

ECE-GY 5253

Applied Matrix Theory

Instructor: Prof. Z. P. Jiang

**Control and Networks (CAN) Lab
ECE Department**

Rm 1001, 370 Jay St, Brooklyn campus
<http://engineering.nyu.edu/people/zhong-ping-jiang>
E-mail: zjiang@nyu.edu

About This Course

1. This is not a Math. Class.

Links to ECE courses:

- **Systems, Control, Robotics, Signal Processing:**
ECE-GY 6113, 6243, 6253, 7133, 7253, ROB-GYxxxx,...
- **Power Engineering:**
ECE-GY 5613, 6603, 6623, 6633, 6653, 6663,...
- **Communications, Networking, CompE:**
ECE-GY 5363, 6023, 6033, 6313, 7353, 5483, ...

About This Course

2. Suitable for both upper-level undergraduates and graduate students from diverse fields of engineering & science:

- **Electrical engineering** (wireless, control/robotics, communications, networking...)
- **Mechanical and chemical engineering**
- **Financial engineering**
- **CS, Applied mathematics, etc**

Background Knowledge

This course assumes only elementary knowledge about:

- o Algebra
- o Calculus

Tips for Getting “A”

- **Well prepared:** Reading the recommended textbook before each class
- **Practice:** exercise, homework
- **Problem-driven:** applications
- **Engaging in teaching:** Ask questions

For advanced topics, consult the recommended textbooks of Horn & Johnson (2013) and of Gantmacher (1960)

What is a Matrix?

A matrix is a rectangular collection of numbers ([J.J. Sylvester](#), 1848). For example,

A $m \times n$ matrix is often written as:

$$A = \left(a_{ij} \right)_{m \times n} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

What is a Matrix?

So, a $m \times n$ matrix has m **row vectors**:

$$(a_{i1}, \dots, a_{in}), 1 \leq i \leq m$$

and n **column vectors** ($1 \leq j \leq n$):

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \doteq \text{col}(a_{1j}, \dots, a_{mj})$$

$$\doteq (a_{1j}, \dots, a_{mj})^T$$

What is a Matrix?

It can be considered as a **linear mapping**
from \mathbb{R}^n to \mathbb{R}^m (A. Cayley, 1855):

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n \mapsto y = Ax \in \mathbb{R}^m$$

$$\text{with } y_i \doteq \sum_{j=1}^n a_{ij} x_j = a_{i1} x_1 + \cdots + a_{in} x_n$$

Comment on the computation of A

Let $e_i^n = \text{col}(0, \dots, 1, 0, \dots, 0)$, with 1 as the i th element, be a vector in the coordinate basis of \mathbb{R}^n .

Then, **for any linear mapping F** (preserving the origin),

$$Fe_i^n = \sum_{j=1}^m a_{ji} e_j^m$$

These coefficients a_{ji} form the $m \times n$ matrix

$A = (a_{ji})_{m \times n}$, associated with F .

Comment on the computation of A

Proof : Let $x = \sum_{i=1}^n x_i e_i^n$. By linearity,

$$y = Fx = \sum_{i=1}^n x_i F e_i^n = \sum_{i=1}^n x_i \sum_{j=1}^m a_{ji} e_j^m$$

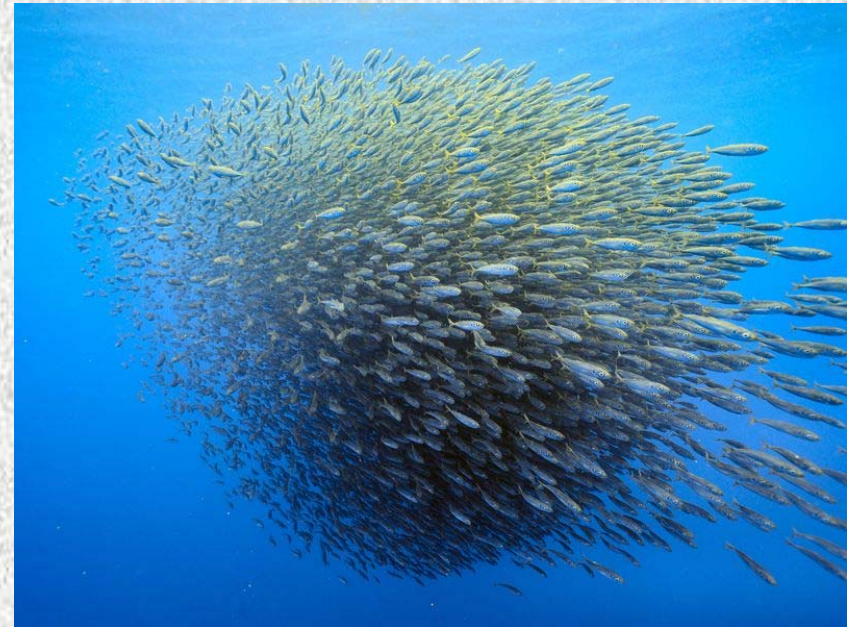
$$= \sum_{j=1}^m \left(\sum_{i=1}^n a_{ji} x_i \right) e_j^m \triangleq \sum_{j=1}^m y_j e_j^m$$

In other words,

$$y = Ax. \quad \text{DONE.}$$

A Modern Engineering Example

We have observed many interesting biological group behaviors:
Bird flocking, fish schooling, etc



- Why?
- How?

A Modern Engineering Example

Hot research topic:

Consider a group of “connected” autonomous agents (say, birds, robots or humans).

- What leads to the desired group behavior?
- Why do local interactions lead to emerging group behavior?
- How to take advantage of it?

Also, see the computer demonstration of *Reynolds's Boids model* of bird flocking at: <http://www.red3d.com/cwr/boids/>

Motivation (cont'd)

Assume that each agent updates his heading using the “average” of his nearby “neighbors”.

Nearest neighbor rule:

$$\theta_i(t+1) = \frac{1}{1+n_i(t)} \left(\theta_i(t) + \sum_{j \in N_i} \theta_j(t) \right)$$

$$\theta := \text{col}(\theta_1, \theta_2, \dots, \theta_n)$$

Motivation (cont'd)

Nearest neighbor rule:

$$\theta_i(t+1) = \frac{1}{1+n_i(t)} \left(\theta_i(t) + \sum_{j \in N_i} \theta_j(t) \right)$$
$$\theta := \text{col}(\theta_1, \theta_2, \dots, \theta_n)$$

Equivalently, in compact matrix notation

$$\theta(t+1) = F_{\sigma(t)} \theta(t), \quad t = 0, 1, 2, \dots$$

$$F_{\sigma(t)} = \left(I + D_{\sigma(t)} \right)^{-1} \left(I + A_{\sigma(t)} \right)$$

Motivation (cont'd)

It is shown in (Jadbabaie, Lin and Morse, IEEE Transactions on Automat. Control, 2003); also see (Bertsekas-Tsitsiklis, 1986) that, under mild assumptions on graph connectivity,

$$\theta(t) = \text{col}(\theta_1, \theta_2, \dots, \theta_n) \rightarrow \theta_{ss} \text{col}(1, 1, \dots, 1)$$

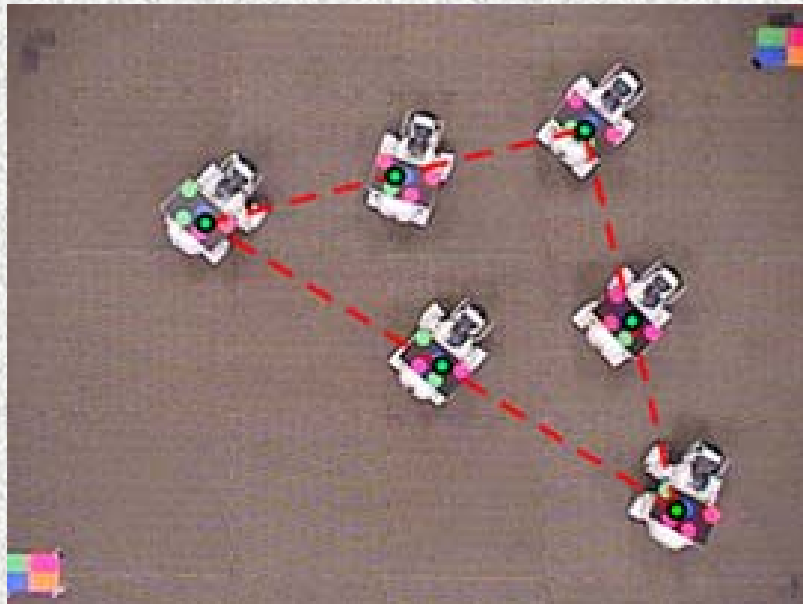
i.e.,

$$\theta_i(t) \rightarrow \theta_{ss}$$

for **all** i and **all** initial conditions $\theta_i(0)$!

Engineering Applications

Coordinated Control of Groups of Unmanned Vehicles:



Special Types of Matrices

$A \in \mathbb{R}^{m \times n}$ or \mathbb{C}^n is a

- Square matrix, if $n = m$.
- Symmetric matrix, if $n = m$ and $a_{ij} = a_{ji}$, or $A = A^T$.
- Hermitian matrix, if $A = A^* (\doteq \bar{A}^T)$.
- Non-square matrix, if $n \neq m$.
- Any scalar number is a 1×1 matrix.
- A column vector is a $m \times 1$ matrix.
- A row vector is a $1 \times n$ matrix.

**Hermitian
adjoint**



Matrix Addition and Multiplication

- 1) Addition of two matrices with the same dimensions:

$$A + B = \left(a_{ij} \right)_{m \times n} + \left(b_{ij} \right)_{m \times n} = \left(a_{ij} + b_{ij} \right)_{m \times n}.$$

- 2) Multiplication of two matrices A , B with matched dimensions:

$$AB = \left(a_{ij} \right)_{m \times n} \left(b_{ij} \right)_{n \times p} = \left(c_{ij} \right)_{m \times p}$$

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

Extension (Semi-tensor product): D. Cheng, 2005

Matrix Addition and Multiplication

3) *Associative* Property:

$$(AB)C = A(BC).$$

4) *Distributive* Property:

$$(A + B)C = AC + BC,$$

$$A(B + C) = AB + AC.$$

Notions about Vectors

- **Linear dependence**

A set of vectors $\{x_1, x_2, \dots, x_k\}$, of the same size, is said to be **linearly dependent**, if \exists constants $\{\alpha_j\}_{j=1}^k$, **not** all zero, s.t.

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k = 0.$$

Or, equivalently, $\exists j \in \{1, \dots, k\}$ such that

$$x_j = \sum_{l \neq j} c_l x_l$$

$$:= c_1 x_1 + \dots + c_{j-1} x_{j-1} + c_{j+1} x_{j+1} + \dots + c_k x_k.$$

Notions about Vectors

- **Linear independence**

A set of vectors $\{x_1, x_2, \dots, x_k\}$, of the same size, is said to be **linearly independent**, if they are **not** linearly dependent.

Or, equivalently,

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k = 0, \quad \forall \alpha_i$$

$$\Rightarrow \alpha_1 = \dots = \alpha_k = 0.$$

Remark

- 1) Ax is a linear combination of the columns of A ,
with the coordinates of x as the coefficients.
- 2) $y^T A$ is a linear combination of the rows of A ,
with the coordinates of y as the coefficients.

Examples

Are the following sets of vectors linearly dependent or independent?

1) $\left\{ (1,0,0)^T, (1,1,0)^T, (1,1,1)^T \right\}$

2) $\{0\}$

3) $\{1\}$

4) $\left\{ (1,1,1)^T, (1,2,3)^T, (2,0,-2)^T \right\}$

Notions about Vectors

- **Basis**

Consider a **subspace** V of \mathbb{R}^n or \mathbb{C}^n (being itself a vector space).

A set of vectors $\{x_1, \dots, x_k\}$ is said to be a **basis** for V , *if* :

- 1) They are linearly independent;
- 2) They span V , *i.e.*

$$\text{span}\{x_1, \dots, x_k\} = \{a_1x_1 + \dots + a_kx_k : \forall a_j\} = V.$$

Standard Basis

The basis $\{e_1, e_2, \dots, e_n\}$ is called the **standard basis** of \mathbb{R}^n , where

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Comments on Basis

- 1) In this case, the dimension of V is k .
- 2) $k \leq n$.
- 3) A subspace has an infinite number of bases.
Nonetheless, all bases must be composed of the same number, k , of vectors.

Remark

It is worth noting that *some* vector spaces may be **infinite-dimensional**, i.e. there does not exist any basis consisting of a finite number of elements. Such an example is

$$V = \{\text{all continuous functions } f : [0, 1] \rightarrow \mathbb{R}\}.$$

Note: such a space occurs in several optimization problems!

Illustrative Example

Consider the following subspace of \mathbb{R}^3 :

$$V = \left\{ (x_1, x_2, 0)^T : \forall x_1, x_2 \in \mathbb{R} \right\}.$$

Examples of a basis for V include:

1) $\left\{ (1, 0, 0)^T, (0, 1, 0)^T \right\};$

2) $\left\{ (1, 1, 0)^T, (0, 1, 0)^T \right\};$

3) $\left\{ (1, 2, 0)^T, (3, 1, 0)^T \right\}.$

1st Application: Solving Linear Equations

Solving n equations for n unknowns x_i , $i = 1, 2, \dots, n$:

$$\left. \begin{array}{l} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = b_n \end{array} \right\} Ax = b.$$

When does a solution exist? When unique?

Special Case 1: $n=1$

In this case, the equation becomes

$$a_{11}x_1 = b_1$$

Clearly,

$$a_{11} \neq 0 \Rightarrow x_1 = \frac{b_1}{a_{11}}, \text{ unique.}$$

However, when $a_{11} = b_1 = 0$,
an infinite number of solutions exist!

Special Case 2: $n=2$

In this case, the equation becomes

$$a_{11}x_1 + a_{12}x_2 = b_1$$

$$a_{21}x_1 + a_{22}x_2 = b_2.$$

Solving for x_1 and x_2 gives

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - b_2a_{12}$$

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - a_{21}b_1$$

Special Case 2: $n=2$

Denote $a_{11}a_{22} - a_{12}a_{21} \doteq \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$.

When $\det A \neq 0$, the equation has the **unique** solution:

$$x_1 = \det \begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix} / \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$x_2 = \det \begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix} / \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Cramer's Rule,
1750

Exercise

Solve

$$\begin{pmatrix} 1 & 2 \\ 3 & 7 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \end{bmatrix}.$$

Determinant

Given any square matrix $A = (a_{ij})_{n \times n}$,

its **determinant** is defined by

$$\det A \triangleq \sum_{(j_1, \dots, j_n)} s(j_1, \dots, j_n) a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

where (j_1, \dots, j_n) is one of the $n!$ **permutations** of $1, \dots, n$, whose sign is given as

$$s(j_1, \dots, j_n) = \text{sign} \prod_{1 \leq p < q \leq n} (j_q - j_p).$$

Facts about Permutations

Fact 1: If two numbers in a permutation is inter-changed, the sign of the permutation is reversed.

For example, $s(3, 1, 4, 2) = -s(4, 1, 3, 2)$.

Examples

Case 1: $n = 2$

$$s(1, 2) = 1, \quad s(2, 1) = -1.$$

Case 2: $n = 3$

$$\begin{aligned} s(1, 2, 3) &= s(3, 1, 2) = s(2, 3, 1) = 1, \\ s(1, 3, 2) &= s(2, 1, 3) = s(3, 2, 1) = -1. \end{aligned}$$

Computing a Determinant

$$\begin{aligned} & \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \\ &= \sum_{(j)=\begin{pmatrix} j_1 & j_2 \end{pmatrix}} s(j) a_{1j_1} a_{2j_2} \\ &= s \begin{pmatrix} 1 & 2 \end{pmatrix} a_{11} a_{22} + s \begin{pmatrix} 2 & 1 \end{pmatrix} a_{12} a_{21} \\ &= a_{11} a_{22} - a_{12} a_{21}. \end{aligned}$$

Facts about Permutations

Fact 2: Let the permutation j_1, \dots, j_n be formed from $1, 2, \dots, n$ by k successive inter-changes of pairs of numbers.

Then, $s(j_1, \dots, j_n) = (-1)^k$.

The permutation j is **even**, if k is even.
Otherwise, it is called an **odd permutation**.

A Puzzle?

Is it possible to rearrange the letters of the alphabet a, b, \dots, z in reverse order z, y, \dots, a by exactly 100 successive interchanges of pairs of letters?

Facts about Determinants

Fact 1: If A is a square matrix, $\det A = \det A^T$.

Proof. It follows directly from the definition of determinant. \square

An Application of Fact 1

For any given $m \times n$ matrix A , the "row rank of A " is equal to its "column rank".

Facts about Determinants

Fact 2: If two rows (or columns) of a square matrix A are interchanged, the sign of the determinant is reversed.

Examples

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = -\det \begin{pmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{pmatrix}$$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = -\det \begin{pmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{pmatrix}$$

Exercise

What is the determinant of

$$\begin{pmatrix} 3 & 2 & 1 \\ 7 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}$$

Proof of Fact 2

Let B be generated by inter-changing the rows r and s of A . That is, $b_{rj} = a_{sj}$, $b_{sj} = a_{rj}$, $b_{ij} = a_{ij}$ if $i \neq r, s$.

By definition,

$$\det B = \sum_{(j)} s(j) b_{1j_1} \dots b_{rj_r} \dots b_{sj_s} \dots b_{nj_n}$$

Let k be the permutation produced from j by interchanging j_r and j_s . Then, $s(k) = -s(j)$.

$$\text{Thus, } \det B = \sum_{(j)} -s(k) b_{1j_1} \dots b_{sj_s} \dots b_{rj_r} \dots b_{nj_n}$$

$$= - \sum_{(k)} s(k) a_{1j_1} \dots a_{rj_s} \dots a_{sj_r} \dots a_{nj_n} = -\det A.$$

An Implication of Fact 2

If a square matrix has two **identical** rows (or columns), then the determinant must be zero.

Facts about Determinants

Fact 3: If a row (or column) of a square matrix is multiplied by a constant c , the determinant is also multiplied by c .

Fact 4: If a multiple of one row (or column) is subtracted from another row (or column) of a square matrix, the determinant is unchanged.

Exercise

Show that

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 0.$$

Proof of Fact 4

Because of Fact 1, we only need to prove the row part.

Let \tilde{A} be the new matrix produced from subtracting row s by λ times row r . So,

$$\begin{aligned}\det \tilde{A} &= \sum_{(j)} s(j) a_{1j_1} \dots \mathbf{a}_{rj_r} \dots \left(a_{sj_s} - \lambda \mathbf{a}_{rj_r} \right) \dots a_{nj_n} \\ &= \sum_{(j)} s(j) a_{1j_1} \dots \mathbf{a}_{rj_r} \dots a_{sj_s} \dots a_{nj_n} \\ &\quad - \lambda \sum_{(j)} s(j) a_{1j_1} \dots \mathbf{a}_{rj_r} \dots \mathbf{a}_{rj_r} \dots a_{nj_n} \\ &= \det A - 0 := \det A.\end{aligned}$$

Facts about Determinants

Fact 5: For any "upper-triangular" square matrix

$$A = (a_{ij})_{n \times n}, \text{ i.e., } a_{ij} = 0 \text{ for } i > j,$$

$$\det A = a_{11}a_{22} \cdots a_{nn}.$$

Fact 6: For any "lower-triangular" square matrix

$$A = (a_{ij})_{n \times n}, \text{ i.e., } a_{ij} = 0 \text{ for } i < j,$$

$$\det A = a_{11}a_{22} \cdots a_{nn}.$$

Facts about Determinants

Fact 7 (Row expansion): Consider any row i of a matrix $A = (a_{ij})_{n \times n}$. Then,

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

where $c_{ik} = (-1)^{i+k} \det A_{ik}$,

$A_{ik} = (n-1) \times (n-1)$ matrix formed by deleting row i and column k from A .

Remark: Because of Fact 1, the same can be stated for “column expansion”.

An Exercise

Give a simple expression for *Vandermonde's Determinant*:

$$V_n(x_1, \dots, x_n) = \det \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$$

Solution

$$V_n(x_1, \dots, x_n) = \prod_{i>j} (x_i - x_j).$$

Idea of Proof :

If $x_i = x_j$ for some $i \neq j$, then the above is obvious.

Otherwise, $x_i, i = 1, \dots, n$, are distinct.

Then, V_n is a polynomial of degree $n-1$ in $x = x_n$,
with distinct roots x_1, x_2, \dots, x_{n-1} . That is,

$$V_n = \alpha (x - x_1) \cdots (x - x_{n-1}),$$

with $\alpha \doteq \begin{cases} \text{the coefficient of } x^{n-1} = x_n^{n-1}, \text{ i.e.} \\ V_{n-1}(x_1, \dots, x_{n-1}). \end{cases}$

Then, the identity follows by induction.

RREF: Row-reduced Echelon Form

This is a canonical form, useful for solving the system of linear equations $Ax=b$.

Indeed, it suffices to bring the augmented matrix $[A, b]$ down to a RREF.

What is a RREF?

Any RREF must meet the following requirements:

- 1) Each nonzero row has 1 as its first nonzero entry;
- 2) All other entries in the column of such a leading 1 are equal to 0;
- 3) Any rows consisting entirely of zeroes occur at the bottom of the matrix;
- 4) The leading 1's occur in a "stairstep" pattern, left to right.

An Example of RREF

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

An Example of RREF

$$\left(\begin{array}{cccc|c} \color{red}{1} & -1 & 0 & 0 & 3 \\ 0 & 0 & \color{red}{1} & 0 & -1 \\ 0 & 0 & 0 & \color{red}{1} & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

$\underbrace{\hspace{10em}}_A \quad \underbrace{\hspace{1em}}_b$

So, the solutions are:

$$\begin{cases} x_1 - x_2 = 3, \\ x_3 = -1, \\ x_4 = 5. \end{cases}$$

Conversion to RREF

Any matrix can be converted into a (**unique**) RREF, via the following elementary (row!) operations:

Type 1) Interchange of two rows;

Type 2) Multiplication of a row by a nonzero scalar;

Type 3) Addition of a scalar multiple of one row to another row.

An Exercise

Give the RREFs of the following matrices:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \text{ and } \begin{pmatrix} 3 & 2 & 1 \\ 1 & 3 & 2 \\ 2 & 1 & 3 \end{pmatrix}.$$

Circulant matrix

Or, Toeplitz matrix

Homework #1

1. Prove the following identities:

$$(A + B)^T = A^T + B^T,$$

$$(AB)^T = B^T A^T,$$

$$(A_1 A_2 \cdots A_n)^T = A_n^T \cdots A_2^T A_1^T.$$

2. Show that AB is not necessarily symmetric if A and B are symmetric.

Homework #1

3. If $A + jB$ is Hermitian, A, B real, then

$$A^T = A, \quad B^T = -B.$$

4. For any square matrix $A = \begin{pmatrix} A_1 & * \\ O & A_2 \end{pmatrix}$

with A_1, A_2 two square submatrices,
show that $\det A = \det A_1 \cdot \det A_2$.

Homework #1

5. (*Optional*) Give a simple expression for

$$\det \begin{bmatrix} a & a & a & a \\ a & b & b & b \\ a & b & c & c \\ a & b & c & d \end{bmatrix}$$