# Applied Matrix Analysis, MAA704

Sergei Silvestrov, Christopher Engström, Daniel Andrén

November 14, 2013

# MAA704 Applied Matrix Analysis

Todays lecture:

- Course overview.
- Repetition of matrices and elementary operations.
- Repetition of solvability of linear equation systems, eigenvectors and eigenvalues.

# Course overview

- 11 lectures of $3 - 4$ hours.
- Examination consists of 2 seminar assignments and one project with presentation.

# Repetition matrices



A matrix is a rectangular array of numbers, symbols or expressions. Some example of matrices can be seen below:

$$\mathrm{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \mathrm{B} = \begin{bmatrix} 1+i & -2 \\ 0 & 1-i \end{bmatrix}$$

$$\mathrm{C} = \begin{bmatrix} 1 & -2 \\ -2 & 3 \\ 3 & 1 \end{bmatrix} \quad \mathrm{D} = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 3 & 0 \end{bmatrix}$$

Note that matrices need not be square as in C and D.

# Repetition matrices: notation

We take a look at matrix $C = \begin{bmatrix} 1 & -2 \\ -2 & 3 \\ 3 & 1 \end{bmatrix}$

- We denote every number or rather element of A as $a_{i,j}$, where $i$ is it's row number and $j$ is it's column number. What is $c_{3,1}$ of C above?
- The size of a matrix is the number of rows and columns (or the index of the bottom right element): We say that C above is a $3 \times 2$ matrix.
- A matrix with only one row or column is called a row vector or column vector respectively.
- The diagonal of a matrix is the elements diagonally from the top left corner towards the bottom left $(a_{1,1}, a_{2,2}, \ldots a_{n,n})$. A matrix where all elements not on the diagonal is zero is called a diagonal matrix.

# Repetition matrices: addition and subtraction



Add every element in the first matrix with correspondign element in the second matrix. We demand that both matrices have the same size.

$$\begin{bmatrix}1 & 0\\0 & 2\end{bmatrix}+\begin{bmatrix}1 & -2\\2 & -1\end{bmatrix}=\begin{bmatrix}1+1 & 0-2\\0+2 & 2-1\end{bmatrix}=\begin{bmatrix}2 & -2\\2 & 1\end{bmatrix}$$

# Repetition matrices: multiplication



Given the two matrices A and B we get the product AB as the matrix whose elements $e_{i,j}$ are found by multiplying the elements in row $i$ of A with the elements of column $j$ of B and adding the results.

$$\begin{bmatrix} 3 & 2 & 0 & 1 \\ \color{red}{1} & \color{red}{2} & \color{red}{3} & \color{red}{4} \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} \color{red}{3} & 2 & 0 & 1 \\ \color{red}{1} & 2 & 3 & 4 \\ \color{red}{0} & 1 & 2 & 1 \\ \color{red}{0} & 0 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 11 & 10 & 7 & 14 \\ \color{red}{5} & 9 & 16 & 24 \\ 1 & 4 & 8 & 9 \\ 0 & 1 & 5 & 10 \end{bmatrix}$$

For the colored element we have: $1 \cdot 3 + 2 \cdot 1 + 3 \cdot 0 + 4 \cdot 0 = 5$

# Repetition matrices: multiplication



We note that we need the number of columns in A to be the same as the number of rows in B but A can have any number of rows and $B$ can have any number of columns. Also we have:

- Generally AB $\neq$ BA, why?
- The size of AB is the number of rows of A times the number of columns of B.
- The Identity matrix I is the matrix with ones on the diagonal and zero elsewere. For the Identity matrix we have: AI $=$ IA $=$ A.

# Repetition matrices: multiplication

Multiplying a scalar with a matrix is done by multiplying the scalar with every element of the matrix.

$$3\begin{bmatrix}1 & 0\\0 & 2\end{bmatrix} = \begin{bmatrix}3\cdot 1 & 3\cdot 0\\3\cdot 0 & 3\cdot 2\end{bmatrix} = \begin{bmatrix}3 & 0\\0 & 6\end{bmatrix}$$

# Linear equation systems and matrices

- Linear equation system

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = b_2 \\ \vdots x_1 & & \vdots x_2 & & \ddots & & \vdots & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = b_m \end{cases}$$

- Represented with a matrix and column vectors

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

- Do you know any interesting linear equation systems?
- Do these system have any special properties?

# Solving a linear equation system

- Gaussian elimination.
- Elementary row operations (adding and multiplying rows).
- Find pivot elements, change matrix to diagonal/echelon row form.

$$\begin{bmatrix} 0 & \circledast & * & * & * & * & * & * \\ 0 & 0 & 0 & \circledast & * & * & * & * \\ 0 & 0 & 0 & 0 & \circledast & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \circledast & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \circledast \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

# When can a linear system be solved?

From previous courses:

## Theorem

*The following statements are equivalent for* $A \in \mathcal{M}_{n\times n}(\mathbf{K})$*:*

1. $AX = B$ *has a unique solution for all* $B \in \mathcal{M}_{n\times m}$ *where* $X \in \mathcal{M}_{n\times m}$*.*
2. $AX = 0$ *only has the trivial solution* ($X = 0$)*.*
3. A *has linearly independent row/column vectors.*
4. A *is invertible/non-degenerate/non-singular.*
5. $\det(A) \neq 0$*.*
6. A *has maximimum rank* ($\text{rank}(A) = n$)*.*
7. A*'s row/column space has dimension* $n$*.*
8. A*'s null space has dimension 0.*
9. A*'s row/column vectors span* $\mathbf{K}$

# Rank

## Definition (1)

For $A \in \mathcal{M}_{n \times n}(\mathbf{K})$ the rank(A) is the number of linearly independent rows or columns. To see that the number or linearly independent rows is the same as the number of linearly independent columns, draw a matrix in echelon form.

## Definition (2)

For $A \in \mathcal{M}_{n \times n}(\mathbf{K})$ the rank(A) is dimension of the column space of A.

# What do we already know about determinants?

Let $A, B \in \mathcal{M}_{n\times n}((K))$

- Know how to calculate it for $1 \times 1$, $2 \times 2$ and $3 \times 3$ matrices.
- Know how to calculate it for triangular matrices and matrices with linearly dependent rows.
- $\det A^\top = \det A$
- $\det AB = \det A \det B$ but $\det(A + B) \neq \det(A) + \det(B)$.
- $\det cA = c^n \det A$
- ... plus a few more things.

# The determinant function

- Determinant function

$$\det : \mathcal{M}_{n \times n}(\mathbf{K}) \mapsto \mathbf{K}$$

$$\det(\mathrm{A}) = \det(A_{.1}, A_{.2}, \ldots, A_{.k}, \ldots, A_{.n})$$

  - det is multilinear
  - det is alternating
  - $\det(\mathrm{I}) = 1$

# Multilinearity

- Linear in each argument

$$\det(A_{.1}, A_{.2}, \ldots, a \cdot A_{.k} + b \cdot B, \ldots, A_{.n}) =$$

$$= a \det(A_{.1}, A_{.2}, \ldots, A_{.k}, \ldots, A_{.n})$$
$$+ b \det(A_{.1}, A_{.2}, \ldots B, \ldots A_{.n})$$

- Example: A is a $3 \times 3$ matrix, B, C are $3 \times 1$ matrices. Let M $= [A_{.1} \ A_{.2} + B \ A_{.3} + C]$ then

$$\det(\mathrm{M}) = \det(A_{.1}, A_{.2} + B, A_{.3} + C) =$$

$$= \det(A_{.1}, A_{.2}, A_{.3} + C) + \det(A_{.1}, B, A_{.3} + C) =$$

$$= \det(\mathrm{A}) + \det(A_{.1}, B, A_{.3}) + \det(A_{.1}, A_{.2}, C) + \det(A_{.1}, B, C)$$

# A consequence of multilinearity

- $\det cA = c^n \det A$
- $\det(A_{.1}, A_{.2}, \ldots, a \cdot A_{.k}, \ldots, b \cdot A_{.i}, \ldots, A_{.n}) = a \cdot b \cdot \det A$

# Alternating

- Switching two columns (or rows) of A switches the sign of det(A)

$$\det(A_{.1}, A_{.2}, \ldots, A_{.k+1}, A_{.k}, \ldots, A_{.n}) = -\det(\mathrm{A})$$

- Most important consequence:

$$\det(A_{.1}, A_{.2}, \ldots, B, B, \ldots, A_{.n}) = 0$$

- Multilinear and alternating gives: linear dependence between columns (or rows) in A $\Leftrightarrow \det(\mathrm{A}) = 0$

# Formula for the determinant

- From previous courses you have formulas for $1 \times 1$, $2 \times 2$ and $3 \times 3$ matrices and you know how to perform column/row expansion.
- Alternative formula for the determinant

$$\det(A) = \sum_{\sigma \in S_n} a_{\sigma_1 1} a_{\sigma_2 2} \dots \ a_{\sigma_n n} \text{sign}(\sigma) \qquad (1)$$

- $\sigma$ is a permutation of columns (rows), $S_n$ is the set of all possible permutations, $a_{\sigma_i i}$ is the element that $a_{ii}$ is switched with.
- sign is a function that is equal to 1 if an even number of rows have been switched and -1 if an odd number of rows have been switched.

# Example

$$\det(A) = \sum_{\sigma \in S_n} a_{\sigma_1 1} a_{\sigma_2 2} \dots \ a_{\sigma_n n} \text{sign}(\sigma)$$

- Determinant of a $2 \times 2$ matrix

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

- Two possible permutations:
  $\sigma^a$ do not switch places of columns, $a_{\sigma_1^a 1} = a_{11}$, $a_{\sigma_2^a 2} = a_{22}$.
  $\sigma^b$ switch places of columns, $a_{\sigma_1^b 1} = a_{12}$, $a_{\sigma_2^b 2} = a_{21}$.
- $\det(A) = a_{\sigma_1^a 1} a_{\sigma_2^a 2} \text{sign}(\sigma^a) + a_{\sigma_1^b 1} a_{\sigma_2^b 2} \text{sign}(\sigma^b) = a_{11} a_{22} - a_{21} a_{12}$

# Basic idea for the formula

- Rewrite columns as sums of columns of the identity matrix

$$A_{.k} = \sum_{\sigma_k=1}^{n} a_{\sigma_k k} I_{.\sigma_k}$$

- Use the multilinearity of the determinant function

$$\det(\sum_{\sigma_1=1}^{n} a_{\sigma_1 1} I_{.\sigma_1}, A_{.2}, \ldots A_{.n}) =$$

$$= \sum_{\sigma_1=1}^{n} a_{\sigma_1 1} \det(I_{.\sigma_1}, A_{.2}, \ldots A_{.n}) =$$

$$= \sum_{\sigma_1=1}^{n} a_{\sigma_1 1} \sum_{\sigma_2=1}^{n} a_{\sigma_2 2} \det(I_{.\sigma_1}, I_{.\sigma_2}, \ldots A_{.n}) =$$

$$= \sum_{\sigma_1=1}^{n} a_{\sigma_1 1} \sum_{\sigma_2=1}^{n} a_{\sigma_2 2} \ldots \sum_{\sigma_n=1}^{n} a_{\sigma_n n} \det(I_{.\sigma_1}, I_{.\sigma_2}, \ldots I_{.\sigma_n})$$

# Basic idea for the formula

- Terms where $\sigma_i = \sigma_k$ for some $i \neq k$, $1 \leq i, k \leq n$ will be equal to zero.

$$\det(A) = \sum_{\sigma \in S_n} a_{\sigma_1 1} a_{\sigma_2 2} \dots \; a_{\sigma_n n} \det(I_{.\sigma_1}, I_{.\sigma_2}, \dots \; I_{.\sigma_n})$$

- The (*sign*) function captures the alternating behaviour.

$$\text{sign}(\sigma) = \frac{\prod_{1 \leq i < j \leq n} (\sigma_j - \sigma_i)}{\prod_{1 \leq i < j \leq n} (j - i)} = \pm 1$$

# Calculating the determinant

- Formula (1) can be unwieldy to calculate, for an $n \times n$ matrix there is $n!$ permutations.
- Mostly we calculate it by column (row) expansion.

$$\det(\mathrm{A}) = \overbrace{\sum_{k=1}^{n} a_{ki} A_{ki}}^{\text{column } i} = \overbrace{\sum_{k=1}^{n} a_{ik} A_{ik}}^{\text{row } i} \tag{2}$$

where $A_{ki}$ is equal to $(-1)^{k+i} \det(\tilde{A}_{ki})$ where $\tilde{A}_{ki}$ is equal to A with the $k$:th row and $i$:th column removed. $A_{ki}$ is called the $ki$-cofactor of A.

# Calculating the determinant

## Theorem

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \dots & a_{nn} \end{vmatrix}$$

## Theorem

$$\begin{vmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{k1} & \dots & a_{ki} & \dots & a_{kn} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{vmatrix} = (-1)^{i+k} a_{ki} \begin{vmatrix} a_{11} & \dots & a_{1(i-1)} & a_{1(i+1)} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{(k-1)1} & \dots & a_{(k-1)(i-1)} & a_{(k-1)(i+1)} & \dots & a_{(k-1)n} \\ a_{(k+1)1} & \dots & a_{(k+1)(i-1)} & a_{(k+1)(i+1)} & \dots & a_{(k+1)n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{n(i-1)} & a_{n(i+1)} & \dots & a_{mn} \end{vmatrix} \quad (3)$$

## Sketch of proof.

Reshape the left hand side so that it looks like the left hand side in the previous theorem. To move $I_{.i}$ to the first column $i-1$ column switches are required. To move the $a_{ki}$ to the first row $k-1$ row switches are required. For each switch the determinant will change sign. Thus the factor in front of the right hand side will be $(-1)^{i+k}$. Then apply the previous theorem. □

# Column (row) expansion



## Theorem

$$\det(A) = \sum_{k=1}^{n} a_{ki} A_{ki}$$

*where $A_{ki}$ is the ki-cofactor of A.*

# Column (row) expansion

## Proof.

Remember

$$A_{.i} = \sum_{k=1}^{n} a_{ki} I_{.k}$$

from which it follows

$$\det(A_{.1}, \ldots, A_{.i}, \ldots, A_{.n}) =$$

$$= \sum_{k=1}^{n} a_{ki} \det(A_{.1}, \ldots, A_{.i-1}, I_{.k}, A_{.i+1}, \ldots, A_{.n}) =$$

use (3)

$$\det(\mathrm{A}) = \sum_{k=1}^{n} a_{ki} A_{ki}$$

□

# An application of the determinant

- Consider a parallelogram defined by two vectors, $x$ and $y$.
- Then the area of the parallelogram is equal to the determinant.
- Can be generalized in the same way to higher dimensions.

# A more interesting application of the determinant

- Jacobian matrix

$$J = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

- det(J) is called the Jacobian.
- Important in many ways, for example when changing variables and integrating

$$\iint_\Omega f(x,y)\mathrm{d}x\mathrm{d}y = \iint_\Omega f(q,r)\det(J)\mathrm{d}q\mathrm{d}r$$

# Inverse

- We know how to check if the system is solvable or not, but how do we solve the system?
- Already discussed Gaussian elimination, one alternative is using Cramer's rule.

## Theorem

*If* $\mathrm{A}\mathbf{x} = \mathbf{b}$ *with* $\mathrm{A} \in \mathcal{M}_{n\times n}(\mathbf{K})$ *and* $\mathbf{x}, \mathbf{b} \in \mathcal{M}_{n\times 1}(\mathbf{K})$ *then*

$$(x)_{i1} = \frac{\det(A_{.1}, \ldots, \mathbf{b}, \ldots, A_{.n})}{\det(A_{.1}, \ldots, A_{.i}, \ldots, A_{.n})}$$

method uses an inverse matrix.

# Inverse



## Definition

The inverse of a matrix $A \in \mathcal{M}_{n \times n}$ is a matrix $X \in \mathcal{M}_{n \times n}$ such that

$$AX = XA = I$$

The inverse is denoted $X = A^{-1}$.

- If $AX = B$ then $A^{-1}AX = A^{-1}B \Leftrightarrow IX = X = A^{-1}B$ and thus we can solve linear equation systems if we know the inverse.

# What do we know about inverses?

- $(A^{-1})^{-1} = A$
- $(A + B)^{-1} \neq A^{-1} + A^{-1}$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(aB)^{-1} = \frac{1}{a}B^{-1}$
- $(A^{-1})^{\top} = (A^{\top})^{-1}$

# Calculate the inverse

- Use gaussian elimination or the adjugate formula

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

where adj(A) is the adjugate matrix.

## Definition

The adjugate matrix adj(A) of a matrix is the transpose of the cofactor matrix where we have $(\mathrm{adj}(A))_{ik} = A_{ki}$.

# Other kinds of inverses

- Left and right inverses

$$BA = I, \; AB = I$$

- Pseudoinverses

$$ABA = A \text{ and } BAB = B$$

- These kinds inverses can exist for non-square matrices, $A \in \mathcal{M}_{m\times n}$, $B \in \mathcal{M}_{n\times m}$

# Eigenvalues and eigenvectors

### Definition

If $A\mathbf{v} = \lambda \mathbf{v}$ where $A \in \mathcal{M}_{n\times n}(\mathbf{K})$, $\lambda \in \mathbf{K}$, $\mathbf{v} \in \mathcal{M}_{n\times 1}(\mathbf{K})$ then $\lambda$ is the *eigenvalue* of A and $\mathbf{v}$ is an *eigenvector* if $\mathbf{v} \neq 0$.

# What do we know about eigenvalues and eigenvectors?

- Can have several different eigenvalues and eigenvectors.
- Different eigenvectors can have the same eigenvalue.
- Can find the eigenvalues by solving $\det(A - \lambda I) = 0$ and the eigenvectors by solving $(A - \lambda I)\mathbf{v} = 0$.

# Spectrum of a matrix

- Set of eigenvalues for a matrix A is called the *spectrum* of A and is denoted by Sp(A).
- The absolute value of the largest eigenvalue is called the *spectral radius*

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

# Finding the eigenvalues

- $A\mathbf{v} = \lambda\mathbf{v} \Leftrightarrow (A - \lambda I)\mathbf{v} = 0$ if $\mathbf{v} \neq 0$ then $\det(A - \lambda I) = 0$.
- Many different methods, for example Gaussian elimination.
- Can find the eigenvalues by solving $\det(A - \lambda I) = 0$ which is an $n$th degree polynomial in $\lambda$.
- $p_A(\lambda) = \det(A - \lambda I)$ is called the *characteristic polynomial* and $p_A(\lambda) = 0$ is called the *characteristic equation*.

# Important results with eigenvalues

- $\text{tr}(A) = \sum_{i=1}^{n} a_{i,i} = \sum_{i=1}^{n} \lambda_i$
- $\det(A) = \prod_{i=1}^{n} \lambda_i$
- If $A \in \mathcal{M}_{n \times n}(\mathbf{K})$ have $n$ different eigenvalues then A is diagonalizable.