UNIVERSITY OF MINNESOTA TWIN CITIES

C S C I 5304 Fall 2024

COMPUTATIONAL ASPECTS OF MATRIX THEORY

Class time	:	*TTh 8:15 – 9:30 am*
Room	:	*Ackerman 209*
Instructor	:	*Yousef Saad*

Lecture notes: *http://www-users.cse.umn.edu/~saad/csci5304/*

September 2, 2024

About this class

- Instructor and Teaching Assistant:

➤ Me: Yousef Saad

➤ TA: Zechen Zhang

- Course title: "Computational Aspects of Matrix Theory"

... Class aims to cover:

"Everything that you can learn in one semester about computations that involve matrices, from theory, to algorithms, matlab/Python implementations, and applications."

➤ Subject is at the core of *most* disciplines requiring numerical computing..

➤ .. and gaining importance in Computer Science (machine learning, robotics, graphics, ...)

Who is in this class today?

➤ Out of ≈ 59

Ugrad: (Total = 43)

- BS/BA CS: 30
- BS-Math: 4
- BS-DS: 3
- B.S. ME 3
- Other: Stat (1), Chem. (1), Biomed E. (1)

Grad: (Total = 16)

- CS: 11
- Civil E.: 2
- ECE: 2
- AEM: 1

Objectives of this course

Set 1 Fundamentals of matrix theory :

- Matrices, subspaces, eigenvectors
- Norms, matrix norms, understanding errors, sensitivity [somewhat theoretical]
- Various decompositions, LU, QR, SVD, ..

set 2 Computational linear algebra / Algorithms

- Solving linear systems, LU factorization
- Solving least-squares problems, QR factorization
- Eigenvalue problems - computing eigenvalues, eigenvectors,

Set 3 Linear algebra in applications

- See how numerical linear algebra is used to solve problems in (a few) computer science-related applications.
- Examples: page-rank, applications in optimization, information retrieval, applications in machine learning, control, ...

Logistics:

- Lecture notes, syllabus, schedule, a matlab/Python folder, and some basic information, can be found here:

http://www-users.cse.umn.edu/~saad/csci5304

- Everything else will be found on Canvas: Homeworks, exam/quizzes info, your grades, etc..

➤ The two sites have links that point to each other

Please Note:

➤ Homeworks, tests, and their solutions are copyrighted

> - *Solutions of HWs and tests are provided to you [Canvas] with the hope of helping you understand the material better. By accessing them you agree not to send them to others, sell them (%#!!$), or otherwise (help) make them available via external web-sites.*

About lecture notes:

➤ Lecture notes (like this first set) will be posted on the class web-site – usually before the lecture. Note: format of notes used in class may be slightly different from the one posted – but contents are identical.

➤ Review them to get some understanding if possible before class.

➤ Read the relevant section (s) in one of the references listed

➤ Lecture note sets are grouped by topics rather than by lecture.

➤ In the notes the symbol ✍1 indicates suggested easy exercises or questions sometimes solved in class.

➤ Each set will include a supplement with solutions to some of these exercises + possibly additional notes/comments (an evolving document)

In-class practice exercises

- Plan for this year: lots of exercises in class –
- ... You will need to review lecture material
- Some exercises posted in advance – [Canvas] $\rightarrow$ Do them before class.

Matlab and/or Python

- You will need to use matlab or Python+numpy for testing algorithms.
- In the past: demos in matlab - Now: in either or sometimes both
- For those interested you can turn in your codes for assignments in Python+numpy.
- Some documentation for matlab is posted in the (class) matlab folder
- Important: I post the matlab diaries used for the demos (if any)...
- ... something similar for Python [under IPython]

• If you do not know matlab at all and have difficulties with it - and you do not know python - talk to me or the TA at office hours. You may need is some initial help to get you started with matlab.

Final remarks on lecture notes

- Please do not hesitate to report errors and/or provide feedback on content.
- On occasion I will repost lecture notes with changes/additions

How to study for this course:

1) Rely primarily on lecture notes as a starting point. Use other sources [e.g., (online) books] to get a deeper understanding.

2) Do and redo the practice exercises done in class and those of the lecture notes. Participate in solving the exercises [at exam no one will be there to help you!]

3) Ask questions! Participate in discussions (office hours, canvas, ...)

Covid, Zoom, Canvas, Office Hours, etc

- Classes are all in-person. Zoom will be used only when necessary.
- If you are sick *please* do not come to class [there is really no need to] !!
- If I get sick - I will schedule the class on Zoom [Assuming I can!] –
- Office hours: See posted information for details (schedule, zoom option, etc.)
- Note: I will be away the week Sept. 16-20. First exam (Out of 5) is scheduled for Sept. 19th

GENERAL INTRODUCTION

- **Background: Linear algebra and numerical linear algebra**
- **Types of problems to be seen in this course**
- **Mathematical background - matrices, eigenvalues, rank, ...**
- **Types of matrices, structutred matrices,**
- **Quick review of Determinants**

Introduction

➤ This course is about *Matrix algorithms* or "matrix computations"

➤ It involves: algorithms for standard matrix computations (e.g. solving linear systems) - and their analysis (e.g., their cost, numerical behavior, ..)

➤ Matrix algorithms pervade most areas of science and engineering.

➤ In computer science: recent increase of interest in matrix algorithms for data mining, information retrieval, search engines, pattern recognition, graphics, ...

General Problems in Numerical Linear Algebra (dense & sparse)

- Linear systems: $Ax = b$. Often: A is large and sparse
- Least-squares problems $\min \|b - Ax\|_2$
- Eigenvalue problem $Ax = \lambda x$. Several variations -
- SVD .. and
- ... Low-rank approximation, subspace approximation, etc.
- Tensors and low-rank tensor approximation
- Matrix equations: Sylvester, Lyapunov, Riccati, ..
- Nonlinear equations – acceleration methods
- Matrix functions and applications
- Many many more ...

Background in linear algebra

➤ Review vector spaces.

➤ A vector subspace of $\mathbb{R}^n$ is a subset of $\mathbb{R}^n$ that is also a real vector space. The set of all linear combinations of a set of vectors $G = \{a_1, a_2, \ldots, a_q\}$ of $\mathbb{R}^n$ is a vector subspace called the linear span of G,

➤ If the a_i's are linearly independent, then each vector of $\mathbf{span}\{G\}$ admits a unique expression as a linear combination of the a_i's. The set G is then called a *basis*.

✍1 Recommended reading: Sections 1.1 – 1.6 of

www.cs.umn.edu/~saad/eig_book_2ndEd.pdf

Matrices

➤ A real $m \times n$ matrix A is an $m \times n$ array of real numbers

$$a_{ij}, \;\; i = 1, \ldots, m, \; j = 1, \ldots, n.$$

Set of $m \times n$ matrices is a real vector space denoted by $\mathbb{R}^{m \times n}$.

➤ Complex matrices defined similarly.

➤ A matrix represents a linear mapping between two vector spaces of finite dimension n and m:

$$x \in \mathbb{R}^n \quad \longrightarrow \quad y = Ax \in \mathbb{R}^m$$

➤ Recall: this mapping is linear [what does it mean?]

➤ Recall: Any linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ *is* a matrix vector product

Operations:

Addition: $C = A + B$, where $A, B, C \in \mathbb{R}^{m \times n}$ and

$$c_{ij} = a_{ij} + b_{ij}, \quad i = 1, 2, \ldots m, \quad j = 1, 2, \ldots n.$$

Multiplication by a scalar: $C = \alpha A$, where

$$c_{ij} = \alpha \, a_{ij}, \quad i = 1, 2, \ldots m, \quad j = 1, 2, \ldots n.$$

Multiplication by another matrix: $C = AB$,

where $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}, C \in \mathbb{R}^{m \times p}$, and

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Transposition: If $A \in \mathbb{R}^{m \times n}$ then its transpose is a matrix $C \in \mathbb{R}^{n \times m}$ with entries

$$c_{ij} = a_{ji}, i = 1, \ldots, n, \; j = 1, \ldots, m$$

Notation : A^T.

Transpose Conjugate: for complex matrices, the **transpose conjugate** matrix denoted by A^H is more relevant: $A^H = \bar{A}^T = \overline{A^T}$.

✍2 $(A^T)^T = ??$

✍3 $(AB)^T = ??$

✍4 $(A^H)^H = ??$

✍5 $(A^H)^T = ??$

✍6 $(ABC)^T = ??$

✍7 True/False: $(AB)C = A(BC)$

✍8 True/False: $AB = BA$

✍9 True/False: $AA^T = A^T A$

➤ Matlab notation - often used in this course:

$A_{:,j}$ or $A(:, j)$ == j-the column of A

$A_{i,:}$ or $A(i, :)$ == i-th row of A

Review: Matrix-matrix and Matrix-vector producs

➤ Recall definition of $C = A \times B$: $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$

➤ Recall what C represents [in terms of mappings]..

➤ Can do the product column-wise [Matlab notation used]:

$$C_{:,j} = \sum_{k=1}^{n} b_{kj} A_{:,k}$$

➤ Can do it row-wise:

$$C_{i,:} = \sum_{k=1}^{n} a_{ik} B_{k,:}$$

➤ Can do it as a sum of 'outer-product' matrices:

$$C = \sum_{k=1}^{n} A_{:,k} B_{k,:}$$

✍10 Verify all 3 formulas above..

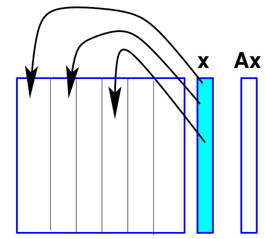
✍11 Complexity? [number of multiplications and additions]

✍12 What happens to these 3 different approches to matrix-matrix multiplication when B has one column ($p = 1$)?

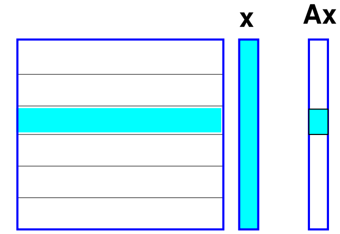
✍13 Characterize the matrices AA^T and A^TA when A is of dimension $n \times 1$.

Matrix-vector product: computing $y = Ax$

Column-form:
Linear combination of columns $A(:, j)$ with coefficients x_j yields y



Row-form:
Dot product of $A(i, :)$ and x gives y_i



Kronecker products of matrices

➤ This is a special product of matrices that can be quite useful in some situations

Definition

For $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{p \times q}$ define:
(A matrix of size $(mp) \times (nq)$).

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \cdots & \cdots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

➤ In Matlab: `kron (A,B)`

➤ Note that the dimensions m, n, p, q, can be any (> 0) integers.

✍14 Show that for 2 vectors u, v we have $v^T \otimes u = uv^T$ and also that $u \otimes v^T = uv^T$

➤ The Kronecker sum of matrices also arises in some applications. If $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times m}$ then their Kronecker sum is: $A \oplus B = A \otimes I_{m,m} + I_{n,n} \otimes B$

Range and null space (for $A \in \mathbb{R}^{m \times n}$)

➤ Range: $\text{Ran}(A) = \{Ax \mid x \in \mathbb{R}^n\} \quad \subseteq \mathbb{R}^m$

➤ Null Space: $\text{Null(A)} = \{x \in \mathbb{R}^n \mid Ax = 0\} \quad \subseteq \mathbb{R}^n$

➤ Range = linear span of the columns of A

➤ Rank of a matrix $\text{rank}(A) = \dim(\text{Ran}(A)) \quad \leq n$

➤ $\text{Ran}(A) \subseteq \mathbb{R}^m \rightarrow$ rank $(A) \leq m \rightarrow$

$$\text{rank } (A) \leq \min\{m, n\}$$

➤ rank (A) = number of linearly independent columns of A = number of linearly independent rows of A

➤ A is of full rank if $\text{rank}(A) = \min\{m, n\}$. Otherwise it is rank-deficient.

Rank+Nullity theorem for an $m \times n$ matrix:

$$dim(Ran(A)) + dim(Null(A)) = n$$

Apply to A^T: $dim(Ran(A^T)) + dim(Null(A^T)) = m \rightarrow$

$$\text{rank}(A) + dim(Null(A^T)) = m$$

➤ Terminology:

- $dim(Null(A))$ is the Nullity of A [Another term: co-rank]

✍15 Show that $A \in \mathbb{R}^{m\times n}$ is of rank one iff [if and only if] there exist two nonzero vectors $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$ such that $A = uv^T$. What are the eigenvalues and eigenvectors of A?

✍16 Is it true that: $\text{rank}(A) = \text{rank}(\bar{A}) = \text{rank}(A^T) = \text{rank}(A^H)$?

✍17 Matlab exercise: explore the matlab function `rank`.

✍18 Matlab exercise: explore the matlab function `rref`.

➤ No `rref` function in numpy – [see sympy]

✍19 Find the range and null space of the following matrix: Verify your result with matlab [hint: use `null`, `rank`, `rref`]

$$\begin{pmatrix} -1 & 1 & 0 \\ 1 & 2 & 3 \\ 1 & -2 & -1 \\ 2 & -1 & 1 \end{pmatrix}$$

Square matrices, matrix inversion, eigenvalues

➤ Square matrix: $n = m$, i.e., $A \in \mathbb{R}^{n\times n}$

➤ Identity matrix: square matrix with

$$a_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

➤ Notation: I.

➤ Property: $AI = IA = A$

➤ Inverse of A (when it exists) is a matrix C such that

$$AC = CA = I$$

Notation: A^{-1} .

Eigenvalues and eigenvectors

A complex scalar λ is called an eigenvalue of a square matrix A if there exists a nonzero vector u in $\mathbb{C}^n$ such that $Au = \lambda u$. The vector u is called an eigenvector of A associated with λ. The set of all eigenvalues of A is the 'spectrum' of A. Notation: $\Lambda(A)$.

➤ λ is an eigenvalue iff the columns of $A - \lambda I$ are linearly dependent.

➤ ... equivalent to saying that its rows are linearly dependent. So: there is a nonzero vector w such that

$$w^H(A - \lambda I) = 0$$

➤ w is a left eigenvector of A (u= right eigenvector)

➤ λ is an eigenvalue iff $\boxed{\det(A - \lambda I) = 0}$

Eigenvalues/vectors

➤ An eigenvalue is a root of the Characteristic polynomial:

$$p_A(\lambda) = \det(A - \lambda I)$$

➤ So there are n eigenvalues (counted with their multiplicities).

➤ The multiplicity of these eigenvalues as roots of p_A are called algebraic multiplicities.

➤ The geometric multiplicity of an eigenvalue λ_i is the number of linearly independent eigenvectors associated with λ_i.

➤ Geometric multiplicity is $\leq$ algebraic multiplicity.

➤ An eigenvalue is simple if its (algebraic) multiplicity is one. It is semi-simple if its geometric and algebraic multiplicities are equal.

➤ Two matrices A and B are similar if there exists a nonsingular matrix X such that $A = XBX^{-1}$

✍20 Eigenvalues of A and B are the same. What about eigenvectors?

➤ Note: A and B represent the same mapping using 2 different bases.

Fundamental Problem: Given A, find X so that B has a simpler structure (e.g., diagonal) $\rightarrow$ Eigenvalues of B easier to compute

Definition: A is diagonalizable if it is similar to a diagonal matrix

➤ We will revisit these notions later in the semester

✍21 Given a polynomial $p(t)$ how would you define $p(A)$?

✍22 Given a function $f(t)$ (e.g., e^t) how would you define $f(A)$? [Leave the full justification for next chapter]

✍23 If A is nonsingular what are the eigenvalues/eigenvectors of A^{-1}?

✍24 What are the eigenvalues/eigenvectors of A^k for a given integer power k?

✍25 What are the eigenvalues/eigenvectors of $p(A)$ for a polynomial p?

✍26 What are the eigenvalues/eigenvectors of $f(A)$ for a function f? [Diagonalizable case]

✍27 For two $n \times n$ matrices A and B are the eigenvalues of AB and BA the same?

✍28 Review the Jordan canonical form; see short description in sec. 1.8.2 of:

http://www.cs.umn.edu/~saad/eig_book_2ndEd.pdf

Define the eigenvalues, and eigenvectors from the Jordan form.

➤ Spectral radius = The maximum modulus of the eigenvalues

$$\rho(A) = \max_{\lambda \in \lambda(A)} |\lambda|.$$

➤ Trace of A = sum of diagonal elements of A.

$$\text{Tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

➤ $\text{tr}(A) =$ sum of all the eigenvalues of A counted with their multiplicities.

➤ Recall that $\det(A)$ = product of all the eigenvalues of A counted with their multiplicities.

✍29 Trace, spectral radius, and determinant of

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

Review of Determinants: summary of main results

➤ [For review only – will *not* be covered in lectures]

➤ A determinant of an $n \times n$ matrix is a number associated with this matrix. Its definition is complex for the general case $\rightarrow$ We start with $n = 2$ and list important properties for this case.

- Determinant of a 2×2 matrix is:

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

- Notation : $\det(A)$ or $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$

➤ Next we list the main properties of determinants. These properties are also true for $n \times n$ case to be defined later.

➤ Properties written for columns (easier to write) but are also true for rows

Notation: We let $A = [u, v]$ columns u, and v are in $\mathbb{R}^2$.

1 If $v = \alpha u$ then $\det(A) = 0$.

➤ Determinant of linearly dependent vectors is zero

➤ If any one column is zero then determinant is zero

2 Interchanging columns or rows:

$$\det[v, u] = -\det[u, v]$$

3 Linearity:

$$\det[u, \alpha v + \beta w] = \alpha \det[u, v] + \beta \det[u, w]$$

➤ $\det(A)$ = linear function of each column (individually)

➤ $\det(A)$ = linear function of each row (individually)

✍30 What is the determinant $\det[u, v + \alpha u]$?

4 Determinant of transpose

$$\det(A) = \det(A^T)$$

5 Determinant of Identity

$$\det(I) = 1$$

6 Determinant of a diagonal:

$$\det(D) = d_1 d_2 \cdots d_n$$

7 Determinant of a triangular matrix (upper or lower)

$$\det(T) = a_{11} a_{22} \cdots a_{nn}$$

8 Determinant of product of matrices [IMPORTANT]

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

9 Consequence: Determinant of inverse

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

✍31 What is the determinant of αA (for 2×2 matrices)?

✍32 What can you say about the determinant of a matrix which satisfies $A^2 = I$?

✍33 Is it true that $\det(A + B) = \det(A) + \det(B)$?

Determinants – 3×3 case

➤ We will define 3×3 determinants from 2×2 determinants:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

➤ This is an expansion of the det. with respect to its 1st row.

1st term = $a_{11} \times$ det of matrix obtained by deleting 1st row and 1st column.

2nd term =$-a_{12} \times$ det of matrix obtained by deleting row 1 and column 2. Note the sign change.

3rd term = $a_{13} \times$ det of matrix obtained by deleting row 1 and column 3.

✍34 Calculate $\begin{vmatrix} 2 & 3 & 0 \\ 1 & 2 & -1 \\ -1 & 2 & 1 \end{vmatrix}$

➤ We will now generalize this definition to any dimension recursively. Need to define following notation.

We denote by A_{ij} the $(n-1) \times (n-1)$ matrix obtained by deleting row i and column j from A.

Example: If $A = \begin{bmatrix} 2 & 3 & 0 \\ 1 & 2 & -1 \\ -1 & 2 & 1 \end{bmatrix}$ Then: $A_{11} = \begin{bmatrix} 2 & -1 \\ 2 & 1 \end{bmatrix}$;

$$A_{12} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}; \; A_{13} = \begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix}; \; A_{23} = \begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix}$$

Definition The determinant of a matrix $A = [a_{ij}]$ is the sum

$$\text{det}(A) = +a_{11}\text{det}(A_{11}) - a_{12}\text{det}(A_{12}) + a_{13}\text{det}(A_{13}) - a_{14}\text{det}(A_{14}) + \cdots + (-1)^{1+n} a_{1n}\text{det}(A_{1n})$$

➤ Note the alternating signs

➤ We can write this as :

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

➤ This is an expansion with respect to the 1st row.

Generalization: Cofactors

Define $c_{ij} = (-1)^{i+j} \text{det } A_{ij}$ = cofactor of entry i, j

➤ Then we get a more general expansion formula:

- det (A) can be expanded with respect to i-th as follows

$$\text{det}(A) = a_{i1}c_{i1} + a_{i2}c_{i2} + \cdots + a_{in}c_{in}$$

➤ Note i is fixed. Can be done for any i [same result each time]

➤ Case $i = 1$ corresponds to definition given earlier

➤ Similar expressions for expanding w.r.t. column j (now j is fixed)