• CSCI 5304 • Fall 2013 •

COMPUTATIONAL ASPECTS OF MATRIX THEORY

Class time	:	MW 9:45-11:00am
Room	:	KHKH 3-111
Instructor	:	Yousef Saad

Class Web-site:

www-users.cselabs.umn.edu/classes/Fall-2013/csci5304/

September 3, 2013

Let us begin ...

- Lecture notes will be posted on the class web-site – usually before the lecture.
- Review them and try to get some understanding (help: text) if possible before class.
- Lecture note sets are grouped by topics - not by lecture.
- Unite students: fonts OK? too small? let me know –
- In the notes the symbol ✍ indicates quick questions or suggested exercises
- Green boxes like this one: 2.3 , refer to related material in the text.

- **A few topics are not covered, or not covered well, in the text (e.g., complexity). Rely on lectures and the notes (when available) for these.**

- **Lecture notes will occasionally contain URL's. These are 'clickable' – For example this one:**

  `http://www.cs.umn.edu/~saad/teaching`

- **I may on occasion refer to material available on internet for supplemental information for example:**

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

- **Do not hesitate to contact me for any questions!**

GENERAL INTRODUCTION

- Linear algebra and numerical linear algebra
- Types of problems to be seen in this course
- Mathematical background - matrices, eigenvalues, ...
- Quick review of Matlab
- Vector norms, matrix norms

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, ...

Examples

➤ **Ancient Chinese Problem (3rd cent BC) [Babylonians had similar problem]:**

There are three types of corn, of which three bundles of the first, two of the second, and one of the third make 38 measures. Two of the first, three of the second and one of the third make 33 measures. And one of the first, two of the second and three of the third make 26 measures. How many measures of corn are contained of one bundle of each type?

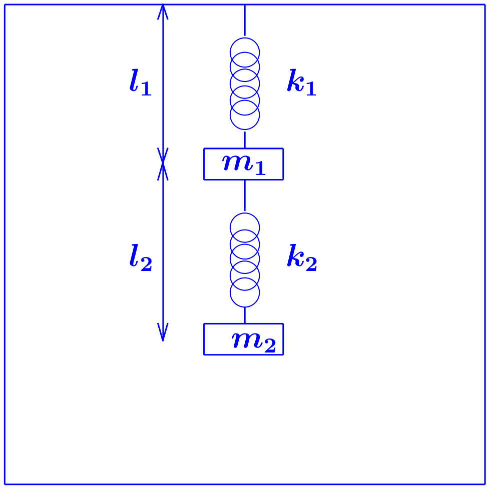
➤ Pagerank of Webpages (21st cent AD)

If one were to do a random walk from web page to web page, following each link on a given web page at random with equal likelihood, which are the pages to be encountered this way most often?

➤ Vibrations in mechanical systems. See:

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

Problem: **Determine the vibration modes of the mechanical system [to avoid resonance]. See details in Chapter 10 (sec. 10.2) of above reference.**



Examples (cont.)

➤ **Method of least-squares (inspired by first use of least squares ever, by Gauss around 1801)**

A planet follows an elliptical orbit according to $ay^2 + bxy + cx + dy + e = x^2$ in cartesian coordinates. Given a set of noisy observations of (x, y) positions, compute a, b, c, d, e, and use to predict future positions of the planet. This least squares problem is nearly rank-deficient and hence very sensitive to perturbations in the observations.

Read Wikipedia's article on planet ceres:
`http://en.wikipedia.org/wiki/Ceres_(dwarf_planet)`

Background in linear algebra

➤ **Review vector spaces. Read section 1.1 of text on vector notation.**

➤ **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 $\text{span}\{G\}$ admits a unique expression as a linear combination of the a_i's. The set G is then called a basis.**

✍ **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

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

The set of all $m \times n$ 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.

Operations:

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

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

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

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

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: $A^H = \bar{A}^T = \overline{A^T}$.

Review: Matrix-matrix and Matrix-vector producs

- Recall definition of $C = A \times B$.
- 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:

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

✍ Verify (prove) all 3 formulas above..

✍ Complexity? [number of multiplications and additions]

✍ What happen to these 3 different approaches when B has one column $(p = 1)$?

Square matrices, matrix inversion, eigenvalues

- **Square matrix:** $n = m$ - so $A \in \mathbb{R}^{n\times n}$
- **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 an **eigenvalue** of the square matrix A if a nonzero vector u of $\mathbb{C}^n$ exists such that

$$Au = \lambda u.$$

The vector u is an **eigenvector** of A associated with λ. The set of all the eigenvalues of A is the **spectrum** of A. Notation: $\lambda(A)$.

➤ λ is an eigenvalue of A if and only if $\det(A - \lambda I) = 0$

➤ $p_A(\lambda) = \det(A - \lambda I)$ is a polynomial of degree n in λ = characteristic polynomial of A.

➤ $\lambda \in \lambda(A)$ if and only if λ is a root of the characteristic polynomial $p_A(\lambda)$.

➤ 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.

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

➤ $\mathrm{tr}(A)$ = 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.

Example: Trace, spectral radius, and determinant of

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

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

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

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

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

✍ Review the Jordan canonical form. Define the eigenvalues, and eigenvectors from the Jordan form.

Range and null space

- **Range:** $\mathrm{Ran}(A) = \{Ax \mid x \in \mathbb{R}^n\}$ $\subseteq \mathbb{R}^m$
- **Null Space:** $\mathrm{Null(A)} = \{x \in \mathbb{R}^n \mid Ax = 0\ \}$ $\subseteq \mathbb{R}^n$
- **Range = linear span of the columns of A**
- **Rank of a matrix** $\mathrm{rank}(A) = \dim(\mathrm{Ran}(A))$ $\leq n$
- **$\mathrm{Ran}(A) \subseteq \mathbb{R}^m \rightarrow$ rank $(A) \leq m \rightarrow$**

  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 $\mathrm{rank}(A) = \min\{m, n\}$. 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: $dim(Ran(A^T)) + dim(Null(A^T)) = m \rightarrow$

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

✍ Show that $A \in \mathbb{R}^{n\times n}$ is of rank one iff [if and only if] there exist two nonzero vectors u and v such that

$$A = uv^T.$$

What are the eigenvalues and eigenvectors of A?

✍ Is it true that

$$\text{rank}(A) = \text{rank}(\bar{A}) = \text{rank}(A^T) = \text{rank}(A^H) \ ?$$

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

✍ Find the range and null space of the matrix

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

Verify your result with matlab.

Types of matrices (square)

- **Symmetric** $A^T = A$.
- **Skew-symmetric** $A^T = -A$.
- **Hermitian** $A^H = A$.
- **Skew-Hermitian** $A^H = -A$.
- **Normal** $A^H A = A A^H$.
- **Nonnegative** $a_{ij} \geq 0, \ i, j = 1, \ldots, n$
- Similarly for nonpositive, positive, and negative matrices
- **Unitary** $Q^H Q = I$.
- **Orthogonal** $Q^H Q = D$ (diagonal)

➤ Note: often term 'orthogonal' is used for 'unitary' in the literature. In this course we will make the distinction..

✍ What is the inverse of a unitary matrix?

✍ What can you say about the diagonal entries of a skew-symmetric (real) matrix?

✍ What can you say about the diagonal entries of a Hermitian (complex) matrix?

✍ What can you say about the diagonal entries of a skew-Hermitian (complex) matrix?

✍ The following types of matrices are normal [true-false]: real symmetric, real skew-symmetric, complex Hermitian, complex skew-Hermitian.

✍ Find all real 2×2 matrices that are normal.

✍ Show that any triangular matrix that is normal is diagonal

Matrices with structure

- **Diagonal** $a_{ij} = 0$ for $j \neq i$. Notation :

$$A = \mathrm{diag}\,(a_{11}, a_{22}, \ldots, a_{nn}) .$$

- **Upper triangular** $a_{ij} = 0$ for $i > j$.
- **Lower triangular** $a_{ij} = 0$ for $i < j$.
- **Upper bidiagonal** $a_{ij} = 0$ for $j \neq i$ or $j \neq i + 1$.
- **Lower bidiagonal** $a_{ij} = 0$ for $j \neq i$ or $j \neq i - 1$.
- **Tridiagonal** $a_{ij} = 0$ when $|i - j| > 1$.

- **Banded** $a_{ij} \neq 0$ only when $i - m_l \leq j \leq i + m_u$, 'Bandwidth' $= m_l + m_u + 1$.

- **Upper Hessenberg** $a_{ij} = 0$ when $i > j + 1$. Lower Hessenberg matrices can be defined similarly.

- **Outer product** $A = uv^T$, where both u and v are vectors.

- **Block tridiagonal** generalizes tridiagonal matrices by replacing each nonzero entry by a square matrix.

Vandermonde :

➤ Given a column of entries $[x_0, x_1, \cdots, x_n]^T$ put its powers into a matrix V:

$$V = \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^2 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{pmatrix}$$

✍ Try the matlab function `vander`

✍ What does the matrix-vector product Va represent?

✍ Interpret the solution of the linear system $Va = y$ where a is the unknown. Sketch a 'fast' solution method based on this.

Toeplitz :

- Entries are constant along diagonals, i.e., $a_{ij} = r_{j-i}$.
- Determined by $m + n - 1$ values r_{j-i}.

$$T = \underbrace{\begin{pmatrix} r_0 & r_1 & r_2 & r_3 & r_4 \\ r_{-1} & r_0 & r_1 & r_2 & r_3 \\ r_{-2} & r_{-1} & r_0 & r_1 & r_2 \\ r_{-3} & r_{-2} & r_{-1} & r_0 & r_1 \\ r_{-4} & r_{-3} & r_{-2} & r_{-1} & r_0 \end{pmatrix}}_{Toeplitz}$$

- Toeplitz systems ($m = n$) can be solver in $O(n^2)$ ops.
- The whole inverse (!) can be determined in $O(n^2)$ ops.

✍ Explore `toeplitz(c,r)` in matlab.

Hankel : Entries are constant along anti-diagonals, i.e., $a_{ij} = h_{j+i-1}$. Determined by $m+n-1$ values h_{j+i-1}.

$$H = \underbrace{\begin{pmatrix} h_1 & h_2 & h_3 & h_4 & h_5 \\ h_2 & h_3 & h_4 & h_5 & h_6 \\ h_3 & h_4 & h_5 & h_6 & h_7 \\ h_4 & h_5 & h_6 & h_7 & h_8 \\ h_5 & h_6 & h_7 & h_8 & h_9 \end{pmatrix}}_{Hankel}$$

Circulant : Entries in a row are cyclically right-shifted to form next row. Determined by n values.

$$C = \underbrace{\begin{pmatrix} v_1 & v_2 & v_3 & v_4 & v_5 \\ v_5 & v_1 & v_2 & v_3 & v_4 \\ v_4 & v_5 & v_1 & v_2 & v_3 \\ v_3 & v_4 & v_5 & v_1 & v_2 \\ v_2 & v_3 & v_4 & v_5 & v_1 \end{pmatrix}}_{Circulant}$$

✍ Explore `hankel(c,r)` in matlab.

✍ How can you generate a circulant matrix in matlab?

Sparse matrices

➤ Matrices with very few nonzero entries – so few that this can be exploited.

➤ Many of the large matrices encountered in applications are sparse.

➤ Main idea of “sparse matrix techniques” is not to represent the zeros.

➤ This will be covered in some detail at the end of the course.