TATA53 Lecture 2

Linear Algebra Honours course

Jonathan Nilsson

Linköping University

Part I

Direct sum

Direct sum (internal)

Let U_1 and U_2 be subspaces of V. If every $v \in V$ can be expressed uniquely as

$$v = u_1 + u_2 \text{ where } u_1 \in U_1,\ u_2 \in U_2,$$

then we write $V = U_1 \oplus U_2$.

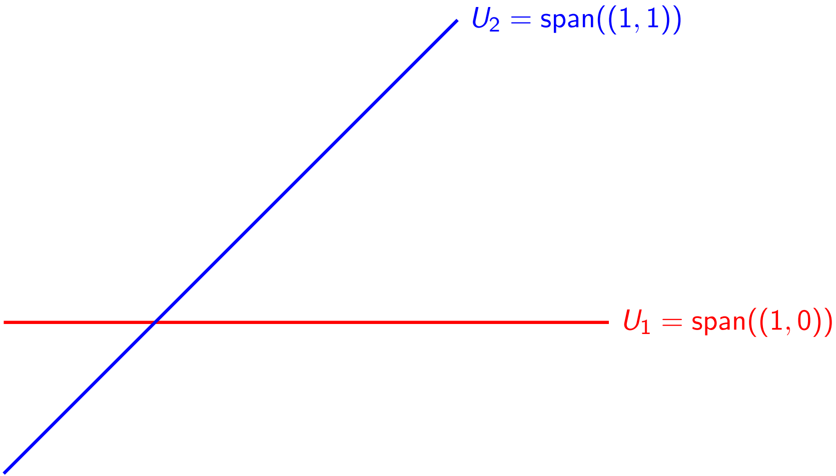
Direct sum (internal)

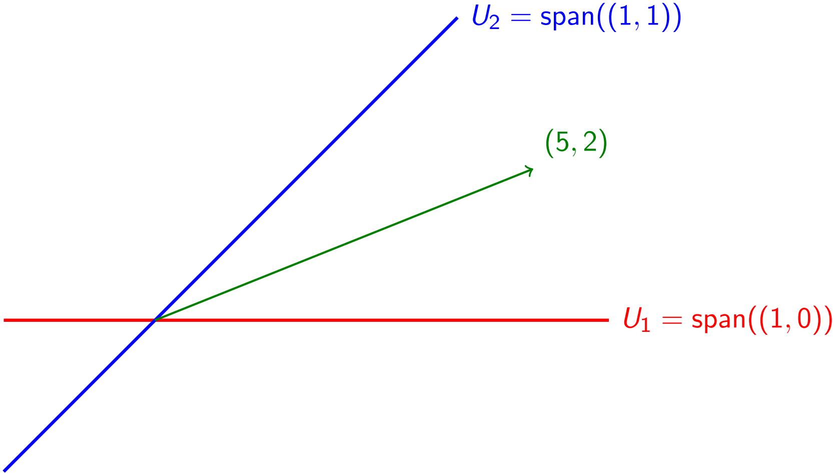
Let U_1 and U_2 be subspaces of V. If every $v \in V$ can be expressed uniquely as

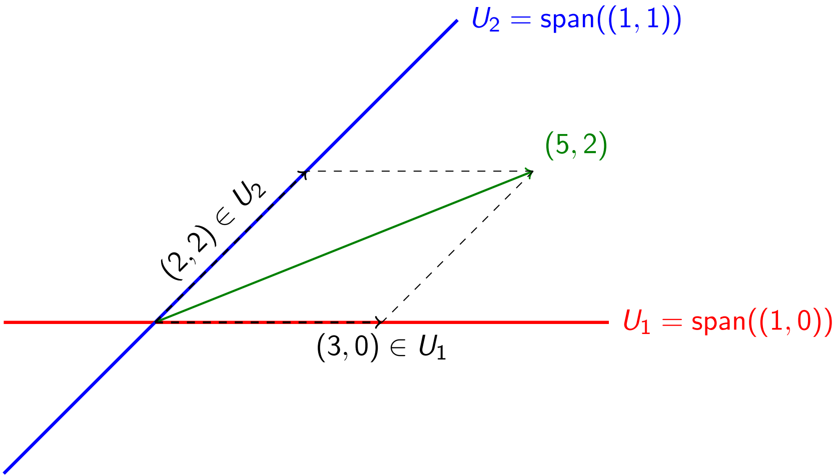
$$v = u_1 + u_2 \text{ where } u_1 \in U_1,\ u_2 \in U_2,$$

then we write $V = U_1 \oplus U_2$.
With respect to this direct sum we can project vectors onto subspaces, if $v = u_1 + u_2$ we write

$$P_{U_1}(v) = u_1 \quad \text{and} \quad P_{U_2}(v) = u_2.$$


$U_2 = \text{span}((1,1))$
$U_1 = \text{span}((1,0))$


$U_2 = \text{span}((1,1))$
$(5,2)$
$U_1 = \text{span}((1,0))$



In this example we have $P_{U_1}(5,2) = (3,0)$ and $P_{U_2}(5,2) = (2,2)$.

Theorem

Let U_1, U_2 be subspaces of V. Then $V = U_1 \oplus U_2$ if and only if

$$U_1 + U_2 = V \quad \text{and} \quad U_1 \cap U_2 = \{0\}.$$

Theorem

Let U_1, U_2 be subspaces of V. Then $V = U_1 \oplus U_2$ if and only if

$$U_1 + U_2 = V \quad \text{and} \quad U_1 \cap U_2 = \{0\}.$$

Where $U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$.

Theorem

Let U_1, U_2 be subspaces of V. Then $V = U_1 \oplus U_2$ if and only if

$$U_1 + U_2 = V \quad \text{and} \quad U_1 \cap U_2 = \{0\}.$$

Where $U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$.

Example

Consider the vector space $\mathrm{Mat}_n(\mathbb{R})$ and consider the subspaces
$\mathcal{S} = \{A \in \mathrm{Mat}_n(\mathbb{R}) | A^T = A\}$ (symmetric matrices)
$\mathcal{S}' = \{A \in \mathrm{Mat}_n(\mathbb{R}) | A^T = -A\}$ (skew-symmetric matrices)
Is $\mathrm{Mat}_n(\mathbb{R}) = \mathcal{S} \oplus \mathcal{S}'$?

External direct sum

Definition

Let V and W be vector spaces over the same field $\mathbb{F}$. We define the **external direct sum** as the $\mathbb{F}$-vector space

$$V \oplus W = V \times W = \{(v, w) \mid v \in V, w \in W\}$$

equipped with the natural addition and scalar multiplication:

$$(v, w) + (v', w') := (v + v', w + w') \qquad \lambda(v, w) = (\lambda v, \lambda w).$$

External direct sum

Definition

Let V and W be vector spaces over the same field $\mathbb{F}$. We define the **external direct sum** as the $\mathbb{F}$-vector space

$$V \oplus W = V \times W = \{(v, w) \mid v \in V, w \in W\}$$

equipped with the natural addition and scalar multiplication:

$$(v, w) + (v', w') := (v + v', w + w') \qquad \lambda(v, w) = (\lambda v, \lambda w).$$

Note that $\dim(V \oplus W) = \dim(V) + \dim(W)$.

Part II

Quotient spaces

Affine subsets

Definition

Let U be a subspace of a vector space V. For each vector $v \in V$ we define the corresponding **affine subset** of V as

$$v + U = \{v + u \mid u \in U\}.$$

Affine subsets

Definition

Let U be a subspace of a vector space V. For each vector $v \in V$ we define the corresponding **affine subset** of V as

$$v + U = \{v + u \mid u \in U\}.$$

Geometrically, $v + U$ looks like U but translated away from the origin by a vector V.

Affine subsets

Definition

Let U be a subspace of a vector space V. For each vector $v \in V$ we define the corresponding **affine subset** of V as

$$v + U = \{v + u \mid u \in U\}.$$

Geometrically, $v + U$ looks like U but translated away from the origin by a vector V.

Definition

Let U be a subspace of a vector space V. The set of all affine subsets

$$V/U := \{v + U \mid v \in V\}$$

is also called the **quotient space** of V by U. V/U is a vector space under the operations

$$(v + U) + (w + U) = (v + w) + U \qquad \lambda(v + U) = \lambda v + U.$$

Affine subsets

Definition

Let U be a subspace of a vector space V. For each vector $v \in V$ we define the corresponding **affine subset** of V as

$$v + U = \{v + u \mid u \in U\}.$$

Geometrically, $v + U$ looks like U but translated away from the origin by a vector V.

Definition

Let U be a subspace of a vector space V. The set of all affine subsets

$$V/U := \{v + U \mid v \in V\}$$

is also called the **quotient space** of V by U. V/U is a vector space under the operations

$$(v + U) + (w + U) = (v + w) + U \qquad \lambda(v + U) = \lambda v + U.$$

Note that $\dim(V/U) = \dim(V) - \dim(U)$.

Part III

Linear maps

Linear maps

Theorem

Let V and W be vector spaces over the same field $\mathbb{F}$. A map $F : V \to W$ is called **linear** if

$$F(u+v) = F(u) + F(v) \quad \text{and} \quad F(\lambda v) = \lambda F(v) \quad \text{for all } u, v \in V,\ \lambda \in \mathbb{F}.$$

Theorem

Let V and W be vector spaces over the same field $\mathbb{F}$. A map $F : V \to W$ is called **linear** if

$$F(u+v) = F(u) + F(v) \quad \text{and} \quad F(\lambda v) = \lambda F(v) \quad \text{for all } u, v \in V,\ \lambda \in \mathbb{F}.$$

Let $\mathcal{B} = (v_1, \ldots, v_n)$ be a basis for V, let $\mathcal{B}' = (w_1, \ldots, w_m)$ be a basis for W. Then

$$F(\lambda_1 v_1 + \cdots \lambda_n v_n) = \lambda_1 F(v_1) + \cdots \lambda_n F(v_n).$$

Linear maps

Theorem

Let V and W be vector spaces over the same field $\mathbb{F}$. A map $F : V \to W$ is called **linear** if

$$F(u+v) = F(u) + F(v) \quad \text{and} \quad F(\lambda v) = \lambda F(v) \quad \text{for all } u, v \in V,\ \lambda \in \mathbb{F}.$$

Let $\mathcal{B} = (v_1, \ldots, v_n)$ be a basis for V, let $\mathcal{B}' = (w_1, \ldots, w_m)$ be a basis for W. Then

$$F(\lambda_1 v_1 + \cdots \lambda_n v_n) = \lambda_1 F(v_1) + \cdots \lambda_n F(v_n).$$

Express each $F(v_i)$ in the basis $\mathcal{B}'$ and put these coordinate vectors as columns in a matrix:

$$A = [F]_{\mathcal{B}',\mathcal{B}} = \begin{pmatrix} | & | & & | \\ F(v_1) & F(v_2) & \cdots & F(v_n) \\ | & | & & | \end{pmatrix} \in \mathrm{Mat}_{m\times n}(\mathbb{F})$$

Linear maps

Theorem

Let V and W be vector spaces over the same field $\mathbb{F}$. A map $F : V \to W$ is called **linear** if

$$F(u+v) = F(u) + F(v) \quad \text{and} \quad F(\lambda v) = \lambda F(v) \quad \text{for all } u, v \in V,\ \lambda \in \mathbb{F}.$$

Let $\mathcal{B} = (v_1, \ldots, v_n)$ be a basis for V, let $\mathcal{B}' = (w_1, \ldots, w_m)$ be a basis for W. Then

$$F(\lambda_1 v_1 + \cdots \lambda_n v_n) = \lambda_1 F(v_1) + \cdots \lambda_n F(v_n).$$

Express each $F(v_i)$ in the basis $\mathcal{B}'$ and put these coordinate vectors as columns in a matrix:

$$A = [F]_{\mathcal{B}',\mathcal{B}} = \begin{pmatrix} | & | & & | \\ F(v_1) & F(v_2) & \cdots & F(v_n) \\ | & | & & | \end{pmatrix} \in \mathrm{Mat}_{m\times n}(\mathbb{F})$$

Then $F(v)_{\mathcal{B}'} = Av_{\mathcal{B}}$, so A is the matrix of F with respect to the bases $\mathcal{B}$ and $\mathcal{B}'$.

Kernel and image

Definition

Let $F : V \to W$ be a linear map (or matrix). We define

- $\mathrm{Ker}(F) = \{v \in V \mid F(v) = 0\}$, the **kernel** or **nullspace** of F.
- $\mathrm{Im}(F) = \{F(v) \mid v \in V\}$, the **image** of F.

Kernel and image

Definition

Let $F : V \to W$ be a linear map (or matrix). We define

- $\mathrm{Ker}(F) = \{v \in V \mid F(v) = 0\}$, the **kernel** or **nullspace** of F.
- $\mathrm{Im}(F) = \{F(v) \mid \ v \in V\}$, the **image** of F.

Then $\mathrm{Ker}(F)$ is a subspace of V and $\mathrm{Im}(F)$ is a subspace of W.

Kernel and image

Definition

Let $F : V \to W$ be a linear map (or matrix). We define

- $\mathrm{Ker}(F) = \{v \in V \mid F(v) = 0\}$, the **kernel** or **nullspace** of F.
- $\mathrm{Im}(F) = \{F(v) \mid v \in V\}$, the **image** of F.

Then $\mathrm{Ker}(F)$ is a subspace of V and $\mathrm{Im}(F)$ is a subspace of W.

Rank nullity theorem (dimension theorem)

$$\dim(\mathrm{Ker}(F)) + \dim(\mathrm{Im}(F)) = \dim(V).$$

$\dim(\mathrm{Im}(F))$ is called the **rank** of F.

Inverses

The **inverse** to a linear map $F : V \to W$ is a linear map $G : W \to V$ such that

$$G(F(v)) = v \text{ for all } v \in V \text{ and } F(G(w)) = w \text{ for all } w \in W.$$

We write $G = F^{-1}$ if such a map exists.

Part IV

Matrices

Matrices

$\mathrm{Mat}_{m \times n}(\mathbb{F})$ = the $\mathbb{F}$-vector space of $m \times n$-matrices with coefficients in $\mathbb{F}$.

Matrices

$\mathrm{Mat}_{m\times n}(\mathbb{F})$ = the $\mathbb{F}$-vector space of $m \times n$-matrices with coefficients in $\mathbb{F}$.

Standard basis: $e_{ij} \in \mathrm{Mat}_{m\times n}(\mathbb{F})$ - matrix with a single 1 in position (i, j), for example:

$$\begin{pmatrix} 0 & 2 & 5 \\ 0 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix} = 2e_{12} + 5e_{13} - 3e_{23}.$$

Matrices

$\mathrm{Mat}_{m\times n}(\mathbb{F})$ = the $\mathbb{F}$-vector space of $m \times n$-matrices with coefficients in $\mathbb{F}$.

Standard basis: $e_{ij} \in \mathrm{Mat}_{m\times n}(\mathbb{F})$ - matrix with a single 1 in position (i, j), for example:

$$\begin{pmatrix} 0 & 2 & 5 \\ 0 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix} = 2e_{12} + 5e_{13} - 3e_{23}.$$

For a matrix $A = (a_{ij})_{ij} = \sum_{i,j} a_{ij}e_{ij}$ we define

- $A^T = (a_{ji})_{ij}$ the **transpose** of A
- $A^* = (\overline{a_{ji}})_{ij}$ the **Hermitian conjugate** (conjugate-transpose) of A (when $\mathbb{F} = \mathbb{C}$)

Matrix properties

A matrix $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{F})$ is called

- **Diagonal** if $a_{ij} = 0$ whenever $i \neq j$
- **Upper triangular** if $a_{ij} = 0$ whenever $i > j$
  (strictly upper triangular if $a_{ii} = 0$ also)
- **Lower triangular** if $a_{ij} = 0$ whenever $j > i$
  (strictly lower triangular if $a_{ii} = 0$ also)
- **Symmetric** if $a_{ij} = a_{ji}$
  (skew-symmetric if $a_{ij} = -a_{ji}$)
- **Hermitian** $a_{ij} = \overline{a_{ji}}$ (when $\mathbb{F} = \mathbb{C}$)
  (skew-Hermitian if $a_{ij} = -\overline{a_{ji}}$)

Multiplication

The standard basis-matrices multiply as

$$e_{ij} e_{kl} = \begin{cases} e_{il} & \text{if } j = k \\ 0 & \text{otherwise} \end{cases} = \delta_{jk} e_{il}.$$

Multiplication

The standard basis-matrices multiply as

$$e_{ij}e_{kl} = \begin{cases} e_{il} & \text{if } j = k \\ 0 & \text{otherwise} \end{cases} = \delta_{jk}e_{il}.$$

Let $A = (a_{ij}) = \sum_{i,j} a_{ij}e_{ij} \in \mathrm{Mat}_{m\times n}(\mathbb{F})$ and let $B = (b_{ij}) = \sum_{i,j} b_{ij}e_{ij} \in \mathrm{Mat}_{n\times k}(\mathbb{F})$. Then

$$AB = \Big(\sum_{r=1}^{n} a_{ir}b_{rj}\Big)_{ij} \in \mathrm{Mat}_{m\times k}(\mathbb{F}).$$

Trace

The **trace** of a $n \times n$-matrix is the sum of the diagonal elements: $\mathrm{tr}(A) = \sum_{k=1}^{n} a_{kk}$.

Trace

The **trace** of a $n \times n$-matrix is the sum of the diagonal elements: $\mathrm{tr}(A) = \sum_{k=1}^{n} a_{kk}$.

Theorem

$\mathrm{tr}(AB) = \mathrm{tr}(BA)$ whenever both products are defined.

Trace

The **trace** of a $n \times n$-matrix is the sum of the diagonal elements: $\mathrm{tr}(A) = \sum_{k=1}^{n} a_{kk}$.

Theorem

$\mathrm{tr}(AB) = \mathrm{tr}(BA)$ whenever both products are defined.

For example, with $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}$ we have

$$\mathrm{tr}(AB) = \mathrm{tr}\begin{pmatrix} 2 & 3 \\ 2 & 5 \end{pmatrix} = 7 \text{ and } \mathrm{tr}(BA) = \mathrm{tr}\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 1 & 4 & 3 \end{pmatrix} = 7.$$

Trace

The **trace** of a $n \times n$-matrix is the sum of the diagonal elements: $\mathrm{tr}(A) = \sum_{k=1}^{n} a_{kk}$.

Theorem

$\mathrm{tr}(AB) = \mathrm{tr}(BA)$ whenever both products are defined.

For example, with $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}$ we have

$$\mathrm{tr}(AB) = \mathrm{tr}\begin{pmatrix} 2 & 3 \\ 2 & 5 \end{pmatrix} = 7 \text{ and } \mathrm{tr}(BA) = \mathrm{tr}\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 1 & 4 & 3 \end{pmatrix} = 7.$$

It follows that $\mathrm{tr}(S^{-1}AS) = \mathrm{tr}(SS^{-1}A) = \mathrm{tr}(A)$, so the trace is basis-independent.

Nilpotent matrices

A matrix A is called **nilpotent** if $A^n = 0$ for some n.

Nilpotent matrices

A matrix A is called **nilpotent** if $A^n = 0$ for some n.

Example

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad A^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad A^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad A^4 = 0$$

Nilpotent matrices

A matrix A is called **nilpotent** if $A^n = 0$ for some n.

Example

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad A^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad A^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad A^4 = 0$$

Similarly, a linear map $F : V \to V$ is called nilpotent if $F^n = 0$ for some n.

Block matrices

It is sometimes useful to consider a matrix as a matrix with matrix-coefficients:

$$X = \left(\begin{array}{cc|cc} 3 & 2 & 3 & 0 \\ 1 & 1 & 0 & 3 \\ \hline 0 & 0 & 6 & 4 \\ 0 & 0 & 2 & 2 \end{array}\right) = \left(\begin{array}{c|c} A & 3I \\ \hline 0 & 2A \end{array}\right) \quad \text{where} \quad A = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}.$$

Block matrices

It is sometimes useful to consider a matrix as a matrix with matrix-coefficients:

$$X = \left(\begin{array}{cc|cc} 3 & 2 & 3 & 0 \\ 1 & 1 & 0 & 3 \\ \hline 0 & 0 & 6 & 4 \\ 0 & 0 & 2 & 2 \end{array}\right) = \left(\begin{array}{c|c} A & 3I \\ \hline 0 & 2A \end{array}\right) \quad \text{where} \quad A = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}.$$

This can simplify calculations, for example:

$$X^2 = \left(\begin{array}{c|c} A & 3I \\ \hline 0 & 2A \end{array}\right)\left(\begin{array}{c|c} A & 3I \\ \hline 0 & 2A \end{array}\right) = \left(\begin{array}{c|c} A \cdot A + 3I \cdot 0 & A \cdot 3I + 3I \cdot 2A \\ \hline 0 \cdot A + 2A \cdot 0 & 0 \cdot 3I + 2A \cdot 2A \end{array}\right)$$

$$= \left(\begin{array}{c|c} A^2 & 9A \\ \hline 0 & 4A^2 \end{array}\right) = \left(\begin{array}{cc|cc} 11 & 8 & 27 & 18 \\ 4 & 3 & 9 & 9 \\ \hline 0 & 0 & 44 & 32 \\ 0 & 0 & 16 & 12 \end{array}\right).$$

Part V

Echelon forms

Matrices in echelon forms

$$A = \begin{pmatrix} \textcircled{1} & 1 & 9 & 2 & 1 & 8 \\ 0 & \textcircled{2} & 4 & 1 & 2 & 5 \\ 0 & 0 & 0 & \textcircled{4} & 1 & 9 \\ 0 & 0 & 0 & 0 & \textcircled{3} & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} \textcircled{1} & 0 & 7 & 0 & 0 & 2 \\ 0 & \textcircled{1} & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & \textcircled{1} & 0 & 2 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Matrices in echelon forms

$$A = \begin{pmatrix} \textcircled{1} & 1 & 9 & 2 & 1 & 8 \\ 0 & \textcircled{2} & 4 & 1 & 2 & 5 \\ 0 & 0 & 0 & \textcircled{4} & 1 & 9 \\ 0 & 0 & 0 & 0 & \textcircled{3} & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} \textcircled{1} & 0 & 7 & 0 & 0 & 2 \\ 0 & \textcircled{1} & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & \textcircled{1} & 0 & 2 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

A is in **row echelon form** (REF) - first nonzero element in each row is to the left of first nonzero element in rows below. Zero-rows at bottom. The encircled elements are **pivots**.

Matrices in echelon forms

$$A = \begin{pmatrix} \textcircled{1} & 1 & 9 & 2 & 1 & 8 \\ 0 & \textcircled{2} & 4 & 1 & 2 & 5 \\ 0 & 0 & 0 & \textcircled{4} & 1 & 9 \\ 0 & 0 & 0 & 0 & \textcircled{3} & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} \textcircled{1} & 0 & 7 & 0 & 0 & 2 \\ 0 & \textcircled{1} & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & \textcircled{1} & 0 & 2 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

A is in **row echelon form** (REF) - first nonzero element in each row is to the left of first nonzero element in rows below. Zero-rows at bottom. The encircled elements are **pivots**.
B is in **reduced row echelon form** (RREF) - also pivots are 1 and have zeros above them.

Matrices in echelon forms

$$A = \begin{pmatrix} \textcircled{1} & 1 & 9 & 2 & 1 & 8 \\ 0 & \textcircled{2} & 4 & 1 & 2 & 5 \\ 0 & 0 & 0 & \textcircled{4} & 1 & 9 \\ 0 & 0 & 0 & 0 & \textcircled{3} & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} \textcircled{1} & 0 & 7 & 0 & 0 & 2 \\ 0 & \textcircled{1} & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & \textcircled{1} & 0 & 2 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

A is in **row echelon form** (REF) - first nonzero element in each row is to the left of first nonzero element in rows below. Zero-rows at bottom. The encircled elements are **pivots**.
B is in **reduced row echelon form** (RREF) - also pivots are 1 and have zeros above them.

Theorem

Every $A \in \mathrm{Mat}_{m\times n}(\mathbb{F})$ can be reduced to RREF by row operations. The RREF is unique.

Matrices in echelon forms

$$A = \begin{pmatrix} \textcircled{1} & 1 & 9 & 2 & 1 & 8 \\ 0 & \textcircled{2} & 4 & 1 & 2 & 5 \\ 0 & 0 & 0 & \textcircled{4} & 1 & 9 \\ 0 & 0 & 0 & 0 & \textcircled{3} & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} \textcircled{1} & 0 & 7 & 0 & 0 & 2 \\ 0 & \textcircled{1} & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & \textcircled{1} & 0 & 2 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

A is in **row echelon form** (REF) - first nonzero element in each row is to the left of first nonzero element in rows below. Zero-rows at bottom. The encircled elements are **pivots**.
B is in **reduced row echelon form** (RREF) - also pivots are 1 and have zeros above them.

Theorem

Every $A \in \mathrm{Mat}_{m\times n}(\mathbb{F})$ can be reduced to RREF by row operations. The RREF is unique.

Note that $\mathrm{Ker}(A) = \mathrm{Ker}(B)$ when the matrices are row equivalent.

Part VI

Elementary matrices

Row operations as matrix products

$$E_1 A = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Row operations as matrix products

$$E_1 A = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Adding (-2) times the first row to the second row of A.

Row operations as matrix products

$$E_1 A = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Adding (-2) times the first row to the second row of A.

$$E_2 A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 3 & -3 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Row operations as matrix products

$$E_1 A = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Adding (-2) times the first row to the second row of A.

$$E_2 A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 3 & -3 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Multiplying the first row of A by 3.

Row operations as matrix products

$$E_1A = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Adding (-2) times the first row to the second row of A.

$$E_2A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 3 & -3 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Multiplying the first row of A by 3.

$$E_3A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 2 & 1 & 0 & 1 \end{pmatrix}$$

Row operations as matrix products

$$E_1A = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Adding (-2) times the first row to the second row of A.

$$E_2A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 3 & -3 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Multiplying the first row of A by 3.

$$E_3A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 2 & 1 & 0 & 1 \end{pmatrix}$$

Switching rows 2 and 3 of A.

Elementary matrices

Elementary matrix	Corresponding row operation
$\begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & \lambda & \ddots & \\ & & & & 1 \end{pmatrix} = I + \lambda e_{ij}$ (λ in position (i, j))	Add λ times row j to row i

Elementary matrices

Elementary matrix	Corresponding row operation
$\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \lambda & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} = I + (\lambda - 1)e_{ii}$ (identity except $\lambda \neq 0$ on position (i, i))	Multiply row i by a nonzero scalar λ

Elementary matrices

Elementary matrix	Corresponding row operation
$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & 0 & & 1 & & \\ & & & & \begin{smallmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{smallmatrix} & & & \\ & & & 1 & & 0 & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix}$ $= I - e_{ii} - e_{jj} + e_{ij} + e_{ji}$ (I but with rows i and j switched)	Switching rows i and j

Products of elementary matrices

Note that all elementary matrices are invertible, and the inverse is also an invertible matrix.

Theorem

Every invertible matrix is the product of elementary matrices.

Products of elementary matrices

Note that all elementary matrices are invertible, and the inverse is also an invertible matrix.

Theorem

Every invertible matrix is the product of elementary matrices.

Proof: Perform row operations on A until it becomes the identity matrix. If the row operations correspond to elementary matrices $E_1, \dots, E_n$ we get

$$E_n \cdots E_2 E_1 A = I \quad \Leftrightarrow \quad A = E_1^{-1} E_2^{-1} \cdots E_n^{-1}.$$

Part VII

LU-decompostion

LU-decomposition

Definition

An **LU-decomposition**, or an **LU-factorization** of $A \in \mathrm{Mat}_{m \times n}(\mathbb{F})$ is a factorization

$$A = LU$$

where

- L is lower triangular $m \times m$-matrix
- U is an upper triangular $m \times n$-matrix

LU-decomposition

Definition

An **LU-decomposition**, or an **LU-factorization** of $A \in \mathrm{Mat}_{m \times n}(\mathbb{F})$ is a factorization

$$A = LU$$

where

- L is lower triangular $m \times m$-matrix
- U is an upper triangular $m \times n$-matrix

Example:

$$\begin{pmatrix} 1 & 1 & 3 \\ 2 & 4 & 7 \\ -1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

The algorithm

Example

Find an LU factorization of

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 3 & -1 \\ -2 & 2 & -1 & 5 \end{pmatrix}.$$

The algorithm

Example

Find an LU factorization of

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 3 & -1 \\ -2 & 2 & -1 & 5 \end{pmatrix}.$$

Idea: Reduce A to REF by standard row operations, the resulting matrix is upper triangular:

$$E_n \cdots E_2 E_1 A = U \Leftrightarrow A = LU = (E_n \cdots E_2 E_1)^{-1} U,$$

where $L = (E_n \cdots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_n^{-1}$ is lower triangular since we only add higher rows to lower rows to achieve the REF.

LDU-decomposition

Definition

An **LDU-decomposition**, or an **LDU-factorization** of $A \in \mathrm{Mat}_{m\times n}(\mathbb{F})$ is a factorization

$$A = LDU$$

- L is lower triangular $m \times m$-matrix with ones on the diagonal
- D is a diagonal $m \times m$-matrix
- U is an upper triangular $m \times n$-matrix with ones on the diagonal

LDU-decomposition

Definition

An **LDU-decomposition**, or an **LDU-factorization** of $A \in \mathrm{Mat}_{m \times n}(\mathbb{F})$ is a factorization

$$A = LDU$$

- L is lower triangular $m \times m$-matrix with ones on the diagonal
- D is a diagonal $m \times m$-matrix
- U is an upper triangular $m \times n$-matrix with ones on the diagonal

This can be obtained from the LU-factorization by factoring $U = DU'$ by dividing out the leading coefficients of each row of U.

Application

Suppose that we want to solve a large linear system $Ax = b$ many times for different right sides b.

Application

Suppose that we want to solve a large linear system $Ax = b$ many times for different right sides b.
By LU-factoring A we see that

$$Ax = b \Leftrightarrow LUx = b \Leftrightarrow Ly = b \quad \text{and} \quad Ux = y.$$

Application

Suppose that we want to solve a large linear system $Ax = b$ many times for different right sides b.
By LU-factoring A we see that

$$Ax = b \Leftrightarrow LUx = b \Leftrightarrow Ly = b \quad \text{and} \quad Ux = y.$$

So we can solve two triangular systems instead with back-substitution, this is significantly faster for large matrices A.