^T : \alpha \in \mathbf{R}\}$ is a subspace of $\mathbf{R}^2$.
and, similarly, $\{\alpha + j2\alpha : \alpha \in \mathbf{R}\}$ is subspace of the vector space $\mathbf{C}$ over the field $\mathbf{R}$.

Let S be a subset of V then $\mathrm{span}(S) = \{\sum_i a_i v_i : a_i \in \mathbf{F}, v_i \in S\}$.
Note that $\mathrm{span}(S)$ is always a subspace even if S may not be.

S is said to span V if $\mathrm{span}(S) = V$.

Bases

A set of vectors $\{v_i\}$ is *linearly dependent* if $\sum_i a_i v_i = 0$ for some $a_i \in \mathbf{F}$. (Otherwise it is *linearly independent*.)

Basis: A subset S of the vector space V is said to span V if every element of V can be represented as a linear combination of elements from S. A *linearly independent* set spanning V is a *basis* for V.

A basis is non-unique BUT, given a basis, any element in V can *uniquely* be represented in terms of that basis.

All bases for V have the same number of elements and that number is the dimension of V, denoted by $\dim(V)$.

The "standard basis" of $\mathbf{R}^n$ (or $\mathbf{C}^n$) is $\{e_1, ..., e_n\}$ where $e_1 = [1\ 0\ 0...]^T$ etc.

Isomorphism

Let U and V be vector spaces over $\mathbf{F}$ and let $f : U \to V$ be an *invertible* function such that $f(ax + by) = af(x) + bf(y);\ \forall x, y \in U$ and $a, b \in \mathbf{F}$.

Then f is said to be an isomorphism and U and V are isomorphic.

If U and V are finite dimensional then they are isomorphic iff they have the same dimension. This implies that all n-dim real vector spaces are isomorphic to $\mathbf{R}^n$.

Example: Consider the vector space V generated by nth order real polynomials with basis $\mathcal{B} = \{1, x, x^2, \ldots, x^n\}$. All elements $p \in V$ can be represented uniquely by $p = \sum_i a_i x^i$ with $a_i \in \mathbf{R}$ and hence we can associate p with $[p]_\mathcal{B} = [a_0, a_1, \ldots, a_n]^T$. The mapping $p \to [p]_\mathcal{B}$ is an isomorphism between V and $\mathbf{R}^{n+1}$ for any basis $\mathcal{B}$.

Matrices

A Matrix: "Array of scalars" or "linear transformation between two vector spaces"

Notation: $A \in M_{m,n}(\mathbf{F})$. Simplifications: $M_{n,n}(\cdot) = M_n(\cdot)$ often $M_{m,n}(\mathbf{C}) = M_{m,n}$.

Linear transformation

Let U (n-dim) and V (m-dim) be vector spaces over $\mathbf{F}$. Further let $\mathcal{B}_U$ and $\mathcal{B}_V$ be bases and let the vectors in U and V be represented by their n- and m-tuples over $\mathbf{F}$.

A linear transformation is a function $T : U \to V$ such that $T(a_1x_1 + a_2x_2) = a_1T(x_1) + a_2T(x_2)$ for all $a_i \in \mathbf{F}$ and $x_i \in U$.

The linear transformation $y = T(x)$ can be represented by a matrix $A \in M_{m,n}(\mathbf{F})$ as follows: $[y]_{\mathcal{B}_V} = A[x]_{\mathcal{B}_U}$.

Note that the matrix representation depends on the bases!

Range and null-space

With no loss of generality we will think of $A \in M_{m,n}(\mathbf{F})$ as a linear transformation from $\mathbf{F}^n$ to $\mathbf{F}^m$. The *domain* is $\mathbf{F}^n$ and the *range* is $\{y \in \mathbf{F}^m : y = Ax,\ x \in \mathbf{F}^n\}$.

The *null-space* of A is $\{x \in \mathbf{F}^n : Ax = 0\}$.

It holds that

$$n = \text{dim null-space of } A + \text{dim range of } A$$

Matrix multiplication (in the usual way) of $A \in M_{m,n}(\mathbf{F})$ and $B \in M_{p,q}(\mathbf{F})$ is only defined if $p = n$. It corresponds to a composition of linear transformations. Note that AB do not in general commute; that is, $AB \neq BA$. Special cases exist, but the (scaled) identity matrix is the only matrix that commutes with any other matrix.

Transpose and Hermitian adjoint

If $A = [a_{ij}] \in M_{m,n}(\mathbf{F})$ then the *transpose* of A, $A^T \in M_{n,m}(\mathbf{F})$, has a_{ij} as its (j,i):th element.

The *Hermitian adjoint* A^* of $A \in M_{m,n}(\mathbf{C})$ is defined as $A^* = \bar{A}^T$ where $\bar{A}$ is the conjugate of A.

Determinants

Let $A = [a_{ij}] \in M_n(\mathbf{F})$ and let A_{ij} denote the submatrix obtained by deleting row i and column j of A.

Laplace expansion:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

$\det(a_{ij}) = a_{ij}$

$\det(A) = 0$ iff a subset of its rows (or equiv. columns) is linearly dependent.

Multiplicativity: $\det(AB) = \det(A)\det(B)$

Elementary operations

- Interchange of two rows
- Multiplication of a row by a scalar
- Addition of a scalar multiple of one row to another row

Each $A \in M_{m,n}(\mathbf{F})$ can be reduced to its RREF (row reduced echelon form) by elementary operations: Canonical (unique) form for matrices

Rank

$\mathrm{rank}(A) =$ is the largest number of linearly independent columns (or rows) of A.

Linear system: Note that $Ax = b$ has either $0,\ 1$, or ∞ many solutions x. If it has solutions it is called *consistent*. That happens iff $\mathrm{rank}([A\ b]) = \mathrm{rank}(A)$.

Characterizations of rank: see book 0.4.4

Rank inequalities: see book 0.4.5

Rank equalities: see book 0.4.6.

Rank cont'd

Note in particular: If $A \in M_{m,n}(\mathbf{F})$ and $\mathrm{rank}(A) = k$ then it can always be written as

$$A = XBY$$

where $X \in M_{m,k}(\mathbf{F})$, $Y \in M_{k,n}(\mathbf{F})$ are full rank, and $B \in M_{k,k}(\mathbf{F})$ is nonsingular.

Nonsingularity

A linear transformation (or matrix) is said to be nonsingular if it produces the output 0 only for the input 0, otherwise it is singular. If $A \in M_{m,n}(\mathbf{F})$ and $m < n$ then A is always singular.

$A \in M_n(\mathbf{F})$ is *invertible* if there exists a matrix A^{-1} such that $A^{-1}A = I$; then also $AA^{-1} = I$ and A^{-1} is unique.

Equivalently, $A \in M_n(\mathbf{F})$ is *invertible* if the linear transformation A is one-to-one and the inverse (linear) transformation exists.

Inner product

Consider elements of $\mathbf{F}^n$ as column vectors ($\mathbf{F}^n = M_{n,1}(\mathbf{F})$).

Let $x, y \in \mathbf{C}^n$. The scalar $y^*x \equiv \langle x, y\rangle$ is the (standard or usual) inner (scalar) product of x and y on $\mathbf{C}^n$ (there are others).

We say $x, y \in \mathbf{C}^n$ are *orthogonal* if $\langle x, y\rangle = 0$.

The *Euclidean length* of $x \in \mathbf{C}^n$ is $\langle x, x\rangle^{1/2}$.

The Cauchy-Schwartz inequality: $|\langle x, y\rangle| \leq \langle x, x\rangle^{1/2}\langle y, y\rangle^{1/2}$ with equality iff x and y are dependent.

The angle between two vectors is defined by: $\cos(\theta) = \frac{|\langle x,y\rangle|}{\langle x,x\rangle^{1/2}\langle y,y\rangle^{1/2}}$

Gram-Schmidt orthonormalization – orthonormal bases – orthogonal complements

PARTITIONED MATRICES

If

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

then the *Schur complement* to A_{11} is $S_{11} = A_{22} - A_{21}A_{11}^{-1}A_{12}$.
Similarly, $S_{22} = A_{11} - A_{12}A_{22}^{-1}A_{21}$ is the Schur complement of A_{22}.
One way of writing the inverse of A is

$$A^{-1} = \begin{bmatrix} S_{22}^{-1} & -A_{11}^{-1}A_{12}S_{11}^{-1} \\ -S_{11}^{-1}A_{21}A_{11}^{-1} & S_{11}^{-1} \end{bmatrix}$$

"MATRIX INVERSION LEMMA"

If $B = A + XRY$, then (assuming the inverses exist)

$$B^{-1} = A^{-1} - A^{-1}X(R^{-1} + YA^{-1}X)^{-1}YA^{-1}$$

MORE TOPICS ...

(Classical) Adjoint of A: $\mathrm{Adj}(A)$ (also called adjugate)

Cramér's rule

Schur complements and determinants

Special matrices :

- Diagonal – triangular etc
- Permutation
- Circulant – Toeplitz – Hankel – Hessenberg – tridiagonal
- Vandermonde

Change of basis