ALGEBRA

1. MATRICES, DETERMINANTS AND LINEAR SYSTEMS

Grau de Matemàtiques. FME
2018-2019

Marta Casanellas
Departament de Matemàtiques

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

OUTLINE

Outline

Definition

An $m \times n$ matrix is a collection A of $m \times n$ (real or complex) numbers arranged into a rectangular array of m rows and n columns. We denote by $a_{i,j}$ the element at row i and column j of A.

Notation: $\quad A = \left(a_{i,j}\right)_{i,j}$.

- If $m = n$ we say that A is a square matrix of size n.
- The set of $m \times n$ matrices with real (resp. complex) entries is denoted as $\mathcal{M}_{m\times n}(\mathbb{R})$ (resp. $\mathcal{M}_{m\times n}(\mathbb{R})$); we use the notation $\mathcal{M}_{m\times n}$ if we do not need to specify $\mathbb{R}$ or $\mathbb{C}$.
- The elements of $\mathcal{M}_{n\times 1}$ are called vectors or column vectors.
- The elements of $\mathcal{M}_{1\times n}$ are called row vectors.

SPECIAL MATRICES

- The matrix **0** is the matrix whose elements are all 0.
- A square matrix A is a diagonal matrix if $a_{i,j} = 0$ for all $i \neq j$.
- The identity matrix Id_n is the diagonal $n \times n$ matrix that has 1's at the diagonal entries.
- A square matrix A is a lower triangular matrix if $a_{i,j} = 0$ for all $i < j$.
- A square matrix A is an upper triangular matrix if $a_{i,j} = 0$ for all $i > j$.

Transpose

Definition

The transpose of $A \in \mathcal{M}_{m\times n}$ is the $n \times m$ matrix A^t whose (i,j)-entry is $a_{j,i}$:

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} \to A^t = \begin{pmatrix} a_{1,1} & \dots & a_{m,1} \\ \vdots & \vdots & \vdots \\ a_{1,n} & \dots & a_{m,n} \end{pmatrix}$$

Definition

- A square matrix is symmetric if $A^t = A$
- A square matrix is skew-symmetric if $A^t = -A$ (in particular, its diagonal entries are 0).

Operations: sum of matrices

If A, B are two $m \times n$ matrices, then the sum $C := A + B$ is the matrix whose (i,j)-entry is $c_{i,j} = a_{i,j} + b_{i,j}$:

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} + \begin{pmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & \vdots & \vdots \\ b_{m,1} & \dots & b_{m,n} \end{pmatrix} = \begin{pmatrix} a_{1,1}+b_{1,1} & \dots & a_{1,n}+b_{1,n} \\ \vdots & \vdots & \vdots \\ a_{m,1}+b_{m,1} & \dots & a_{m,n}+b_{m,n} \end{pmatrix}$$

Properties: associative, commutative, neutral element $\mathbf{0}$, opposite element $-A = (-a_{i,j})_{i,j}$,

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} - \begin{pmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & \vdots & \vdots \\ b_{m,1} & \dots & b_{m,n} \end{pmatrix} = \begin{pmatrix} a_{1,1}-b_{1,1} & \dots & a_{1,n}-b_{1,n} \\ \vdots & \vdots & \vdots \\ a_{m,1}-b_{m,1} & \dots & a_{m,n}-b_{m,n} \end{pmatrix}$$

$$(A+B)^t = A^t + B^t$$

Product by a scalar

Let $A \in \mathcal{M}_{m\times n}$ and let $x \in \mathbb{R}$ be a number (scalar), then $x \cdot A$ is the $m \times n$ matrix whose (i,j)-element is $xa_{i,j}$ for all $i \in \{1,\dots,m\}$, $j \in \{1,\dots,n\}$:

$$x \cdot \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} = \begin{pmatrix} xa_{1,1} & \dots & xa_{1,n} \\ \vdots & \vdots & \vdots \\ xa_{m,1} & \dots & xa_{m,n} \end{pmatrix}$$

Properties: $0 \cdot A = \mathbf{0}$, $x \cdot (A+B) = x \cdot A + x \cdot B$, $(x \cdot A)^t = x \cdot A^t$, $x \cdot Id_n = diag(x,\dots,x)$

Multiplication of matrices

Let $A \in \mathcal{M}_{m\times n}$ and $B \in \mathcal{M}_{n\times p}$, then AB is the matrix C such that

$$c_{i,j} = a_{i,1}b_{1,j}+a_{i,2}b_{2,j}+\cdots+a_{i,n}b_{n,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j} = (a_{i,1}\ a_{i,2} \ldots a_{i,n}) \begin{pmatrix} b_{1,j} \\ \vdots \\ b_{n,j} \end{pmatrix}$$

Example:

a) $\begin{bmatrix} 1 & 2 & 4 \\ 3 & 5 & 1 \\ 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 2 & 3 & 2 \\ 4 & 6 & 1 \\ 3 & 2 & 5 \end{bmatrix} = \begin{bmatrix} 22 & 23 & 24 \\ 29 & 41 & 16 \\ 17 & 18 & 20 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 2 & 4 \\ 3 & 5 & 1 \\ 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 2 & 3 & 2 \\ 4 & 6 & 1 \\ 3 & 2 & 5 \end{bmatrix} = \begin{bmatrix} 22 & 23 & 24 \\ 29 & 41 & 16 \\ 17 & 18 & 20 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 2 & 4 \\ 3 & 5 & 1 \\ 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 2 & 3 & 2 \\ 4 & 6 & 1 \\ 3 & 2 & 5 \end{bmatrix} = \begin{bmatrix} 22 & 23 & 24 \\ 29 & 41 & 16 \\ 17 & 18 & 20 \end{bmatrix}$

Remark: Each column (resp. row) of C is a *linear combination* of the columns (rep. rows) of A (resp. B):

$$\begin{pmatrix} c_{1,j} \\ \vdots \\ c_{m,j} \end{pmatrix} = b_{1,j} \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \end{pmatrix} + b_{2,j} \begin{pmatrix} a_{1,2} \\ \vdots \\ a_{m,2} \end{pmatrix} + b_{n,j} \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{m,n} \end{pmatrix}.$$

- $Id_m A = A Id_n = A$ (neutral element)
- $A(BC) = (AB)C$ (associative)
- $A(B + C) = AB + AC$ (distributive law)
- $(A + B)C = AC + BC$ (distributive law)
- $AB \neq BA$
- $(AB)^t = B^t A^t$
- **it does not exist an inverse element in general**: given a matrix A, under which conditions does there exist a matrix B such that

$$AB = BA = Id_n?$$

First, we need that A and B can be multiplied as AB and BA, so we require A, B to be **square** matrices of the same size.

Inverse

Definition

Let A be an $n \times n$ matrix. If there exists a matrix B such that

$$AB = BA = Id_n$$

then B is called the inverse of A and is denoted as A^{-1}. A matrix is called invertible (or non-singular) if it has an inverse and is called singular if it does NOT have an inverse.

Remark: We'll prove later that only $AB = Id_n$ or $BA = Id_n$ is necessary (the other comes for free).

Properties of the inverse

If A and B are $n \times n$ invertible matrices, then

1. The inverse is unique
2. $(A^{-1})^{-1} = A$
3. $(A^t)^{-1} = (A^{-1})^t$
4. $(AB)^{-1} = B^{-1}A^{-1}$
5. $(A^k)^{-1} = (A^{-1})^k$ for $k \in \mathbb{N}$ (and if $m \in \mathbb{N}$, A^{-m} is defined as $(A^{-1})^m$).

Inverse in the 2×2 case

- If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $ad - bc \neq 0$, then the following matrix is the inverse of A:
$$\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$
- Computing the inverse for larger matrices: see the section “Determinant” and “Simultaneous Linear Systems”.

Outline

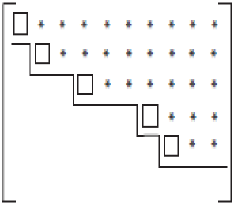
Elementary operations

Given an $m \times n$ matrix A, the following are called elementary operations:

- E_1 Exchange two rows, $r_i \leftrightarrow r_j$
- E_2 Multiply a row by a nonzero constant, $r_i \leftarrow cr_i$.
- E_3 Add a multiple of one row to another row, $r_i \leftarrow r_i + cr_j$.

Row echelon form

Gaussian elimination is an algorithm that uses elementary operations to transform a matrix into a matrix in row echelon form:



- squares are called pivots = first non-zero element of each row.
- $*$: can be 0 or not.

Definition

A matrix is in row echelon form if:

- any rows consisting entirely of zeros are at the bottom.
- in each row, the pivot is always further to the right than the pivot of the previous row.

Given A an $m \times n$ matrix, the following elementary operations transform A into a row echelon form:

(1) Start with the leftmost non-zero column.

(2) Permute rows if necessary (type E_1), so that the top position is non-zero (hence a pivot).

(3) Use operations of type E_3 so obtain zeros below the pivot in that column.

(4) Repeat steps (1),(2),(3) with the submatrix below the pivot.

Computational issues: dividing by small numbers is never a good idea; there are better implementations of this algorithm.

Definition

An elementary matrix is any matrix that can be obtained by performing an elementary row operation on an identity matrix.

There are elementary matrices of 3 types.
Examples: on blackboard

Theorem

Let E be the elementary matrix obtained by performing an elementary row operation on Id_n. If the same elementary operation is performed on an $n \times m$ matrix A, the result is the same as the matrix EA.

Elementary matrices are not good for practical purposes but are important from the theoretical point of view to prove certain results.
Properties:

LEMMA

Every elementary matrix is invertible and its inverse is an elementary matrix of the same type.

Proof: elementary row operations can be reversed.

The rank of an $m \times n$ matrix is the number of pivots in a row echelon form of A.

Properties (to be proved in the "Linear Systems" section):

- The rank of A is the same no matters which process we choose to reduce the matrix (so the definition of rank does not depend on the echelon form we obtain).
- $rank(A) = rank(A^t)$.

Remark. Elementary transformations can also be performed on the *columns* of the matrix.

Outline

Definition

A system of m linear equations in the n variables $x_1, \ldots, x_n$ is a collection of equations

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

where the *coefficients* a_{ij} and the *independent terms* $b_1, b_2, \ldots, b_m$ are constants (usually in $\mathbb{R}$). A system is homogeneous if $b_i = 0$ for $i = 1, \ldots, m$.

A solution is an n-tuple $(s_1, \ldots, s_n)$ which satisfies the equations when we substitute $x_1 = s_1, \ldots, x_n = s_n$.
The solution set is the set of all solutions of the system.

In matrix notation, we write the system as $Ax = b$, where

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}, \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}$$

The matrix A is the matrix of the system.
The augmented matrix is $(A \mid b)$.

THEOREM

Any system of linear equations has either (i) no solution, (ii) a unique solution, or (iii) an infinite number of solutions.

A system of linear equations is consistent if it has one or more solutions. Otherwise, it is inconsistent.
If it is consistent, then

- it is determinate if there is a unique solution
- it is indeterminate if there are infinite solutions.

- A homogeneous system is always consistent (0 is always a solution).
- If y is a solution of $Ax = 0 \Rightarrow cy$ is also a solution for any $c \in \mathbb{R}$.
- If y and z are solutions of $Ax = 0 \Rightarrow y + z$ is also a solution.

Proposition

For a consistent system, $Ax = b$, the solution set is formed by the sum of any particular solution plus the solutions of the corresponding homogeneous system $Ax = 0$. That is, if x_0 is a solution of $Ax = b$, then all the solutions x of $Ax = b$ can be written as

$$x = x_0 + y$$

where y is a solution of the homogeneous system $Ax = 0$.

Geometric interpretation

From a geometric point of view, the solution set of a linear system describes a linear variety.

DEFINITION

Two systems of linear equations are called equivalent if they have the same set of solutions.

PROPOSITION

Elementary operations on the augmented matrix transform the system into an equivalent system.

Goal: Use elementary operations to transform a system to a simpler system (for example a system in row echelon form).

To solve a linear system $Ax = b$ we consider the augmented matrix $(A \mid b)$ and apply the following 2-step algorithm:

1ST STEP Reduce $(A \mid b)$ to row echelon form.

2ND STEP Solve the system by back substitution if it is consistent.

- The number of pivots (rank) of the row echelon form of A and $(A|b)$ tells us whether the system is consistent or not.
- If the system is consistent, then the leading variables corresponding to pivots can be written in terms of the other variables (called free variables).
- The number of free variables is the degrees of freedom of the system.

Back substitution and Gauss-Jordan elimination

The back substitution step can also be performed by elementary row operations on the row echelon form of $(A|b)$ by an algorithm known as Gauss-Jordan elimination:

Once we have a matrix in *row echelon form*, do:

1. start with the rightmost pivot and use an operation of type E_2 to convert it to 1.
2. from bottom to top: make all the entries above the pivot equal to zero using type E_3.
3. Repeat the previous steps the next column to the left (so, from right to left).

Reduced row echelon form

In this way we obtain a matrix in reduced row echelon form, that is a matrix of the following form:

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

Definition

A matrix is in reduced row echelon form if it is in row echelon form and

- all pivots are 1
- the pivots are the only non-zero entries in its column.

Reduced row echelon form

- If A is a square matrix and and the reduced row echelon form is the identity Id_n, then the system $Ax = b$ can be trivially solved: the solution is the new independent term

$$(A \,|\, b) \sim \cdots \sim (Id_n \,|\, b') \quad \text{so} \quad Ax = b \Leftrightarrow Id_n x = b' \Leftrightarrow x = b'$$

- Whereas the row echelon form of A is not unique, the *reduced* row echelon form is unique:

Theorem

Each matrix is equivalent, via elementary transformations, to a unique reduced row echelon form.

Corollary

The rank of a matrix is well defined (is independent of the elementary transformations we perform).

Rouché-Frobenius Theorem

Theorem (Rouché-Frobenius)

Consider $A \in \mathcal{M}_{m\times n}$ and $b \in \mathcal{M}_{n\times 1}$. Then,

- *The system $Ax = b$ is consistent if and only if $rank(A) = rank(A|b)$.*
  In this case, the set of solutions depends on $n - rank(A)$ free variables;
  $n - rank(A) =$ degrees of freedom of the system.
- *In particular, if $n = rank(A)$, the solution is unique.*

Outline

Let A be an $n \times n$ matrix, we define the determinant of A, det(A), as follows (notation $|A| = det(A)$):

- If $n = 1$, $A = (a_{1,1})$, then $\det(A) = a_{1,1}$.
- If $n = 2$, $\det(A) = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$.
- If $n = 3$,

$$\det(A) = a_{11}\begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} - a_{1,2}\begin{vmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} + a_{1,3}\begin{vmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{vmatrix}$$

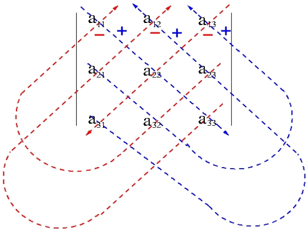
- Recursively, if $A_{i,j}$ is the matrix obtained by removing row i and column j from A,

$$|A| = a_{11}\det A_{1,1} - a_{1,2}\det A_{1,2} + \cdots + (-1)^{n+1}a_{1,n}\det A_{1,n}.$$

This is called the Laplace *expansion of the determinant by the first row*.

Sarrus Rule:

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = a_{1,1}a_{2,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3} - a_{1,3}a_{2,2}a_{3,1} - a_{2,3}a_{3,2}a_{1,1} - a_{3,3}a_{1,2}a_{2,1}$$



Warning: Not valid for $n \geq 4$.

We can compute det it by expanding by any row or column:
Given a square matrix A, we define the cofactor matrix of A as the matrix $co(A)$ whose (i,j) entry is $C_{i,j} = (-1)^{i+j} \det A_{i,j}$, where $A_{i,j}$ is the matrix obtained by removing row i and column j of A.

THEOREM (LAPLACE EXPANSION)

The determinant of an $n \times n$ matrix A can be computed as the cofactor expansion along the ith row,

$$\det A = a_{i,1} C_{i,1} + \ldots + a_{i,n} C_{i,n}$$

and also as the cofactor expansion along the jth column:

$$\det A = a_{1,j} C_{1,j} + \ldots + a_{n,j} C_{n,j}.$$

Let A be a square matrix. Then,

Lemma (Effect of E_1)

If we exchange two rows of A and call B the new matrix, then:

$$\det(B) = -\det(A).$$

The following properties characterize det:

(1) If two rows (or columns) of A are equal, then $\det(A) = 0$.

(2) If we multiply a row (or column) by $x \neq 0$ and call B the new matrix we have $\det(B) = x\det(A)$.

(3) If A,B, C are identical except that the ith row (or column) of C is the sum of the ith rows (columns) of A and B, then $\det(C) = \det(A) + \det(B)$.

(4) $\det(Id_n) = 1$.

If A is an $n \times n$ matrix, then to compute the determinant, the number of multiplications involved is:

- $(n-1)n!$
  if we compute the determinant using Laplace expansion.
- $\frac{(n-1)n(2n-1)}{6} + \frac{(n-1)n}{2} + n - 1 = \frac{n^3}{3} + \text{lower order terms}$
  if we use Gaussian elimination first.

For example, if $n = 8$ the first number is 282240 whereas the second is 175.
Goal: Transformations of type E_3 do not change the value of $det(A) \Rightarrow$ do E_3 transformations (and of type E_1 if necessary) to compute efficiently $det(A)$.

Other properties of the determinant:

- If one row or column is 0, then $\det(A) = 0$.
- If A is a triangular matrix, $\det(A)$ is the product of elements in the diagonal.
- $\det(A^t) = \det(A)$
- $\det(c \cdot A) = c^n \det(A)$.
- $\det(AB) = \det(A)\det(B)$ (proof: slide 43)

Consequence:

If A is invertible (non-singular) $\Rightarrow \det(A) \neq 0$ and $\det(A^{-1}) = 1/\det(A)$.

Proposition

An $n \times n$ matrix A is invertible if and only if $\det(A) \neq 0$.

Note that the adjugate or adjoint matrix $co(A)^t$ satisfies

$$Aco(A)^t = co(A)^t A = \det(A) Id_n.$$

Therefore, if $\det(A) \neq 0$, the matrix

$$\frac{1}{\det(A)} co(A)^t$$

is the inverse of A.

Note: this is not the optimal way to compute the inverse for $n \geq 4$; for a more efficient way: see ”Simultaneous linear Systems“.

Proposition

An $n \times n$ matrix A has rank n if and only if $\det(A) \neq 0$.

A minor of A is a the determinant of a square submatrix of A obtained by using only some rows and columns of A.

Proposition

The rank of A equals the maximum size of non-zero minors of A.

Corollary

$$rank(A) = rank(A^t)$$

Theorem

Let $A \in \mathcal{M}_{n\times n}$ be a square matrix. Then the following are equivalent:

1. *A is invertible.*
2. *$Ax = b$ has a unique solution for every b.*
3. *$Ax = 0$ has only the trivial solution 0.*
4. *The reduced row echelon form of A is Id_n.*
5. *$rank(A) = n$.*
6. $\det(A) \neq 0$.

WE CAN PROVE NOW:

COROLLARY

If A is a square matrix and B is a square matrix such that $AB = Id_n$ or $BA = Id_n$ then, B is the inverse of A.

COROLLARY

$$det(AB) = det(A)det(B)$$

Outline

Goal: solve systems with the same $m \times n$ matrix A but different independent terms,

$$Ax^{(1)} = b^{(1)}, Ax^{(2)} = b^{(2)}, \ldots, Ax^{(r)} = b^{(r)}.$$

Equivalently: find X $m \times r$ matrix such that

$$AX = \underbrace{\begin{pmatrix} b^{(1)} \, b^{(2)} \, \ldots \, b^{(r)} \end{pmatrix}}_{B}.$$

matrix equation $AX = B$

Efficient solution: Gauss-Jordan elimination to the following augmented matrix

$$\left(A \,|\, b^{(1)} \, b^{(2)} \, \ldots \, b^{(r)}\right)$$

To find the inverse of an $n \times n$ matrix, we can think of solving n systems simultaneously: $Ax^{(i)} = e_i$, $i = 1, \ldots, n$ (e_i is a column vector with 1 at the ith entry and 0 elsewhere). Then use Gauss-Jordan:
Input: an $n \times n$ matrix A.
Output: the inverse of A if A is nonsingular.

1. Form the $n \times 2n$ matrix $M = (A \mid Id_n)$, where A is the left half and the identity matrix Id_n is the right half.
2. Reduce M to row echelon form. If the process generates a zero row in the left block of M, then A has no inverse $\Rightarrow$ STOP.
3. Further reduce M to its *reduced* row echelon form $M \sim (Id_n \mid B)$, where the identity matrix Id_n has replaced A in the left half (*back substitution*).
4. Then $A^{-1} =$ the matrix B that is now in the right half.

Outline

Basic:

- D. Poole, Algebra lineal, una introducción moderna (3a ed.), Brooks/Cole, 2013.
- Lay et al., Algebra lineal y sus aplicaciones (4a ed.), Pearson Educación 2012.

Additional

- Castellet, LLerena, Algebra lineal i geometria, Edicions UAB, 2000.