Chapter 3. Matrix algebra

C.O.S. Sorzano

Biomedical Engineering

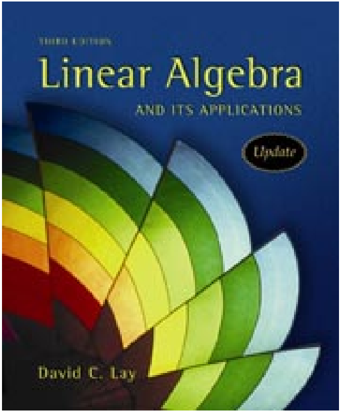
December 3, 2013

CEU
Universidad
San Pablo

Outline

3. Matrix algebra
   - Matrix operations (a)
   - Inverse of a matrix (b)
   - Elementary matrices (b)
   - An algorithm to invert matrices (b)
   - Characterization of invertible matrices (c)
   - Invertible linear transformations (c)
   - Partitioned matrices (c)
   - LU factorization (d)
   - An application to computer graphics and image processing (d)
   - Subspaces of $\mathbb{R}^n$ (e)
   - Dimension and rank (e)

References



D. Lay. Linear algebra and its applications (3rd ed). Pearson (2006). Chapter 2.

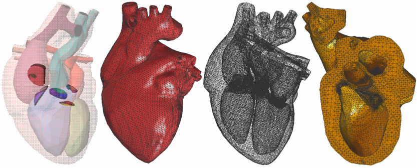
A little bit of history

Matrices appeared as a regular arrangement of numbers more than 2,000 years ago. However, it was during the XVIIth, XVIIIth and XIXth centuries that they developed in the way we know them now. Some important names in their modern development are Seki Takakazu (1683), Gottfried Leibniz (1693), Gabriel Cramer (1750), James Sylvester (1850), and Arthur Cayley (1858). They were applied in all kind of mathematical problems as a way to organize calculations.

To know more about the history of matrix algebra visit

- http://www-groups.dcs.st-and.ac.uk/~history/PrintHT/Matrices_and_determinants.html

Finite elements has been one of the most successful approaches to biomechanical modeling. In the figure we show one of such a model for the heart. Using this model, all kind of local stresses can be calculated.



(a) (b) (c) (d)

$$\underset{n\times 1}{\begin{pmatrix} f_{bc} \\ f_{int} \\ f_{vis} \\ f_{con} \end{pmatrix}} = \underset{n\times n}{\begin{bmatrix} K_{bc1} & K_{bc2} & K_{bc3} & K_{bc4} \\ K_{int1} & K_{int2} & K_{int3} & K_{int4} \\ K_{vis1} & K_{vis2} & K_{vis3} & K_{vis4} \\ K_{con1} & K_{con2} & K_{con3} & K_{con4} \end{bmatrix}} \underset{n\times 1}{\begin{pmatrix} d_{bc} \\ d_{int} \\ d_{vis} \\ d_{con} \end{pmatrix}}$$

(a)

$$\begin{pmatrix} f'_{bc} \\ f'_{int} \\ f'_{vis} \\ f'_{con} \end{pmatrix} = \begin{bmatrix} K'_{bc1} & K'_{bc2} & K'_{bc3} & K'_{bc4} \\ 0 & K'_{int2} & K'_{int3} & K'_{int4} \\ 0 & 0 & K'_{vis3} & K'_{vis4} \\ 0 & 0 & 0 & K'_{con4} \end{bmatrix} \begin{pmatrix} d_{bc} \\ d_{int} \\ d_{vis} \\ d_{con} \end{pmatrix}$$

(b)

$$\begin{pmatrix} f''_{bc} \\ f''_{int} \\ f''_{vis} \\ f''_{con} \end{pmatrix} = \begin{bmatrix} K''_{bc} & 0 & 0 & 0 \\ 0 & K''_{int} & 0 & 0 \\ 0 & 0 & K''_{vis} & 0 \\ 0 & 0 & 0 & K''_{con} \end{bmatrix} \begin{pmatrix} d_{bc} \\ d_{int} \\ d_{vis} \\ d_{con} \end{pmatrix}$$

(c)

J. Berkley, S. Weghorst, H. Gladstone, G. Raugi, D. Berg, M. Ganter. Banded Matrix Approach to Finite Element Modeling for Soft Tissue Simulation.

Outline

Basic definitions

Definition 1.1 (Matrix)

Informally, we can define a ***matrix*** *as a regular arrangement of numbers that are laid out in a grid of m rows and n columns. More formally, we say that* $A \in \mathcal{M}_{m\times n}$*. We denote as* $\mathbf{a}_j$ *as its j-th* ***column****, and* a_{ij} *the element in the i-th row and the j-th column.*

$$A = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

The ***main diagonal*** *is the vector given by* $(a_{11}, a_{22}, \ldots)$*. Two important special matrices are the* ***identity matrix*** *(*$I \in \mathcal{M}_{n\times n}$*) that is zero everywhere except the main diagonal that is full of 1s; and the* ***zero matrix*** *(*$0 \in \mathcal{M}_{m\times n}$*) that is zero everywhere.*

Example

MATLAB: `A=[1 2 3; 4 5 6]`

Matrix operations

Definition 1.2 (Sum with a scalar)

*We define the **sum operator between a scalar and a matrix** as:*

$$\begin{array}{rccl} + : & \mathbb{R} \times \mathcal{M}_{m\times n} & \to & \mathcal{M}_{m\times n} \\ & +(k, A) & \to & B = k + A \quad |b_{ij} = k + a_{ij} \end{array}$$

*We overload the notation to define the **sum operator between a matrix and a scalar** as*

$$\begin{array}{rccl} + : & \mathcal{M}_{m\times n} \times \mathbb{R} & \to & \mathcal{M}_{m\times n} \\ & +(A, k) & \to & B = A + k \quad |b_{ij} = a_{ij} + k \end{array}$$

Example

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

$$B = 1 + A = \begin{pmatrix} 2 & 3 & 4 \\ 0 & -1 & -2 \end{pmatrix}$$

MATLAB: `B=1+A`

Properties

$$\begin{array}{rcl} k + A & = & A + k \\ (k_1 + k_2) + A & = & k_1 + (k_2 + A) \end{array}$$

Definition 1.3 (Multiplication with a scalar)

*We define the **multiplication operator between a scalar and a matrix** as:*

$$\begin{array}{rccl} \cdot: & \mathbb{R} \times \mathcal{M}_{m\times n} & \to & \mathcal{M}_{m\times n} \\ & \cdot(k, A) & \to & B = k + A \quad | b_{ij} = k a_{ij} \end{array}$$

*We overload the notation to define the **multiplication operator between a matrix and a scalar** as*

$$\begin{array}{rccl} \cdot: & \mathcal{M}_{m\times n} \times \mathbb{R} & \to & \mathcal{M}_{m\times n} \\ & \cdot(A, k) & \to & B = Ak \quad | b_{ij} = a_{ij} k \end{array}$$

Example

$$\begin{aligned} A &= \begin{pmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \end{pmatrix} \\ B = 2A &= \begin{pmatrix} 2 & 4 & 6 \\ -2 & -4 & -6 \end{pmatrix} \end{aligned}$$

MATLAB: `B=2*A`

Properties

$$\begin{aligned} kA &= Ak \\ (k_1 k_2)A &= k_1(k_2 A) \\ (k_1 + k_2)A &= k_1 A + k_2 A \end{aligned}$$

Matrix operations

Definition 1.4 (Sum of two matrices)

*We define the **sum operator between two matrices** as:*

$$\begin{array}{rcl} + : \quad \mathcal{M}_{m\times n} \times \mathcal{M}_{m\times n} & \to & \mathcal{M}_{m\times n} \\ +(A,B) & \to & C = A + B \quad |c_{ij} = a_{ij} + b_{ij} \end{array}$$

Example

$$\begin{aligned} A &= \begin{pmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \end{pmatrix} \\ B &= \begin{pmatrix} 4 & 5 & 6 \\ 0 & 1 & 1 \end{pmatrix} \\ C = A + B &= \begin{pmatrix} 5 & 7 & 9 \\ -1 & -1 & -2 \end{pmatrix} \end{aligned}$$

MATLAB: `C=A+B`

Properties

$$\begin{aligned} A + B &= B + A \\ A + (B + C) &= (A + B) + C \\ A + 0 &= A \\ k(A + B) &= kA + kB \end{aligned}$$

Proof of the properties

We are not proving all properties, although all of them follow the same strategy. Let's see an example

$$k(A+B) = kA + kB$$

Proof

Let us develop the left hand side

$$\begin{array}{r|l} C = A + B & c_{ij} = a_{ij} + b_{ij} \\ D = kC = k(A+B) & d_{ij} = kc_{ij} = k(a_{ij} + b_{ij}) = ka_{ij} + kb_{ij} \end{array}$$

Now, the right hand side

$$\begin{array}{r|l} E = kA & e_{ij} = ka_{ij} \\ F = kB & f_{ij} = kb_{ij} \\ G = E + F = kA + kB & g_{ij} = e_{ij} + f_{ij} = ka_{ij} + kb_{ij} \end{array}$$

It is obvious that $d_{ij} = g_{ij}$, and consequently $k(A+B) = kA + kB$. (q.e.d.)

Definition 1.5 (Multiplication of two matrices)

*We define the **multiplication operator between two matrices** as:*

$$\begin{array}{rccl} \cdot : & \mathcal{M}_{m\times n} \times \mathcal{M}_{n\times p} & \rightarrow & \mathcal{M}_{m\times p} \\ & \cdot(A, B) & \rightarrow & C = AB \quad \left| c_{ij} = \sum\limits_{k=1}^{n} a_{ik} b_{kj} \right. \end{array}$$

If we consider the different columns of B, then we have

$$B = \begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_p \end{pmatrix} \Rightarrow AB = \begin{pmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \ldots & A\mathbf{b}_p \end{pmatrix}$$

That can be interpreted as "the j-th column of AB is a weighted sum of the columns of matrix A using the weights defined by the j-th column of B".

Example

MATLAB: `A*B`

Matrix operations

Example

Let $A = \begin{pmatrix} 2 & 3 \\ 1 & -5 \end{pmatrix}$ and $B = \begin{pmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{pmatrix}$. Then,

$$\begin{aligned}
A\mathbf{b}_1 &= \begin{pmatrix} 2 & 3 \\ 1 & -5 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 11 \\ -1 \end{pmatrix} \\
A\mathbf{b}_2 &= \begin{pmatrix} 2 & 3 \\ 1 & -5 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 13 \end{pmatrix} \\
A\mathbf{b}_3 &= \begin{pmatrix} 2 & 3 \\ 1 & -5 \end{pmatrix} \begin{pmatrix} 6 \\ 3 \end{pmatrix} = \begin{pmatrix} 21 \\ -9 \end{pmatrix} \\
AB &= \begin{pmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & A\mathbf{b}_3 \end{pmatrix} = \begin{pmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{pmatrix}
\end{aligned}$$

To directly compute a specific entry, for instance, $(AB)_{23}$ we have to multiply the 2nd row of A and the third column of B

$$(AB)_{23} = \left[\begin{pmatrix} 2 & 3 \\ 1 & -5 \end{pmatrix} \begin{pmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{pmatrix} \right] = 1 \cdot 6 + (-5) \cdot 3 = -9$$

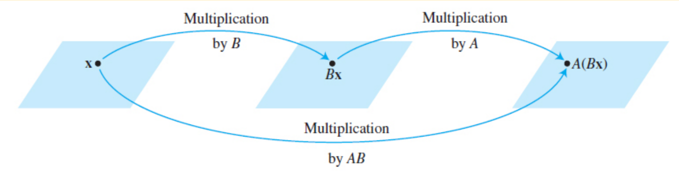
Matrix operations

Geometrical interpretation

Consider the linear transformations

$$T_A(\mathbf{x}) = A\mathbf{x}$$
$$T_B(\mathbf{x}) = B\mathbf{x}$$

that map any input vector using the matrix A or B, respectively. Now consider the sequential application of first T_B, and then T_A, as shown in the following figure:



Matrix multiplication helps us to define a single transformation such that we can transform the original vectors in a single step:

$$T_{AB}(\mathbf{x}) = (AB)\mathbf{x} = A(B\mathbf{x}) = T_A(T_B(\mathbf{x}))$$

Matrix operations

Property

$\mathrm{row}_i(AB) = \mathrm{row}_i(A)B$

Example (continued)

$$\mathrm{row}_1(AB) = \mathrm{row}_1(A)B = \begin{pmatrix} 2 & 3 \end{pmatrix} \begin{pmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{pmatrix} = \begin{pmatrix} 11 & 0 & 21 \end{pmatrix}$$

More properties

$A(BC) = (AB)C$	Associativity
$A(B+C) = AB + AC$	Left distributivity
$(A+B)C = AC + BC$	Right distributivity
$r(AB) = (rA)B = A(rB)$	For any scalar r
$I_m A = A = A I_n$	For $A \in \mathcal{M}_{m\times n}$

Matrix operations

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

Let us consider the column decomposition of matrix C.

$$C = \begin{pmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_p \end{pmatrix} \Rightarrow$$
$$BC = \begin{pmatrix} B\mathbf{c}_1 & B\mathbf{c}_2 & \dots & B\mathbf{c}_p \end{pmatrix} \Rightarrow$$
$$A(BC) = \begin{pmatrix} A(B\mathbf{c}_1) & A(B\mathbf{c}_2) & \dots & A(B\mathbf{c}_p) \end{pmatrix}$$

But we have seen in the geometrical interpretation of matrix multiplication that $A(B\mathbf{c}_i) = (AB)\mathbf{c}_i$, therefore

$$A(BC) = \begin{pmatrix} (AB)\mathbf{c}_1 & (AB)\mathbf{c}_2 & \dots & (AB)\mathbf{c}_p \end{pmatrix} = (AB)C$$

Warnings

- $AB \neq BA$, matrix multiplication is not commutative.
- $AB = AC \nRightarrow B = C$.
- $AB = 0 \nRightarrow B = 0$ or $C = 0$.

Definition 1.6 (Power of a matrix)

If $A \in \mathcal{M}_{n\times n}$, then the k-th power of the matrix is defined as

$$A^k = \underbrace{A \cdot A \cdot ... \cdot A}_{k\ times}$$

Note: $A^0 = I_n$

Example

MATLAB: `A^k`

Matrix operations

Definition 1.7 (Transpose)

If $A \in \mathcal{M}_{m\times n}$, then the transpose of A (A^T) is a matrix in $\mathcal{M}_{n\times m}$ such that the rows of A are the columns of A^T, or more formally

$$(A^T)_{ij} = A_{ji}$$

Example

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

MATLAB: `A'`

Properties

$$(A^T)^T = A$$
$$(A+B)^T = A^T + B^T$$
$$(rA)^T = rA^T$$
$$(AB)^T = B^T A^T$$

Matrix operations

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

Let $A \in \mathcal{M}_{m\times n}$ and $B \in \mathcal{M}_{n\times p}$ By the definition of matrix multiplication we know that

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Let $B' = B^T$ and $A' = A^T$. For the same reason

$$(B^T A^T)_{ij} = (B'A')_{ij} = \sum_{k=1}^{n} b'_{ik} a'_{kj}$$

But $b'_{ik} = b_{ki}$ and $a'_{kj} = a_{jk}$, consequently

$$(B^T A^T)_{ij} = \sum_{k=1}^{n} b_{ki} a_{jk} = \sum_{k=1}^{n} a_{jk} b_{ki} = (AB)_{ji}$$

or what is the same

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

Exercises

Exercises

From Lay (3rd ed.), Chapter 2, Section 1:

- 2.1.3
- 2.1.10
- 2.1.12
- 2.1.18
- 2.1.19
- 2.1.20
- 2.1.22
- 2.1.23
- 2.1.24
- 2.1.25
- 2.1.26
- 2.1.27
- 2.1.39 (bring computer)
- 2.1.40 (bring computer)

Outline

3 Matrix algebra

Matrix inverse

Example

The inverse of a number is a clear concept

$$5\tfrac{1}{5} = 5 \cdot 5^{-1} = 1 = 5^{-1} \cdot 5$$

Definition 2.1 (Inverse of a matrix)

*A matrix $A \in \mathcal{M}_{n\times n}$ is **invertible** (or **non-singular**) if there exists another matrix $C \in \mathcal{M}_{n\times n}$ such that $AC = I_n = CA$. C is called the inverse of A and it is denoted as A^{-1}. If A is not invertible, it is said to be **singular**. (MATLAB: `inv(A)`)*

Properties

The inverse of a matrix is unique.

Proof

Let us assume that there exist two different inverses C_1 and C_2. Then,

$$C_2 = C_2 I = C_2(AC_1) = (C_2 A)C_1 = IC_1 = C_1$$

which is a contradiction and, therefore, the inverse must be unique. (q.e.d.)

Matrix inverse

Example

$$\text{Let } A = \begin{pmatrix} 2 & 5 \\ -3 & -7 \end{pmatrix} \text{ and } A^{-1} = \begin{pmatrix} -7 & -5 \\ 3 & 2 \end{pmatrix}$$

It can easily be verified that

$$AA^{-1} = A^{-1}A = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Theorem 2.1 (Inverse of a 2×2 matrix)

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. *If* $ad - bc \neq 0$, *then* A *is invertible and its inverse is*

$$A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Proof
It is easy to verify that $AA^{-1} = A^{-1}A = I_2$.

Matrix inverse

Theorem 2.2

If $A \in \mathcal{M}_{n\times n}$ is invertible, then for every $b \in \mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution that is $\mathbf{x} = A^{-1}\mathbf{b}$.

Proof

Proof $\mathbf{x} = A^{-1}\mathbf{b}$ is a solution

If we substitute the solution in the equation we have

$$A\mathbf{x} = A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = \mathbf{b} \text{ (q.e.d.)}$$

Proof $\mathbf{x} = A^{-1}\mathbf{b}$ is the unique solution

Let us assume that $\mathbf{x}' \neq \mathbf{x}$ is also a solution, then

$$A\mathbf{x}' = \mathbf{b}$$

If we multiply on the left by A^{-1}, we have

$$A^{-1}A\mathbf{x}' = A^{-1}\mathbf{b} \Rightarrow \mathbf{x}' = \mathbf{x}$$

which is obviously a contradiction and, therefore, $\mathbf{x} = A^{-1}\mathbf{b}$ must be the unique solution. (q.e.d.)

Example

EXAMPLE 3 A horizontal elastic beam is supported at each end and is subjected to forces at points 1, 2, 3, as shown in Fig. 1. Let $\mathbf{f}$ in $\mathbb{R}^3$ list the forces at these points, and let $\mathbf{y}$ in $\mathbb{R}^3$ list the amounts of deflection (that is, movement) of the beam at the three points. Using Hooke's law from physics, it can be shown that

$$\mathbf{y} = D\mathbf{f}$$

where D is a *flexibility matrix*. Its inverse is called the *stiffness matrix*. Describe the physical significance of the columns of D and D^{-1}.

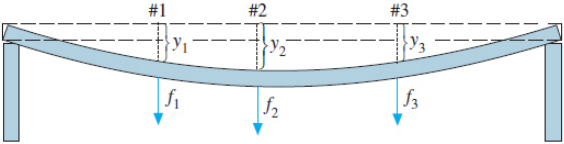


FIGURE 1 Deflection of an elastic beam.

Matrix inverse

Example (continued)

Consider the equation $\mathbf{y} = D\mathbf{f}$, $D = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & 1 \end{pmatrix}$ and the fact that

$$D = DI = \begin{pmatrix} D\mathbf{e}_1 & D\mathbf{e}_2 & D\mathbf{e}_3 \end{pmatrix}$$

Therefore, the i-th column of D can be interpreted as the deflection at the different points when a unit force ($\mathbf{e}_i$) is applied onto the i-th point. In our example when we apply a unit force at point 1, the first column of D is $(1, \frac{1}{2}, \frac{1}{4})$ meaning that the first point displaces 1 m., the second point $\frac{1}{2}$ m., and the third point $\frac{1}{4}$ m.

Matrix inverse

Example (continued)

If we now consider that $\mathbf{f} = D^{-1}\mathbf{y}$, $D^{-1} = \begin{pmatrix} \frac{4}{3} & -\frac{2}{3} & 0 \\ -\frac{2}{3} & \frac{4}{3} & -\frac{2}{3} \\ 0 & -\frac{2}{3} & \frac{4}{3} \end{pmatrix}$ and the fact that

$$D^{-1} = D^{-1}I = \begin{pmatrix} D^{-1}\mathbf{e}_1 & D^{-1}\mathbf{e}_2 & D^{-1}\mathbf{e}_3 \end{pmatrix}$$

Therefore, the i-th column of D^{-1} can be interpreted as the forces needed to be applied at the different points to produce a unit deformation ($\mathbf{e}_i$) at the i-th point. In our example, to produce a displacement of 1 m. in the first point and none at the other points ($\mathbf{e}_1 = (1, 0, 0)$, we need to push point 1 with a force of $\frac{4}{3}$ N., to pull point 2 with a force of $-\frac{2}{3}$ N., and we do not need to apply any force at point 3.

Matrix inverse

Theorem 2.3

1. *If A is invertible, then A^{-1} is also invertible and its inverse is A.*
2. *If A and B are invertible, then AB is also invertible and its inverse is $B^{-1}A^{-1}$*
3. *If A is invertible, then A^T is also invertible and its inverse is $(A^{-1})^T$.*

Proof 1)

The definition of A^{-1} is that it is a matrix such that

$$AA^{-1} = A^{-1}A = I$$

The inverse of A^{-1} must be a matrix C such that

$$CA^{-1} = A^{-1}C = I$$

If we compare this equation with the previous one, we easily see that $C = A$ is the inverse of A^{-1}.

Matrix inverse

Proof 2)

Let us check that $B^{-1}A^{-1}$ is actually the inverse of AB

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$
$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

Proof 3)

Let us check that $(A^{-1})^T$ is actually the inverse of A^T

$$A^T(A^{-1})^T = [(AB)^T = B^TA^T] = (A^{-1}A)^T = I^T = I$$
$$(A^{-1})^TA^T = [(AB)^T = B^TA^T] = (AA^{-1})^T = I^T = I$$

Theorem 2.4

We may generalize the previous theorem and state that

$$(A_1A_2...A_p)^{-1} = A_p^{-1}A_{p-1}^{-1}...A_2^{-1}A_1^{-1}$$

Matrix inverse

Proof

Let's prove it by weak induction. That is, we know that the statement is true for $p = 2$ (by the previous theorem). Let us assume it is true for $p - 1$, that is

$$(A_1 A_2 ... A_{p-1})^{-1} = A_{p-1}^{-1} ... A_2^{-1} A_1^{-1}$$

We wonder if it is also true for p. Let us define $B = A_1 A_2 ... A_{p-1}$. Then, we can rewrite the left hand side of the theorem as

$$(A_1 A_2 ... A_p)^{-1} = (BA_p)^{-1}$$

This is the inverse of the multiplication of two matrices. We know by the previous theorem that $(BA_p)^{-1} = A_p^{-1} B^{-1}$ But we presumed that

$$B^{-1} = (A_1 A_2 ... A_{p-1})^{-1} = A_{p-1}^{-1} ... A_2^{-1} A_1^{-1}$$

And consequently

$$(BA_p)^{-1} = A_p^{-1} A_{p-1}^{-1} ... A_2^{-1} A_1^{-1} \text{ (q.e.d.)}$$

Outline

3. Matrix algebra
   - Matrix operations (a)
   - Inverse of a matrix (b)
   - Elementary matrices (b)
   - An algorithm to invert matrices (b)
   - Characterization of invertible matrices (c)
   - Invertible linear transformations (c)
   - Partitioned matrices (c)
   - LU factorization (d)
   - An application to computer graphics and image processing (d)
   - Subspaces of $\mathbb{R}^n$ (e)
   - Dimension and rank (e)

Elementary matrices

The elementary operations we can perform on the rows of a matrix are

1. Multiply by a scalar
2. Swap two rows
3. Replace a row by a linear combination of two or several rows

All these operations can be represented as matrix multiplications.

Example

Consider the matrix $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$

1. We can multiply the third row by a scalar r by multiplying on the left by the matrix

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & r \end{pmatrix} \Rightarrow E_1 A = \begin{pmatrix} a & b & c \\ d & e & f \\ rg & rh & ri \end{pmatrix}$$

Elementary matrices

Example (continued)

2. We can swap the first and second rows of the matrix by multiplying on the left by the matrix

$$E_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow E_2A = \begin{pmatrix} d & e & f \\ a & b & c \\ g & h & i \end{pmatrix}$$

3. We can substitute the third row by $\mathbf{r}_3 + k_1\mathbf{r}_1$ by multiplying on the left by the matrix

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k_1 & 0 & 1 \end{pmatrix} \Rightarrow E_3A = \begin{pmatrix} a & b & c \\ d & e & f \\ g + k_1 a & h + k_1 b & i + k_1 c \end{pmatrix}$$

Definition 3.1 (Elementary matrix)

*An **elementary matrix** is one that differs from the identity matrix by one single, elementary row operation.*

Elementary matrices

Theorem 3.1

The inverse of an elementary matrix is another elementary matrix of the same type. That is, row operations can be undone.

Example (continued)

1. $E_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{r} \end{pmatrix}$
2. $E_2^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
3. $E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -k_1 & 0 & 1 \end{pmatrix}$

MATLAB:

```
syms k1
E3=[1 0 0; 0 1 0; k1 0 1];
inv(E3)
```

Elementary matrices

Theorem 3.2

A matrix $A \in \mathcal{M}_{n\times n}$ is invertible iff it is row-equivalent to I_n. In this case, the sequence of operations that transforms A into I_n is also the one that transforms I_n into A^{-1}.

Proof $\Rightarrow$

If A is invertible, then by theorem 2.2 we know that the equation system $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$. If it has a solution for every $\mathbf{b}$, then it must have a pivot in every row, that must be in the diagonal and, consequently the reduced echelon form of A must be I_n.

Proof $\Leftarrow$

If A is row-equivalent I_n, then there exists a sequence of elementary matrices that transform A into I_n

$$A \sim E_1A \sim E_2E_1A \sim ... \sim E_nE_{n-1}...E_2E_1A = I_n$$

$E = E_nE_{n-1}...E_2E_1$ is a candidate to be the inverse of A. Since each of the elementary matrices is invertible, and the product of invertible matrices is invertible, then E is invertible and A must be its (unique) inverse. Conversely, E is the inverse of A and A is invertible.

Outline

3. Matrix algebra
   - Matrix operations (a)
   - Inverse of a matrix (b)
   - Elementary matrices (b)
   - **An algorithm to invert matrices (b)**
   - Characterization of invertible matrices (c)
   - Invertible linear transformations (c)
   - Partitioned matrices (c)
   - LU factorization (d)
   - An application to computer graphics and image processing (d)
   - Subspaces of $\mathbb{R}^n$ (e)
   - Dimension and rank (e)

An algorithm to invert matrices

Algorithm

Algorithm: Reduce the augmented matrix $\left(\begin{array}{c|c} A & I \end{array} \right)$
If A is invertible, then $\left(\begin{array}{c|c} A & I \end{array} \right) \sim \left(\begin{array}{c|c} I & A^{-1} \end{array} \right)$.
If A is not invertible, then we will not be able to reduce A into I.
This algorithm is very much used in practice because it is numerically stable and rather efficient.

Example

Let $A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{pmatrix}$.
We construct the augmented matrix

$$\left(\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array} \right)$$

An algorithm to invert matrices

Example (continued)

And now we transform it

$$\left(\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array}\right)$$

$$\mathbf{r}_1 \leftrightarrow \mathbf{r}_2 \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array}\right)$$

$$\mathbf{r}_3 \leftarrow \mathbf{r}_3 - 4\mathbf{r}_1 \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{array}\right)$$

$$\mathbf{r}_3 \leftarrow \mathbf{r}_3 + 3\mathbf{r}_2 \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{array}\right)$$

$$\mathbf{r}_3 \leftarrow \tfrac{1}{2}\mathbf{r}_3 \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{3}{2} & -2 & \frac{1}{2} \end{array}\right)$$

An algorithm to invert matrices

Example (continued)

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{3}{2} & -2 & \frac{1}{2} \end{array}\right)$$

$$\mathbf{r}_2 \leftarrow \mathbf{r}_2 - 2\mathbf{r}_3 \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & \frac{3}{2} & -2 & \frac{1}{2} \end{array}\right)$$

$$\mathbf{r}_1 \leftarrow \mathbf{r}_1 - 3\mathbf{r}_3 \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{9}{2} & 7 & -\frac{3}{2} \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & \frac{3}{2} & -2 & \frac{1}{2} \end{array}\right)$$

Since A is row-equivalent to I_3, then A is invertible and its inverse is $A^{-1} = \begin{pmatrix} -\frac{9}{2} & 7 & -\frac{3}{2} \\ -2 & 4 & -1 \\ \frac{3}{2} & -2 & \frac{1}{2} \end{pmatrix}$. To finalize the exercise we should check that

$$AA^{-1} = A^{-1}A = I_3$$

An algorithm to invert matrices

A new interpretation of matrix inversion

By constructing the augmented matrix $\left(\begin{array}{c|c} A & I \end{array}\right)$ we are simultaneously solving multiple equation systems

$$A\mathbf{x} = \mathbf{e}_1 \quad A\mathbf{x} = \mathbf{e}_2 \quad A\mathbf{x} = \mathbf{e}_3 \quad \ldots$$

Example (continued)

$$\left(\begin{array}{ccc|c} 0 & 1 & 2 & 1 \\ 1 & 0 & 3 & 0 \\ 4 & -3 & 8 & 0 \end{array}\right) \quad \left(\begin{array}{ccc|c} 0 & 1 & 2 & 0 \\ 1 & 0 & 3 & 1 \\ 4 & -3 & 8 & 0 \end{array}\right) \quad \left(\begin{array}{ccc|c} 0 & 1 & 2 & 0 \\ 1 & 0 & 3 & 0 \\ 4 & -3 & 8 & 1 \end{array}\right)$$

This note is important because if we want to compute only the i-th column of A^{-1} it is enough to solve the equation system

$$A\mathbf{x} = \mathbf{e}_i$$

Exercises

Exercises

From Lay (3rd ed.), Chapter 2, Section 2:

- 2.2.7
- 2.2.11
- 2.2.13
- 2.2.17
- 2.2.19
- 2.2.21
- 2.2.25
- 2.2.36

Outline

3. Matrix algebra
   - Matrix operations (a)
   - Inverse of a matrix (b)
   - Elementary matrices (b)
   - An algorithm to invert matrices (b)
   - Characterization of invertible matrices (c)
   - Invertible linear transformations (c)
   - Partitioned matrices (c)
   - LU factorization (d)
   - An application to computer graphics and image processing (d)
   - Subspaces of $\mathbb{R}^n$ (e)
   - Dimension and rank (e)

Characterization of invertible matrices

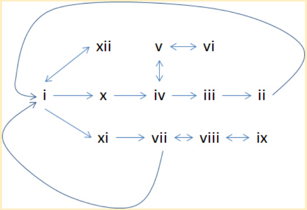
Theorem 5.1 (The invertible matrix theorem)

Let $A \in \mathcal{M}_{n\times n}$. The following statements are equivalent (either they are all true or they are all false):

i. *A is invertible.*
ii. *A is row-equivalent to I_n.*
iii. *A has n pivot positions.*
iv. *$A\mathbf{x} = \mathbf{0}$ only has the trivial solution $\mathbf{x} = \mathbf{0}$.*
v. *The columns of A are linearly independent.*
vi. *The transformation $T(\mathbf{x}) = A\mathbf{x}$ is injective.*
vii. *The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for every $\mathbf{b} \in \mathbb{R}^n$.*
viii. *The columns of A span $\mathbb{R}^n$.*
ix. *The transformation $T(\mathbf{x}) = A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.*
x. *There exists a matrix $C \in \mathcal{M}_{n\times n}$ such that $CA = I_n$.*
xi. *There exists a matrix $D \in \mathcal{M}_{n\times n}$ such that $AD = I_n$.*
xii. *A^T is an invertible matrix*

Characterization of invertible matrices

To prove the theorem we will follow the reasoning scheme below:



Proof i $\Rightarrow$ x

If i is true, then x is true simply by doing $C = A^{-1}$.

Proof x $\Rightarrow$ iv

See Exercise 2.1.23 in Lay.

Proof iv $\Rightarrow$ iii

See Exercise 2.2.23 in Lay.

Proof iii $\Rightarrow$ ii

If iii is true, then the n pivots have to be in the main diagonal and in this case, the reduced echelon form must be I_n.

Characterization of invertible matrices

Proof ii $\Rightarrow$ i
If ii is true, then i is true thanks to Theorem 3.2.
Proof i $\Rightarrow$ xi
If i is true, then xi is true simply by doing $D = A^{-1}$.
Proof xi $\Rightarrow$ vii
See Exercise 2.1.24 in Lay.
Proof vii $\Rightarrow$ i
See Exercise 2.2.24 in Lay.
Proof vii $\Leftrightarrow$ viii $\Leftrightarrow$ ix
See Theorems 3.2 and 8.2 in Chapter 2.
Proof iv $\Leftrightarrow$ v $\Leftrightarrow$ vi
See Theorems 3.2, 5.1 and 8.1 in Chapter 2.
Proof i $\Rightarrow$ xii
See Theorem 2.3.
Proof i $\Leftarrow$ xii
See Theorem 2.3 interchanging the roles of A and A^T.

The power of this theorem is that it connects equation systems to invertibility, linear independence and subspace bases.

Characterization of invertible matrices

Corollary

1. If A is invertible, then $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b} \in \mathbb{R}^n$.
2. If $A, B \in \mathcal{M}_{n\times n}$ and $AB = I_n$, then A and B are invertible and $B = A^{-1}$ and $A = B^{-1}$.

Watch out that this corollary only applies to square matrices.

Outline

Invertible linear transformations

Consider the linear transformation

$$\begin{array}{rcl} T:\mathbb{R}^n & \to & \mathbb{R}^n \\ \mathbf{x} & \to & A\mathbf{x} \end{array}$$

Definition 6.1 (Inverse transformation)

T is invertible iff there exists $S:\mathbb{R}^n \to \mathbb{R}^n$ such that $\forall \mathbf{x} \in \mathbb{R}^n$:

$$S(T(\mathbf{x})) = \mathbf{x} = T(S(\mathbf{x}))$$

Example

$T(\mathbf{x}) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x}$ is invertible and its inverse is $S(\mathbf{x}) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x}$.

Proof

$$S(T(\mathbf{x})) = S\left(\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\mathbf{x}\right) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\mathbf{x} = \mathbf{x}$$

$$T(S(\mathbf{x})) = T\left(\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\mathbf{x}\right) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\mathbf{x} = \mathbf{x}$$

Invertible linear transformations

Example

$T(\mathbf{x}) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{x}$ is not invertible because $T((1,0)) = T((1,1)) = (1,0)$, so given the "output" (1,0), we cannot recover the input vector that originated this output.

Theorem 6.1

If T is invertible, then it is surjective.

Proof

Consider any vector $\mathbf{b} \in \mathbb{R}^n$, we can always apply the transformation S to get a new vector $\mathbf{x} = S(\mathbf{b})$. And then, recover $\mathbf{b}$ making use of the fact that T is the inverse of S, that is, $\mathbf{b} = T(\mathbf{x})$. In other words, any vector $\mathbf{b}$ is in the range of T and, therefore, T is surjective.

Invertible linear transformations

Theorem 6.2

T is invertible iff A is invertible. If T is invertible, then the only function that satisfies the previous definition is

$$S(\mathbf{x}) = A^{-1}\mathbf{x}$$

Proof $\Rightarrow$

If T is invertible, then it is surjective (see previous Theorem). Then, A is invertible by Theorem 5.1 (items i and ix).

Proof $\Leftarrow$

If A is invertible, then we may construct the linear transformation $S = A^{-1}\mathbf{x}$. S is an inverse of T since

$$S(T(\mathbf{x})) = S(A\mathbf{x}) = A^{-1}(A\mathbf{x}) = (A^{-1}A)\mathbf{x} = \mathbf{x}$$
$$T(S(\mathbf{x})) = T(A^{-1}\mathbf{x}) = A(A^{-1}\mathbf{x}) = (AA^{-1})\mathbf{x} = \mathbf{x}$$

Invertible linear transformations

Proof uniqueness

Let us assume that there are two inverses $S_1(\mathbf{x}) = B_1\mathbf{x}$ and $S_2(\mathbf{x}) = B_2\mathbf{x}$ with $B_1 \neq B_2$. Let $\mathbf{v} \in \mathbb{R}^n$ and $\mathbf{v} = T(\mathbf{x})$ for some $\mathbf{x} \in \mathbb{R}^n$ (since T is invertible and, therefore, surjective, we are guaranteed that there exists at least one such $\mathbf{x}$). Now

$$\left.\begin{array}{l} S_1(\mathbf{v}) = B_1A\mathbf{x} = \mathbf{x} = B_1\mathbf{v} \\ S_2(\mathbf{v}) = B_2A\mathbf{x} = \mathbf{x} = B_2\mathbf{v} \end{array}\right\} \Rightarrow B_1\mathbf{v} = B_2\mathbf{v}\ [\forall \mathbf{v} \in \mathbb{R}^n] \Rightarrow B_1 = B_2$$

which is a contradiction and, consequently, there exists only one inverse (q.e.d.)

Definition 6.2 (Ill-conditioned matrix)

*Informally, we say that a matrix A is **ill-conditioned** if it is "nearly singular". In practice, this implies that the equation system $A\mathbf{x} = \mathbf{b}$ may have large variations in the solution ($\mathbf{x}$) when $\mathbf{b}$ varies slightly.*

Exercises

Exercises

From Lay (3rd ed.), Chapter 2, Section 3:

- 2.3.13
- 2.3.16
- 2.3.17
- 2.3.33
- 2.3.41

Outline

3 Matrix algebra
- Matrix operations (a)
- Inverse of a matrix (b)
- Elementary matrices (b)
- An algorithm to invert matrices (b)
- Characterization of invertible matrices (c)
- Invertible linear transformations (c)
- Partitioned matrices (c)
- LU factorization (d)
- An application to computer graphics and image processing (d)
- Subspaces of $\mathbb{R}^n$ (e)
- Dimension and rank (e)

Partitioned matrices

Partitioned matrices sometimes help us to gain insight into the structure of the problem by identifying blocks within the matrix.

Example

$$A = \left(\begin{array}{ccc|cc|c} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -3 & 1 \\ \hline -8 & -6 & 3 & 1 & 7 & -4 \end{array}\right) = \left(\begin{array}{c|c|c} A_{11} & A_{12} & A_{13} \\ \hline A_{21} & A_{22} & A_{23} \end{array}\right)$$

$A \in \mathcal{M}_{3\times 6}$,
$A_{11} \in \mathcal{M}_{2\times 3}$, $A_{12} \in \mathcal{M}_{2\times 2}$, $A_{13} \in \mathcal{M}_{2\times 1}$,
$A_{21} \in \mathcal{M}_{1\times 3}$, $A_{22} \in \mathcal{M}_{1\times 2}$, $A_{23} \in \mathcal{M}_{1\times 1}$.
MATLAB:

```
A=[3 0 -1 5 9 -2; -5 2 4 0 -3 1; -8 -6 3 1 7 -4];
A11=A(1:2,1:3)
A12=A(1:2,4:5)
A13=A(1:2,6)
A21=A(3,1:3)
A22=A(3,4:5)
A23=A(3,6)
```

Partitioned matrices

Definition 7.1 (Sum of partitioned matrices)

Let A and B be two matrices partitioned in the same way. Then the blocks of A + B are simply the sum of the corresponding blocks.

$$A + B = \left(\begin{array}{c|c|c} & & \\ \hline & A_{ij} & \\ \hline & & \end{array}\right) + \left(\begin{array}{c|c|c} & & \\ \hline & B_{ij} & \\ \hline & & \end{array}\right) = \left(\begin{array}{c|c|c} & & \\ \hline & A_{ij} + B_{ij} & \\ \hline & & \end{array}\right)$$

Definition 7.2 (Multiplication by scalar)

The multiplication by a scalar simply multiplies each one of the blocks independently

$$rA = r\left(\begin{array}{c|c|c} & & \\ \hline & A_{ij} & \\ \hline & & \end{array}\right) = \left(\begin{array}{c|c|c} & & \\ \hline & rA_{ij} & \\ \hline & & \end{array}\right)$$

Partitioned matrices

Definition 7.3 (Multiplication of partitioned matrices)

Multiply the different block as if they were scalars (but applying matrix multiplication).

Example

Let $A = \left(\begin{array}{ccc|cc} 2 & -3 & 1 & 0 & -4 \\ 1 & 5 & -2 & 3 & -1 \\ \hline 0 & -4 & -2 & 7 & -1 \end{array}\right) = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$

and $B = \left(\begin{array}{cc} 6 & 4 \\ -2 & 1 \\ -3 & 7 \\ \hline -1 & 3 \\ 5 & 2 \end{array}\right) = \left(\begin{array}{c} B_1 \\ \hline B_2 \end{array}\right)$.

Then, $AB = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{pmatrix} = \begin{pmatrix} -5 & 4 \\ -6 & 2 \\ 2 & 1 \end{pmatrix}$

Partitioned matrices

Theorem 7.1 (Multiplication of matrices)

Let $A \in \mathcal{M}_{m\times n}$ and $B \in \mathcal{M}_{n\times p}$, then

$$AB = \sum_{k=1}^{n} \mathrm{column}_k(A)\mathrm{row}_k(B)$$

Proof

Let us analyze each one of the terms in the sum

$$\mathrm{column}_k(A)\mathrm{row}_k(B) = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \dots \\ a_{mk} \end{pmatrix} \begin{pmatrix} b_{k1} & b_{k2} & \dots & b_{kp} \end{pmatrix} =$$

$$\begin{pmatrix} a_{1k}b_{k1} & a_{1k}b_{k2} & \dots & a_{1k}b_{kp} \\ a_{2k}b_{k1} & a_{2k}b_{k2} & \dots & a_{2k}b_{kp} \\ \dots & \dots & \dots & \dots \\ a_{mk}b_{k1} & a_{mk}b_{k2} & \dots & a_{mk}b_{kp} \end{pmatrix}$$

Partitioned matrices

In general, the *ij*-th term is

$$(\mathrm{column}_k(A)\mathrm{row}_k(B))_{ij} = a_{ik}b_{kj}$$

If we now analyze the *ij*-th element of the sum

$$\left(\sum_{k=1}^{n} \mathrm{column}_k(A)\mathrm{row}_k(B)\right)_{ij} = \sum_{k=1}^{n} (\mathrm{column}_k(A)\mathrm{row}_k(B))_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

But this is the definition of matrix multiplication and, therefore,

$$\left(\sum_{k=1}^{n} \mathrm{column}_k(A)\mathrm{row}_k(B)\right)_{ij} = (AB)_{ij} \text{ (q.e.d.)}$$

Partitioned matrices

Definition 7.4 (Transpose of partitioned matrices)

Transpose the partitioned matrix as if it were composed of scalars, and transpose each one of the blocks.

Example

$$A = \left(\begin{array}{c|c|c} A_{11} & A_{12} & A_{13} \\ \hline A_{21} & A_{22} & A_{23} \\ \hline A_{31} & A_{32} & A_{33} \end{array}\right) \Rightarrow A^T = \left(\begin{array}{c|c|c} A_{11}^T & A_{21}^T & A_{31}^T \\ \hline A_{12}^T & A_{22}^T & A_{32}^T \\ \hline A_{13}^T & A_{23}^T & A_{33}^T \end{array}\right)$$

Example

$$A = \left(\begin{array}{ccc|cc} 2 & -3 & 1 & 0 & -4 \\ 1 & 5 & -2 & 3 & -1 \\ \hline 0 & -4 & -2 & 7 & -1 \end{array}\right) \Rightarrow A^T = \left(\begin{array}{cc|c} 2 & 1 & 0 \\ -3 & 5 & -4 \\ 1 & -2 & -2 \\ \hline 0 & 3 & 7 \\ -4 & -1 & -1 \end{array}\right)$$

Partitioned matrices

Definition 7.5 (Inverse of partitioned matrices)

The formula for each one of the cases is worked out particularly for that case. Here go a couple of examples.

Example

Let $A = \left(\begin{array}{c|c|c} A_{11} & 0 & 0 \\ \hline 0 & A_{22} & 0 \\ \hline 0 & 0 & A_{33} \end{array}\right)$.

$A \in \mathcal{M}_{n\times n}$, $A_{11} \in \mathcal{M}_{p\times p}$, $A_{22} \in \mathcal{M}_{q\times q}$, $A_{33} \in \mathcal{M}_{r\times r}$ such that $p+q+r=n$.
We look for a matrix B such that

$$\left(\begin{array}{c|c|c} A_{11} & 0 & 0 \\ \hline 0 & A_{22} & 0 \\ \hline 0 & 0 & A_{33} \end{array}\right) \left(\begin{array}{c|c|c} B_{11} & B_{12} & B_{13} \\ \hline B_{21} & B_{22} & B_{23} \\ \hline B_{31} & B_{32} & B_{33} \end{array}\right) = \left(\begin{array}{c|c|c} I_p & 0 & 0 \\ \hline 0 & I_q & 0 \\ \hline 0 & 0 & I_r \end{array}\right) \Rightarrow$$

$$\left(\begin{array}{c|c|c} A_{11}B_{11} & A_{11}B_{12} & A_{11}B_{13} \\ \hline A_{22}B_{21} & A_{22}B_{22} & A_{22}B_{23} \\ \hline A_{33}B_{31} & A_{33}B_{32} & A_{33}B_{33} \end{array}\right) = \left(\begin{array}{c|c|c} I_p & 0 & 0 \\ \hline 0 & I_q & 0 \\ \hline 0 & 0 & I_r \end{array}\right)$$

Partitioned matrices

Example (continued)

For each one of the entries we have a set of equations:

$$
\begin{aligned}
&\forall A_{11} \in \mathcal{M}_{p\times p}\ A_{11}B_{11} = I_p \Rightarrow B_{11} = A_{11}^{-1}\\
&\forall A_{11} \in \mathcal{M}_{p\times p}\ A_{11}B_{12} = 0 \Rightarrow B_{12} = 0\\
&\forall A_{11} \in \mathcal{M}_{p\times p}\ A_{11}B_{13} = 0 \Rightarrow B_{13} = 0\\
&\forall A_{22} \in \mathcal{M}_{q\times q}\ A_{22}B_{21} = 0 \Rightarrow B_{21} = 0\\
&\forall A_{22} \in \mathcal{M}_{q\times q}\ A_{22}B_{22} = I_q \Rightarrow B_{22} = A_{22}^{-1}\\
&\forall A_{22} \in \mathcal{M}_{q\times q}\ A_{22}B_{23} = 0 \Rightarrow B_{23} = 0\\
&\forall A_{33} \in \mathcal{M}_{r\times r}\ A_{33}B_{31} = 0 \Rightarrow B_{31} = 0\\
&\forall A_{33} \in \mathcal{M}_{r\times r}\ A_{33}B_{32} = 0 \Rightarrow B_{32} = 0\\
&\forall A_{33} \in \mathcal{M}_{r\times r}\ A_{33}B_{33} = I_r \Rightarrow B_{33} = A_{33}^{-1}
\end{aligned}
$$

Finally,

$$
B = \left(\begin{array}{c|c|c} A_{11}^{-1} & 0 & 0 \\ \hline 0 & A_{22}^{-1} & 0 \\ \hline 0 & 0 & A_{33}^{-1} \end{array}\right)
$$

Example

Let $A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right)$.

$A \in \mathcal{M}_{n\times n}$, $A_{11} \in \mathcal{M}_{p\times p}$, $A_{12} \in \mathcal{M}_{p\times q}$, $A_{22} \in \mathcal{M}_{q\times q}$ such that $p + q = n$.

We look for a matrix B such that

$$= \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right) \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right) = \left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & I_q \end{array}\right) \Rightarrow$$

$$\left(\begin{array}{c|c} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ \hline A_{22}B_{21} & A_{22}B_{22} \end{array}\right) = \left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & I_q \end{array}\right)$$

Partitioned matrices

Example (continued)

For each one of the entries we have a set of equations:

$$
\begin{array}{ll}
\forall A_{22} \in \mathcal{M}_{q\times q} \; A_{22}B_{21} = 0 \Rightarrow B_{21} = 0 & \\
\forall A_{22} \in \mathcal{M}_{q\times q} \; A_{22}B_{22} = I_q \Rightarrow B_{22} = A_{22}^{-1} & \\
\forall A_{11} \in \mathcal{M}_{q\times q}, A_{12} \in \mathcal{M}_{p\times q} & A_{11}B_{11} + A_{12}B_{21} = I_p \Rightarrow [B_{21} = 0] \Rightarrow \\
 & A_{11}B_{11} = I_p \Rightarrow B_{11} = A_{11}^{-1} \\
\forall A_{11} \in \mathcal{M}_{q\times q}, A_{12} \in \mathcal{M}_{p\times q} & A_{11}B_{12} + A_{12}B_{22} = 0 \Rightarrow [B_{22} = A_{22}^{-1}] \Rightarrow \\
 & A_{11}B_{12} + A_{12}A_{22}^{-1} = 0 \Rightarrow A_{11}B_{12} = -A_{12}A_{22}^{-1} \Rightarrow \\
 & B_{12} = -A_{11}^{-1}A_{12}A_{22}^{-1}
\end{array}
$$

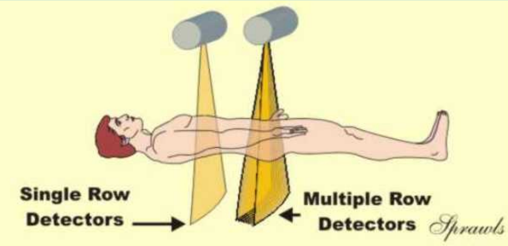
Finally,

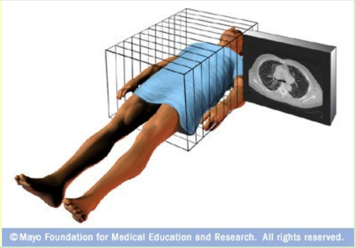
$$
B = \left(\begin{array}{c|c} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ \hline 0 & A_{22}^{-1} \end{array} \right)
$$

Partitioned matrices

Example

Computational Tomography (CT) with multiple rows gives a non-block structure for the system matrix that forces the problem to be solved in 3D. However, with a single row detector, the system matrix has a block structure so that the problem can be solved as a series of 2D problems strongly accelerating the process (on the other side the redundancy introduced by multiple row offers better resolution and robustness to noise).





Exercises

Exercises

From Lay (3rd ed.), Chapter 2, Section 4:

- 2.4.15
- 2.4.16
- 2.4.18
- 2.4.19

Outline

LU factorization

Example

Let us presume that we have a collection of equation systems

$$A\mathbf{x} = \mathbf{b}_1$$
$$A\mathbf{x} = \mathbf{b}_2$$
$$\ldots$$

and A is not invertible, which could be an efficient way of solving all of them together? Factorize A as $A = LU$ (see below) and solve the equation system in two steps. In fact the method is so efficient it is even used to solve a single equation system.

Definition 8.1 (LU factorization)

Let $A \in \mathcal{M}_{m\times n}$ that can be reduced to a reduced echelon form without row permutations. We can factorize A as $A = LU$, where L is an invertible, lower triangular matrix (with 1s in the main diagonal) of size $m \times m$ and U is an upper triangular matrix of size $m \times n$.
MATLAB: `[L,U]=lu(A)`

Example

Let $A \in \mathcal{M}_{4\times 5}$. LU factorization will produce two matrices L and U may be of the following structure

$$A = LU = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \heartsuit & 1 & 0 & 0 \\ \heartsuit & \heartsuit & 1 & 0 \\ \heartsuit & \heartsuit & \heartsuit & 1 \end{pmatrix} \begin{pmatrix} \Diamond & \heartsuit & \heartsuit & \heartsuit & \heartsuit \\ 0 & \Diamond & \heartsuit & \heartsuit & \heartsuit \\ 0 & 0 & 0 & \Diamond & \heartsuit \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

LU factorization

Solving a linear equation system using the LU decomposition

Consider the equation system $A\mathbf{x} = \mathbf{b}$, and assume we have decomposed A as $A = LU$. Then, we can solve the equation system in two steps:

$$A\mathbf{x} = \mathbf{b} \Rightarrow (LU)\mathbf{x} = L(U\mathbf{x}) = \mathbf{b} \Rightarrow \begin{cases} L\mathbf{y} = \mathbf{b} \\ U\mathbf{x} = \mathbf{y} \end{cases}$$

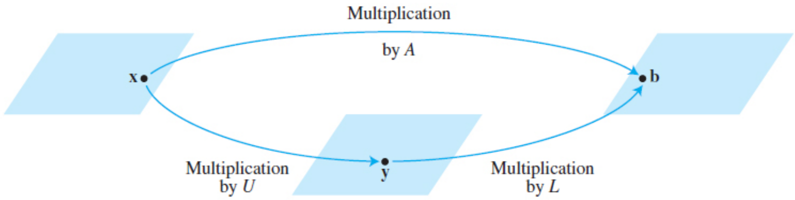


FIGURE 2 Factorization of the mapping $\mathbf{x} \mapsto A\mathbf{x}$.

LU factorization

Example

Consider

$$A = \begin{pmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{pmatrix} \begin{pmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and $\mathbf{b} = (-9, 5, 7, 11)$. We first solve $L\mathbf{y} = \mathbf{b}$

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -9 \\ -1 & 1 & 0 & 0 & 5 \\ 2 & -5 & 1 & 0 & 7 \\ -3 & 8 & 3 & 1 & 11 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

and now we solve $U\mathbf{x} = \mathbf{y}$

$$\left(\begin{array}{cccc|c} 3 & -7 & -2 & 2 & -9 \\ 0 & -2 & -1 & 2 & -4 \\ 0 & 0 & -1 & 1 & 5 \\ 0 & 0 & 0 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & -1 \end{array}\right)$$

The trick is that, thanks to the triangular structure, solving these two equation systems is rather fast.

An algorithm to simple LU factorizations

Algorithm

Let us assume that A is row-equivalent to U only using row replacement only with the rows above the replaced row. Then, there must be a sequence of elementary matrices such that

$$A \sim U \Rightarrow E_p \ldots E_2 E_1 A = U \Rightarrow A = (E_p \ldots E_2 E_1)^{-1} U$$

By inspection, we note that $L = (E_p \ldots E_2 E_1)^{-1}$.

In the previous algorithm we are making using of the following theorem:

Theorem 8.1

1. *The product of two lower triangular matrices is lower triangular.*
2. *The inverse of a lower triangular matrix is lower triangular.*

An algorithm to simple LU factorizations

Example

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

$$\mathbf{r}_2 \leftarrow \mathbf{r}_2 - \tfrac{1}{2}\mathbf{r}_1 \quad E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 & 0 \\ 0 & \frac{3}{2} & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

$$\mathbf{r}_3 \leftarrow \mathbf{r}_3 - \tfrac{2}{3}\mathbf{r}_2 \quad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{pmatrix} \quad U = \begin{pmatrix} 2 & 1 & 0 \\ 0 & \frac{3}{2} & 1 \\ 0 & 0 & \frac{4}{3} \end{pmatrix}$$

Now, we calculate L as

$$L = (E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{2}{3} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & \frac{2}{3} & 1 \end{pmatrix}$$

An algorithm to simple LU factorizations

Example

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

$$\mathbf{r}_2 \leftarrow \mathbf{r}_2 - \tfrac{1}{2}\mathbf{r}_1 \quad E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 2 & 1 & 0 \\ 0 & \frac{3}{2} & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

$$\mathbf{r}_3 \leftarrow \mathbf{r}_3 - \tfrac{2}{3}\mathbf{r}_2 \quad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{pmatrix} \qquad U = \begin{pmatrix} 2 & 1 & 0 \\ 0 & \frac{3}{2} & 1 \\ 0 & 0 & \frac{4}{3} \end{pmatrix}$$

Now, we calculate L as

$$L = (E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{2}{3} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & \frac{2}{3} & 1 \end{pmatrix}$$

LDU factorization

Example (continued)

Note that the L and U matrices found so far are assymetric in the sense that L has 1s in its main diagonal, but U has not. We can extract the elements in the main diagonal of U to a separate matrix D by simply dividing the corresponding row of U by that element:

$$\begin{aligned} A &= LU = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & \frac{2}{3} & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & \frac{3}{2} & 1 \\ 0 & 0 & \frac{4}{3} \end{pmatrix} \\ &= LDU = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & \frac{2}{3} & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & \frac{4}{3} \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 1 & \frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

where D is always a diagonal matrix.

Other factorization examples

Other factorizations

There are many other possibilities to factorize a matrix $A \in \mathcal{M}_{m \times n}$. See http://en.wikipedia.org/wiki/Matrix_decomposition. Among the most important are:

- QR: $A = QR$ where $Q \in \mathcal{M}_{m \times m}$ is orthogonal ($Q^t Q = D$) and $R \in \mathcal{M}_{m \times n}$ is upper triangular.
- SVD: $A = UDV^t$ where $U \in \mathcal{M}_{m \times m}$ is unitary ($U^t U = I_m$), $D \in \mathcal{M}_{m \times n}$ is diagonal, and $V \in \mathcal{M}_{n \times n}$ is also unitary ($V^t V = I_n$).
- Spectral: $A = PDP^{-1}$ (only for square matrices) where $P \in \mathcal{M}_{n \times n}$ and $D \in \mathcal{M}_{n \times n}$ is diagonal.

Exercises

Exercises

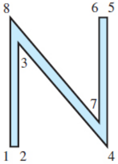
From Lay (3rd ed.), Chapter 2, Section 5:

- 2.5.9
- 2.5.Practice problem

Outline

Example

In vectorial graphics, graphics are described as a set of connected points (whose coordinates are known).

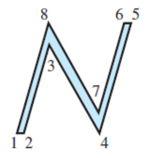


EXAMPLE 1 The capital letter N in Fig. 1 is determined by eight points, or *vertices*. The coordinates of the points can be stored in a data matrix, D.

$$\begin{array}{r} \text{Vertex:} \\ \\ x\text{-coordinate} \\ y\text{-coordinate} \end{array} \begin{array}{c} \\ \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{matrix} \\ \begin{bmatrix} 0 & .5 & .5 & 6 & 6 & 5.5 & 5.5 & 0 \\ 0 & 0 & 6.42 & 0 & 8 & 8 & 1.58 & 8 \end{bmatrix} = D \end{array}$$

In addition to D, it is necessary to specify which vertices are connected by lines, but we omit this detail. ■

We may produce "italic" fonts by shearing the standard coordinates $T(\mathbf{x}) = A\mathbf{x}$ where $A = \begin{pmatrix} 1 & 0.25 \\ 0 & 1 \end{pmatrix}$.



An application to computer graphics and image processing

Example

Coordinate translations can be expressed as $T(\mathbf{x}) = \mathbf{x} + \mathbf{x}_0$. But this is not a linear transformation:

$$\begin{array}{rcl} T(\mathbf{u}) & = & \mathbf{u} + \mathbf{x}_0 \\ T(\mathbf{v}) & = & \mathbf{v} + \mathbf{x}_0 \\ T(\mathbf{u}+\mathbf{v}) & = & \mathbf{u} + \mathbf{v} + \mathbf{x}_0 \\ T(\mathbf{u}) + T(\mathbf{v}) & = & (\mathbf{u} + \mathbf{x}_0) + (\mathbf{v} + \mathbf{x}_0) = \mathbf{u} + \mathbf{v} + 2\mathbf{x}_0 \\ T(\mathbf{u}+\mathbf{v}) & \neq & T(\mathbf{u}) + T(\mathbf{v}) \end{array}$$

We can solve this problem with homogeneous coordinates.

An application to computer graphics and image processing

Definition 9.1 (Homogeneous coordinates)

Given a point with coordinates $\mathbf{x}$ *we can construct its* ***homogeneous coordinates*** *as*

$$\tilde{\mathbf{x}} = \begin{pmatrix} h\mathbf{x} \\ h \end{pmatrix}$$

Or in other words, given the homogeneous coordinates $\tilde{\mathbf{u}} = \begin{pmatrix} \mathbf{u} \\ h \end{pmatrix}$*, they represent the point at* $\frac{\mathbf{u}}{h}$*. It is customary to use* $h = 1$ *(but it is not compulsory, and in certain applications it is better to use other* h*'s).*

Example

The 2D point $(1, 2)$ can be represented in homogeneous coordinates as $(1, 2, 1)$, as $(2, 4, 2)$ and, even, as $(-2, -4, -2)$. They all represent the same point.

Example

Now, coordinate translations in homogeneous coordinates is a linear transformation. For instance, in 2D:

$$T(\tilde{\mathbf{x}}) = A\tilde{\mathbf{x}} = \begin{pmatrix} 1 & 0 & \Delta x \\ 0 & 1 & \Delta y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x + \Delta x \\ y + \Delta y \\ 1 \end{pmatrix}$$

2D transformations in homogeneous coordinates

In general, any 2D transformation of the form $T(\mathbf{x}) = A\mathbf{x}$ can be represented in homogeneous coordinates as

$$T(\tilde{\mathbf{x}}) = \begin{pmatrix} A & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \tilde{\mathbf{x}}$$

An application to computer graphics and image processing

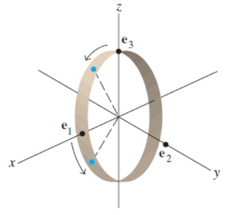
Example

An application in 3D graphics:
http://www.youtube.com/watch?v=EsNmiiKlRXQ

Example

Let's say we want to

1. Rotate a point 30°about the Y axis.



2. then, translate by $(-6, 4, 5)$

An application to computer graphics and image processing

Example (continued)

We need to use the transformation $T(\tilde{\mathbf{x}}) = \tilde{A}\tilde{\mathbf{x}}$ with

$$\tilde{A} = \begin{pmatrix} 1 & 0 & 0 & -6 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(30^\circ) & 0 & \sin(30^\circ) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(30^\circ) & 0 & \cos(30^\circ) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and

$$\tilde{x} = \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

An application to computer graphics and image processing

Example

Let's say we want to produce perspective projections. Let's imagine that the screen is on the XY plane and the viewer's eye is at $(0, 0, d)$ (the distance to the screen is d). Any object between the viewer and the screen is projected onto the screen as in the figure below

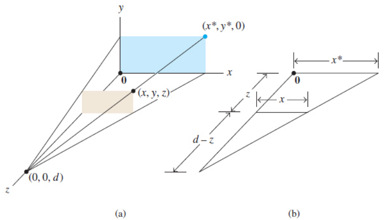


FIGURE 6 Perspective projection of (x, y, z) onto $(x^*, y^*, 0)$.

By similar triangles we have

$$\tan\alpha = \frac{x^*}{d} = \frac{x}{d-z} \Rightarrow x^* = \frac{x}{1-\frac{z}{d}}$$

An application to computer graphics and image processing

Example (continued)

Similarly, $y^* = \frac{y}{1-\frac{z}{d}}$. Using homogeneous coordinates we want that $(x, y, z, 1)$ maps onto $\left(\frac{x}{1-\frac{z}{d}}, \frac{y}{1-\frac{z}{d}}, 0, 1\right)$, or what is the same $\left(x, y, 0, 1 - \frac{z}{d}\right)$. We can achieve this with the perspective transformation:

$$\tilde{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{d} & 1 \end{pmatrix}$$

Exercises

Exercises

From Lay (3rd ed.), Chapter 2, Section 7:

- 2.7.2
- 2.7.3
- 2.7.10
- 2.7.12
- 2.7.22

Outline

3. Matrix algebra
   - Matrix operations (a)
   - Inverse of a matrix (b)
   - Elementary matrices (b)
   - An algorithm to invert matrices (b)
   - Characterization of invertible matrices (c)
   - Invertible linear transformations (c)
   - Partitioned matrices (c)
   - LU factorization (d)
   - An application to computer graphics and image processing (d)
   - Subspaces of $\mathbb{R}^n$ (e)
   - Dimension and rank (e)

Subspace

Definition 10.1 (Subspace of $\mathbb{R}^n$)

$H \subseteq \mathbb{R}^n$ is a subspace of $\mathbb{R}^n$ if:

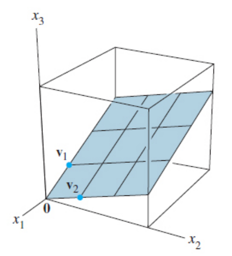
1. $\mathbf{0} \in H$
2. $\forall \mathbf{u}, \mathbf{v} \in H \quad \mathbf{u} + \mathbf{v} \in H$ *(H is closed under vector addition)*
3. $\forall \mathbf{u} \in H \, \forall r \in \mathbb{R} \quad r\mathbf{u} \in H$ *(H is closed under multiplication by a scalar)*

Example: Special subspaces

The following two sets are subspaces of $\mathbb{R}^n$:

1. $H = \{\mathbf{0}\}$
2. $H = \mathbb{R}^n$

Example: Plane



A plane is defined as

$H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\} = \{\mathbf{v} \in \mathbb{R}^n | \mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2\}$

This plane is a subspace of $\mathbb{R}^3$

Proof

1. Proof $\mathbf{0} \in H$

   If $\lambda_1 = \lambda_2 = 0$, then $\mathbf{v} = \mathbf{0}$.

2. Proof $\mathbf{u} + \mathbf{v} \in H$

   $\mathbf{u} \in H \Rightarrow \mathbf{u} = \lambda_{1u}\mathbf{v}_1 + \lambda_{2u}\mathbf{v}_2$

   $\mathbf{v} \in H \Rightarrow \mathbf{v} = \lambda_{1v}\mathbf{v}_1 + \lambda_{2v}\mathbf{v}_2$

   $$\begin{aligned} \mathbf{u} + \mathbf{v} &= (\lambda_{1u}\mathbf{v}_1 + \lambda_{2u}\mathbf{v}_2) + (\lambda_{1v}\mathbf{v}_1 + \lambda_{2v}\mathbf{v}_2) \\ &= (\lambda_{1u} + \lambda_{1v})\mathbf{v}_1 + (\lambda_{2u} + \lambda_{2v})\mathbf{v}_2 \in H \end{aligned}$$

3. Proof $r\mathbf{u} \in H$

   $\mathbf{u} \in H \Rightarrow \mathbf{u} = \lambda_{1u}\mathbf{v}_1 + \lambda_{2u}\mathbf{v}_2$

   $$\begin{aligned} r\mathbf{u} &= r(\lambda_{1u}\mathbf{v}_1 + \lambda_{2u}\mathbf{v}_2) \\ &= r\lambda_{1u}\mathbf{v}_1 + r\lambda_{2u}\mathbf{v}_2 \in H \end{aligned}$$

Subspace

Example: Line not through the origin

A line (L) that does not pass through the origin is not a subspace, because

1. $\mathbf{0} \notin L$
2. If we take two points belonging to the line ($\mathbf{u}$ and $\mathbf{v}$), $\mathbf{u}+\mathbf{v} \notin L$.
3. If we take a point belonging to the line ($\mathbf{w}$), $2\mathbf{w} \notin L$.

u + v is not on L 2w is not on L

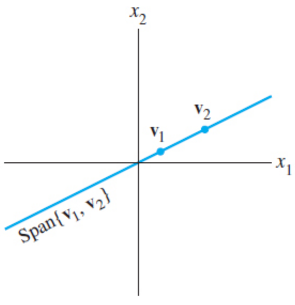
Subspace

Example: Line through the origin

Consider $\mathbf{v}_1$ and $\mathbf{v}_2 = k\mathbf{v}_1$. Then,

$$H = \mathrm{Span}\left\{\mathbf{v}_1, \mathbf{v}_2\right\} = \mathrm{Span}\left\{\mathbf{v}_1\right\}$$

is a line. It is easy to prove that this line is a subspace of $\mathbb{R}^n$.



$\mathbf{v}_1 \neq \mathbf{0}$, $\mathbf{v}_2 = k\,\mathbf{v}_1$.

Column space

Definition 10.2 (Column space of a matrix)

*Let $A \in \mathcal{M}_{m\times n}$. Let $\mathbf{a}_i \in \mathbb{R}^m$ be the columns of A. The **column space** of A is defined as*

$$\mathrm{Col}\{A\} = \mathrm{Span}\left\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\right\} \subseteq \mathbb{R}^m$$

Theorem 10.1

$\mathrm{Col}\{A\}$ *is a subspace of* $\mathbb{R}^m$.

Column space

Example

Let $A = \begin{pmatrix} 1 & -3 & -4 \\ -4 & 6 & -2 \\ -3 & 7 & 6 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 3 \\ 3 \\ -4 \end{pmatrix}$.

Determine if $\mathbf{b}$ belongs to $\mathrm{Col}\{A\}$.

Solution

If $\mathbf{b} \in \mathrm{Col}\{A\}$ there must be some coefficients x_1, x_2 and x_3 such that

$$\mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3$$

To find these coefficients we simply have to solve the equation system $A\mathbf{x} = \mathbf{b}$.

$$\left(\begin{array}{ccc|c} 1 & -3 & -4 & 3 \\ -4 & 6 & -2 & 3 \\ -3 & 7 & 6 & -4 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -3 & -4 & 3 \\ 0 & -6 & -18 & 15 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

In fact, there are infinite solutions to the equation system and, consequently, $\mathbf{b} \in \mathrm{Col}\{A\}$.

Null space

Definition 10.3 (Null space of a matrix)

*Let $A \in \mathcal{M}_{m\times n}$. The **null space** of A is defined as*

$$\mathrm{Nul}\{A\} = \{\mathbf{v} \in \mathbb{R}^n | A\mathbf{v} = \mathbf{0}\}$$

Theorem 10.2

$\mathrm{Nul}\{A\}$ *is a subspace of* $\mathbb{R}^n$.

Proof

1. *Proof* $\mathbf{0} \in \mathrm{Nul}\{A\}$
   $A\mathbf{0} = \mathbf{0} \Rightarrow \mathbf{0} \in \mathrm{Nul}\{A\}$ *(q.e.d.)*
2. *Proof* $\mathbf{u}+\mathbf{v} \in \mathrm{Nul}\{A\}$
   $\mathbf{u} \in \mathrm{Nul}\{A\} \Rightarrow A\mathbf{u} = \mathbf{0}$
   $\mathbf{v} \in \mathrm{Nul}\{A\} \Rightarrow A\mathbf{v} = \mathbf{0}$
   $A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0}+\mathbf{0} = \mathbf{0} \Rightarrow \mathbf{u}+\mathbf{v} \in \mathrm{Nul}\{A\}$ *(q.e.d.)*
3. *Proof* $r\mathbf{u} \in \mathrm{Nul}\{A\}$
   $\mathbf{u} \in \mathrm{Nul}\{A\} \Rightarrow A\mathbf{u} = \mathbf{0}$
   $A(r\mathbf{u}) = rA\mathbf{u} = r\mathbf{0} = \mathbf{0} \Rightarrow r\mathbf{u} \in \mathrm{Nul}\{A\}$ *(q.e.d.)*

Basis of a subspace

Definition 10.4 (Basis of a subspace)

Let $H \subseteq \mathbb{R}^n$. The set of vectors B is a basis of H if:

1. *All vectors in B are linearly independent*
2. $H = \mathrm{Span}\{B\}$

Standard basis of $\mathbb{R}^n$

Let be the vectors

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \ldots \\ 0 \end{pmatrix} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \ldots \\ 0 \end{pmatrix} \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \ldots \\ 0 \end{pmatrix} \quad \ldots \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \ldots \\ 1 \end{pmatrix}$$

The set $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is the **standard basis** of $\mathbb{R}^n$.

Basis of a subspace

Example

Find a basis for the null space of $A = \begin{pmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix}$.

Solution

The null space of A are all those vectors satisfying $A\mathbf{x} = \mathbf{0}$.

$$\left(\begin{array}{c|c} A & \mathbf{0} \end{array} \right) \sim \left(\begin{array}{ccccc|c} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

So the solution of the equation system is $\left.\begin{array}{c} x_1 = 2x_2 + x_4 - 3x_5 \\ x_3 = -2x_4 + 2x_5 \end{array}\right\} \Rightarrow$

$$\mathbf{x} = \begin{pmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{pmatrix} = x_2 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$

Null space and equation systems

Example (continued)

The set $B = \{(2,1,0,0,0),(1,0,-2,1,0),(-3,0,2,0,1)\}$ is a basis of $\mathrm{Nul}\{A\}$. By construction, we have chosen them to be linearly independent.

Example: Null space and equation systems

Consider $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

- $\{\mathbf{e}_3\}$ is a basis for $\mathrm{Nul}\{A\}$
- Consider $\mathbf{b} = (7,3,0)$. The general solution of $A\mathbf{x} = \mathbf{b}$ is of the form

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_{Nul}$$

where $\mathbf{x}_0$ is a solution of $A\mathbf{x} = \mathbf{b}$ that does not belong to $\mathrm{Nul}\{A\}$ and $\mathbf{x}_{Nul}$ belongs to $\mathrm{Nul}\{A\}$. In this particular case,

$$\mathbf{x} = (7,3,0) + x_3\mathbf{e}_3$$

Null space and equation systems

Example: Null space and equation systems (continued)

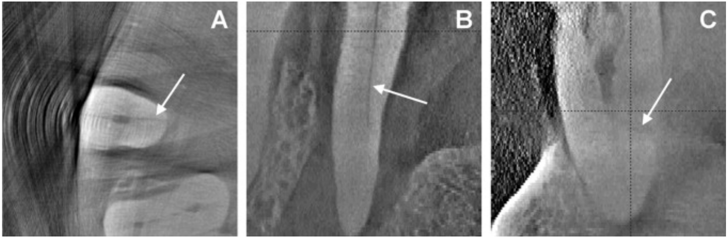
Let us prove that the general solution is actually a solution of $A\mathbf{x} = \mathbf{b}$

$$A\mathbf{x} = A(\mathbf{x}_0 + \mathbf{x}_{Nul}) = A\mathbf{x}_0 + A\mathbf{x}_{Nul} = \mathbf{b} + \mathbf{0} = \mathbf{b}$$

Intuititively we can say that the null space is the set of all solutions for which we have no measurements. The equation system only impose some constraints on those coefficients for which we have measurements. This is a problem in real situations as shown in the following slide.

Null space and equation systems

In this example, the authors describe how the exact location of a tooth fracture is uncertain (Fig. C) due to the artifacts introduced by the null space of the tomographic problem.



Mora, M. A.; Mol, A.; Tyndall, D. A., Rivera, E. M. *In vitro assessment of local computed tomography for the detection of longitudinal tooth fractures*.

Oral Surg Oral Med Oral Pathol Oral Radiol Endod, **2007**, 103, 825-829.

Basis of a subspace

Example

Find a basis for the column space of $B = \begin{pmatrix} 1 & 0 & -3 & 5 & 0 \\ 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

Solution

From the columns with non-pivot positions of matrix B we learn that

$$\begin{array}{c} \mathbf{b}_3 = -3\mathbf{b}_1 + 2\mathbf{b}_2 \\ \mathbf{b}_4 = 5\mathbf{b}_1 - \mathbf{b}_2 \end{array}$$

Then,

$$\begin{array}{rcl} \mathrm{Col}\{B\} & = & \left\{\mathbf{v} \in \mathbb{R}^4 | \mathbf{v} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + x_3\mathbf{b}_3 + x_4\mathbf{b}_4 + x_5\mathbf{b}_5\right\} \\ & = & \left\{\mathbf{v} \in \mathbb{R}^4 \left| \begin{array}{rcl} \mathbf{v} & = & x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + x_3(-3\mathbf{b}_1 + 2\mathbf{b}_2) + \\ & & x_4(5\mathbf{b}_1 - \mathbf{b}_2) + x_5\mathbf{b}_5 \end{array} \right. \right\} \\ & = & \left\{\mathbf{v} \in \mathbb{R}^4 | \mathbf{v} = x_1'\mathbf{b}_1 + x_2'\mathbf{b}_2 + x_5\mathbf{b}_5\right\} \end{array}$$

And, consequently, $\mathrm{Basis}\{\mathrm{Col}\{B\}\} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_5\}$

Example

Find a basis for the column space of $A = \begin{pmatrix} 1 & 3 & 3 & 2 & -9 \\ -2 & -2 & 2 & -8 & 2 \\ 2 & 3 & 0 & 7 & 1 \\ 3 & 4 & -1 & 11 & -8 \end{pmatrix}$.

Solution

It turns out that $A \sim B$ (B in the previous example). Since row operations do not affect linear dependence relations among the columns of the matrix, we should have

$$\begin{aligned} \mathbf{a}_3 &= -3\mathbf{a}_1 + 2\mathbf{a}_2 \\ \mathbf{a}_4 &= 5\mathbf{a}_1 - \mathbf{a}_2 \end{aligned}$$

and $\mathrm{Basis}\{\mathrm{Col}\{A\}\} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_5\}$

Theorem 10.3

The pivot columns of A form a basis of $\mathrm{Col}\{A\}$.

Exercises

Exercises

From Lay (3rd ed.), Chapter 2, Section 1:

- 2.8.1
- 2.8.2
- 2.8.5

Outline

3. Matrix algebra
   - Matrix operations (a)
   - Inverse of a matrix (b)
   - Elementary matrices (b)
   - An algorithm to invert matrices (b)
   - Characterization of invertible matrices (c)
   - Invertible linear transformations (c)
   - Partitioned matrices (c)
   - LU factorization (d)
   - An application to computer graphics and image processing (d)
   - Subspaces of $\mathbb{R}^n$ (e)
   - Dimension and rank (e)

Coordinate system

Definition 11.1 (Coordinates of a vector in the basis B)

Suppose $B = \{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_p\}$ *is a basis for the subspace* $H \subseteq \mathbb{R}^n$*. For each* $\mathbf{x} \in H$*, the coordinates of* $\mathbf{x}$ *relative to the basis* B *are the weights* c_i *such that*

$$\mathbf{x} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + ... + c_p\mathbf{b}_p$$

The ***coordinates of*** $\mathbf{x}$ ***with respect to the basis*** B *is the vector in* $\mathbb{R}^p$

$$[\mathbf{x}]_B = \begin{pmatrix} c_1 \\ c_2 \\ ... \\ c_p \end{pmatrix}$$

Coordinate system

Example

Let $\mathbf{x} = (3, 12, 7)$, $\mathbf{v}_1 = (3, 6, 2)$, $\mathbf{v}_2 = (-1, 0, 1)$, $B = \{\mathbf{v}_1, \mathbf{v}_2\}$.

1. Show that B is a linearly independent set
2. Find the coordinates of $\mathbf{x}$ in the coordinate system B

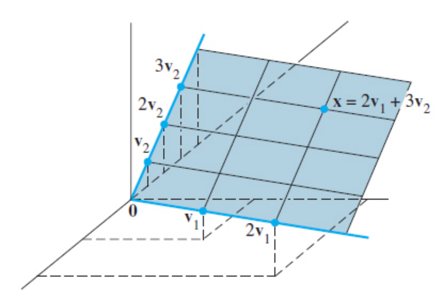
Solution

1. We need to prove that the only solution of the equation system $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}$ is $c_1 = c_2 = 0$.
$$\left(\begin{array}{cc|c} 3 & -1 & 0 \\ 6 & 0 & 0 \\ 2 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right)$$
And, therefore, the unique solution is $c_1 = c_2 = 0$ (q.e.d.)
2. We need to find c_1 and c_2 such that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{x}$
$$\left(\begin{array}{cc|c} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right)$$
And, therefore, $[\mathbf{x}]_B = (2, 3)$.

Coordinate system

Example (continued)

The following figure shows how $\mathbf{x}$ is equal to $2\mathbf{v}_1 + 3\mathbf{v}_2$



Coordinate system

Theorem 11.1

The coordinates of a given vector with respect to a given basis are unique.
Proof
Let us assume they are not unique. Then, there must be two different sets of coordinates such that

$$\mathbf{x} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + ... + c_p\mathbf{b}_p$$
$$\mathbf{x} = c_1'\mathbf{b}_1 + c_2'\mathbf{b}_2 + ... + c_p'\mathbf{b}_p$$

If we subtract both equations, we have

$$\mathbf{0} = (c_1 - c_1')\mathbf{b}_1 + (c_2 - c_2')\mathbf{b}_2 + ... + (c_p - c_p')\mathbf{b}_p$$

But because the basis is a linearly independent set of vectors, it must be

$$\begin{array}{c} c_1 - c_1' = 0 \Rightarrow c_1 = c_1' \\ c_2 - c_2' = 0 \Rightarrow c_2 = c_2' \\ ... \\ c_p - c_p' = 0 \Rightarrow c_p = c_p' \end{array}$$

This is a contradiction with the hypothesis that there were two different sets of coordinates, and therefore, the coordinates of the vector $\mathbf{x}$ *must be unique.*

Subspace dimension

Isomorphism to $\mathbb{R}^p$

For any given subspace H and its corresponding basis B, the mapping

$$\begin{array}{rcl} T : H & \rightarrow & \mathbb{R}^p \\ \mathbf{x} & \rightarrow & [\mathbf{x}]_B \end{array}$$

is a linear, injective transformation that makes H to behave as $\mathbb{R}^p$.

Definition 11.2 (Dimension)

The ***dimension of a subspace*** *H ($\dim\{H\}$) is the number of vectors of any of its basis.*
The dimension of $H = \{\mathbf{0}\}$ is 0.

Example (continued)

In our previous example in which $B = \{\mathbf{v}_1, \mathbf{v}_2\}$, the dimension is 2, in fact H behaves like a plane (see previous figure in the example).

Rank of a matrix

Definition 11.3 (Rank of a matrix)

The ***rank of a matrix*** *A is* $\mathrm{rank}\{A\} = \dim\{\mathrm{Col}\{\mathrm{A}\}\}$, *that is, the dimension of the column space of the matrix.*
MATLAB: `rank(A)`

Theorem 11.2

The rank of a matrix is the number of pivot columns it has.
Proof
Since the pivot columns form a basis of the column space of A, the number of pivot columns is the rank of the matrix.

Example

$$A = \begin{pmatrix} 1 & 3 & 3 & 2 & -9 \\ -2 & -2 & 2 & -8 & 2 \\ 2 & 3 & 0 & 7 & 1 \\ 3 & 4 & -1 & 11 & -8 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -3 & 5 & 0 \\ 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Therefore, the rank of A is 3.

Rank of a matrix

Theorem 11.3 (Rank theorem)

If A has n columns, then

$$\mathrm{Rank}\{A\} + \dim\{\mathrm{Nul}\{\mathrm{A}\}\} = n$$

Theorem 11.4 (Basis theorem)

Let H be a subspace of dimension p. Any linearly independent set of p vectors of H is a basis of H. Any set of p vectors that span H is a basis of H.

Characterization of invertible matrices (continued)

Theorem 11.5 (The invertible matrix theorem)

Let $A \in \mathcal{M}_{n\times n}$. The following statements are equivalent (either they are all true or they are all false):

xiii. *The columns of A form a basis of* $\mathbb{R}^n$

xiv. $\mathrm{Col}\{\mathrm{A}\} = \mathbb{R}^n$

xv. $\dim\{\mathrm{Col}\{\mathrm{A}\}\} = n$

xvi. $\mathrm{Rank}\{A\} = n$

xvii. $\mathrm{Nul}\{\mathrm{A}\} = \{\mathbf{0}\}$

xviii. $\dim\{\mathrm{Nul}\{\mathrm{A}\}\} = 0$

Characterization of invertible matrices

Proof v $\Rightarrow$ xiii

This is true by the basis theorem.

Proof xiii $\Rightarrow$ xiv

By the definition of basis.

Proof xiii $\Rightarrow$ xv

By the definition of dimension.

Proof xv $\Rightarrow$ xvi

By the definition of rank.

Proof xvi $\Rightarrow$ xviii

By the rank theorem.

Proof xvii $\Rightarrow$ iv

By the definition of null space.

Exercises

Exercises

From Lay (3rd ed.), Chapter 2, Section 9:

- 2.9.1
- 2.9.3
- 2.9.9
- 2.9.19
- 2.9.27

Outline

3. Matrix algebra
   - Matrix operations (a)
   - Inverse of a matrix (b)
   - Elementary matrices (b)
   - An algorithm to invert matrices (b)
   - Characterization of invertible matrices (c)
   - Invertible linear transformations (c)
   - Partitioned matrices (c)
   - LU factorization (d)
   - An application to computer graphics and image processing (d)
   - Subspaces of $\mathbb{R}^n$ (e)
   - Dimension and rank (e)