TMV035 Analysis and Linear Algebra B, 2005

LECTURE 5.1

In this lecture we present the vector space \mathbb{R}^n , Gauss elimination method. This covers AMBS Ch 42.3–6, 42.13.

1. The vector space \mathbf{R}^n

Linear combination. \mathbf{R}^n is the set of all column vectors of n elements, where each element is a real number:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad x_1, x_2, \dots, x_n \in \mathbf{R}.$$

The algebraic operations addition and multiplication by a scalar are defined elementwise:

$$x+y = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{bmatrix}, \quad \alpha x = \alpha \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{bmatrix}, \quad \alpha \in \mathbf{R}.$$

By combining these operations we can form *linear combinations* of column vectors:

$$\alpha x + \beta y = \begin{bmatrix} \alpha x_1 + \beta y_1 \\ \vdots \\ \alpha x_n + \beta y_n \end{bmatrix}, \quad \alpha, \beta \in \mathbf{R}.$$

This is an important way of creating new column vectors. Of course, we can form linear combinations of more than two column vectors:

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_m x_m.$$

The numbers α_j are called *coefficients*.

The zero column vector is

$$0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

It satisfies

$$x + 0 = x.$$

It is easy to check that all the familiar rules of algebra apply: the commutative and associative laws of addition, and the distributive law.

Notice that if x, y belong to \mathbb{R}^n then any linear combination of them, $\alpha x + \beta y$, also belongs to \mathbb{R}^n . A set with this property, that it preserves linear combinations, or linear combinations remain inside the set, is called a *vector space* or *linear space* ("vektorrum, linjärt rum"). Elements of such a set are called *vectors*. So \mathbb{R}^n is a linear space and column vectors are a special kind of vectors.

The set of all row vectors of n real numbers:

$$x = [x_1, \ldots, x_n]$$

is also denoted \mathbb{R}^n , and it is also a linear space with addition and multiplication by a scalar defined elementwise. We usually work with column vectors.

We will meet other examples of vector spaces and vectors later, i.e., vectors which are not column or row vectors.

Date: November 20, 2005, Stig Larsson, Mathematical Sciences, Chalmers University of Technology.

LECTURE 5.1

Let $a_1, \ldots, a_j \in \mathbf{R}^n$ be j vectors. The set of all linear combinations of the given vectors:

$$S(a_1,\ldots,a_j) = \{x = \alpha_1 a_1 + \cdots + \alpha_j a_j : \alpha_1,\ldots,\alpha_j \in \mathbf{R}\}$$

is called the *space spanned by*, or the *space generated by*, or simply the *span* of the given vectors. ("Rummet som genereras av de givna vektorerna.") This is also a vector space.

Theorem. The set $S(a_1, \ldots, a_j)$ is a linear space.

Proof. We must show that all linear combinations of elements of $S(a_1, \ldots, a_j)$ remain inside $S(a_1, \ldots, a_j)$. Let

$$x = \alpha_1 a_1 + \dots + \alpha_j a_j, \quad y = \beta_1 a_1 + \dots + \beta_j a_j$$

be two elements of $S(a_1, \ldots, a_j)$. Form a linear combination of them:

$$\alpha x + \beta y = \alpha (\alpha_1 a_1 + \dots + \alpha_j a_j) + \beta (\beta_1 a_1 + \dots + \beta_j a_j)$$

= $(\alpha \alpha_1 a_1 + \dots + \alpha \alpha_j a_j) + (\beta_1 \beta a_1 + \dots + \beta_j \beta a_j)$
= $(\alpha \alpha_1 + \beta \beta_1) a_1 + \dots + (\alpha \alpha_j + \beta \beta_j) a_j.$

We see that $\alpha x + \beta y$ is also a linear combination of a_1, \ldots, a_j , in other words $\alpha x + \beta y$ belongs to $S(a_1, \ldots, a_j)$.

Note in particular that the zero vector belongs to $S(a_1, \ldots, a_j)$. Proof: Take all coefficients $\alpha_1 = \cdots = \alpha_j = 0$. Then $0 = 0a_1 + \cdots + 0a_j \in S(a_1, \ldots, a_j)$. Important: here the zero on the left-hand side is the zero vector, while the zeros on the right-hand side are the number zero:

$$\begin{bmatrix} 0\\ \vdots\\ 0 \end{bmatrix} = 0 \begin{bmatrix} a_{11}\\ \vdots\\ a_{n1} \end{bmatrix} + \dots + 0 \begin{bmatrix} a_{1j}\\ \vdots\\ a_{nj} \end{bmatrix}.$$

Since the vector space $S(a_1, \ldots, a_j)$ is a subset of the vector space \mathbf{R}^n , $S(a_1, \ldots, a_j) \subset \mathbf{R}^n$, we say that $S(a_1, \ldots, a_j)$ is a subspace of \mathbf{R}^n ("underrum"). We will often work with this kind of supspace.

Note: in mathematics the word "space" means "a set with some extra structure". So linear space (vector space) is a set where you can form linear combinations.

1.1. Scalar product. We define the *scalar product* or *innner product* of two column vectors ("skalärprodukt"):

$$(x,y) = \sum_{j=1}^{n} x_j y_j$$

The scalar product is also denoted by $x \cdot y$. Note that it can also be computed as:

$$(x,y) = y^T x = \begin{bmatrix} y_1, \dots, y_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

In MATLAB:

>> y'*x

We also define the *length* or *norm* of a vector:

$$||x|| = \sqrt{(x,x)} = \sqrt{\sum_{j=1}^{n} x_j^2}$$

In MATLAB: norm(x). Note that this is not the same as abs(x), which is max $|x_j|$.

The scalar product has the following properties.

Theorem. If $x, y, z \in \mathbf{R}^n$ and $\alpha \in \mathbf{R}$ then

(1)
$$(x,x) \begin{cases} = 0, & \text{if } x = 0, \\ > 0, & \text{if } x \neq 0, \end{cases}$$
 $\left\{ \text{positive definite} \right\}$
(2)
$$(x,y) = (y,x)$$
 $\left\{ \text{symmetric} \right\}$
(3)
$$(x+y,z) = (x,z) + (y,z)$$
 $\left\{ \text{linear} \right\}$
(4)
$$(\alpha x, y) = \alpha(x, y)$$
 $\left\{ \text{linear} \right\}$

$$(2) \qquad (x,y) = (y,x)$$

(3)
$$(x+y,z) = (x,z) + (y,z)$$

(4)
$$(\alpha x, y) = \alpha(x, y)$$

Prove them by direct calculation! The next theorem gives a bound for the absolute value of the scalar product.

Theorem. (Cauchy's inequality) If $x, y \in \mathbf{R}^n$ then

$$|(x,y)| \le ||x|| ||y||$$

Proof. If one of the x, y is the zero vector then both sides of the inequality are equal to zero (the number zero), so the inequality is true that case. We then assume that one of them is not the zero vector, for example, $y \neq 0$. Then, for any $\alpha \in \mathbf{R}$, we calculate using the properties (2), (3), (4),

$$0 \le \|x + \alpha y\|^2 = (x + \alpha y, x + \alpha y) = (x, x) + 2\alpha(x, y) + \alpha^2(y, y) = \|x\|^2 + 2\alpha(x, y) + \alpha^2 \|y\|^2.$$

Since $y \neq 0$ implies $||y||^2 = (y, y) \neq 0$, see (1), we may choose

$$\alpha = -\frac{(x,y)}{\|y\|^2}$$

to get

$$0 \le \|x\|^2 - 2\frac{(x,y)^2}{\|y\|^2} + \frac{(x,y)^2}{\|y\|^4} \|y\|^2 = \|x\|^2 - \frac{(x,y)^2}{\|y\|^2}$$

or

$$(x,y)^2 \le ||x||^2 ||y||^2$$

This implies $|(x, y)| \le ||x|| ||y||$.

The norm has the following properties.

Theorem.

(5)
$$||x|| \begin{cases} = 0, & \text{if } x = 0, \\ > 0, & \text{if } x \neq 0, \end{cases}$$

$$\|\alpha x\| = |\alpha| \|x|$$

(7)
$$||x+y|| \le ||x|| + ||y|| \quad \left\{ triangle \ inequality \right\}$$

Proof. The first statement (5) is the same as (1). To prove the other two we compute with the square of the norm:

$$\begin{aligned} \|\alpha x\|^2 &= (\alpha x, \alpha x) = \alpha^2 (x, x) = \alpha^2 \|x\|^2 \\ \|x + y\|^2 &= (x + y, x + y) = \|x\|^2 + 2(x, y) + \|y\|^2 \le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \\ \text{In the last one we used Cauchy's inequality.} \end{aligned}$$

From Cauchy's inequality it follows that

$$-1 \le \frac{(x,y)}{\|x\| \|y\|} \le 1$$

so we can define the angle θ between x and y by the equation:

$$\cos(\theta) = \frac{(x,y)}{\|x\| \|y\|}$$

The most important case is when $\theta = \pi/2$. We say that the vectors x, y are orthogonal ("ortogonala, vinkelräta") if

$$(x, y) = 0$$

2. The Gauss elimination method

We shall solve a linear system of m equations in n unknowns:

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases}$$

We will use the Gauss elimination method, which is a systematic way of eliminating one variable after the other. We explain it by example.

Example.

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 1\\ 4x_1 + 3x_2 + 2x_3 + x_4 = 1\\ -10x_1 - 5x_2 + 5x_4 = -1 \end{cases}$$

We begin by eliminating x_1 from equations 2 and 3. This is achieved by multiplying equation 1 by -4 and adding it to equation 2. Similarly, we multiply equation 1 by 10 and add it to equation 3.

$$\begin{bmatrix} x_1 \\ +2x_2 \\ +3x_3 \\ +4x_4 \\ +3x_2 \\ +2x_3 \\ +x_4 \\ =1 \\ \swarrow \\ -10x_1 \\ -5x_2 \\ +5x_4 \\ =-1 \\ \swarrow \\ \checkmark$$

This results in

$$\begin{cases} \hline x_1 \\ -5x_2 \\ 15x_2 \\ -10x_3 \\ -15x_4 \\ -$$

Next we clean up by multiplying equation 2 by $-\frac{1}{5}$ and equation 3 by $\frac{1}{15}$. This results in

$$\begin{cases} \boxed{x_1} + 2x_2 + 3x_3 + 4x_4 = 1 \\ \hline x_2 + 2x_3 + 3x_4 = \frac{3}{5} & \boxed{-1} \\ x_2 + 2x_3 + 3x_4 = \frac{3}{5} & \checkmark \end{cases}$$

Finally, we eliminate x_2 from equation 3 by multiplying equation 2 by -1 and adding it to equation 3:

$$\begin{cases} \boxed{x_1} +2x_2 +3x_3 +4x_4 =1 \\ \hline x_2 +2x_3 +3x_4 =\frac{5}{5} \\ 0 =0 \end{cases}$$

We now have

 $\begin{cases} \boxed{x_1} + 2x_2 + 3x_3 + 4x_4 = 1 \\ \boxed{x_2} + 2x_3 + 3x_4 = \frac{3}{5} \end{cases}$

The boxed terms in the equations are called *pivots*. This system has the same solutions as the original one. We can easily solve it as follows. The variable associated with pivots, here x_1 and x_2 are called *bound variables* ("bundna variabler"). The other variables, here x_3, x_4 are called *free variables*. They can be chosen arbitrarily:

$$x_3 = s, \quad x_4 = t$$

where s, t are arbitrary numbers. We then solve for x_2 from equation 2:

$$x_2 = \frac{3}{5} - 2x_3 - 3x_4 = \frac{3}{5} - 2s - 3t$$

and then x_1 is given by equation 1:

$$x_1 = 1 - 2x_2 - 3x_3 - 4x_4 = -\frac{1}{5} + s + 2t$$

We conclude

$$x_{1} = s + 2t - \frac{1}{5}$$

$$x_{2} = -2s - 3t + \frac{3}{5}$$

$$x_{3} = s$$

$$x_{4} = t$$

where s, t are arbitrary numbers.

Example. The previous calculations are easier to do if we use matrix notation. The system of equations can be written:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ -10 & -5 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

It is of the form

$$Ax = b, \quad \text{with } A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ -10 & -5 & 0 & 1 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

The idea is that it is not necessary to keep writing the x's when we perform the calculations. We form the extended matrix:

and perform the eliminations in the form of *row operations* on this matrix. The first step is to multiply row 1 by -4 and add it to row 2. Then we multiply row 1 by 10 and add it to row 3. The result is:

$$\begin{bmatrix} 1 & 2 & 3 & 4 & | & 1 \\ 0 & -5 & -10 & -15 & | & -3 \\ 0 & 15 & 30 & 45 & | & 9 \end{bmatrix} \begin{bmatrix} -\frac{1}{5} \\ \frac{1}{15} \end{bmatrix}$$

We then multiply rows 2 and 3 by $-\frac{1}{5}$ and $\frac{1}{15}$. The result is

$$\begin{bmatrix} 1 & 2 & 3 & 4 & | & 1 \\ 0 & 1 & 2 & 3 & | & -\frac{3}{5} \\ 0 & 1 & 2 & 3 & | & -\frac{3}{5} \end{bmatrix}$$

Next we multiply row 2 by -1 and add it to row 3:

$$\begin{bmatrix} 1 & 2 & 3 & 4 & | & 1 \\ 0 & 1 & 2 & 3 & | & -\frac{3}{5} \\ 0 & 1 & 2 & 3 & | & -\frac{3}{5} \end{bmatrix} \begin{bmatrix} -1 \\ \swarrow \end{bmatrix}$$

The result is

$$[\hat{A}|\hat{b}] = \begin{bmatrix} 1 & 2 & 3 & 4 & | & 1 \\ 0 & 1 & 2 & 3 & | & -\frac{3}{5} \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix}$$

LECTURE 5.1

We have now finished the elimination process. The matrix is an *echelon matrix* ("trappstegsmatris"). The boxed matrix elements are called *pivot elements* ("pivå-element"). The system $\hat{A} = \hat{b}$ has the same solutions as the original system Ax = b. It is easily solved if we write it out:

$$\begin{cases} \boxed{x_1} + 2x_2 + 3x_3 + 4x_4 = 1 \\ \hline x_2 + 2x_3 + 3x_4 = \frac{3}{5} \end{cases}$$

and calculate as before. The result is

$$x = \begin{bmatrix} s + 2t - \frac{1}{5} \\ -2s - 3t + \frac{3}{5} \\ s \\ t \end{bmatrix} = s \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -\frac{1}{5} \\ \frac{3}{5} \\ 0 \\ 0 \end{bmatrix}$$

There are infinitely many solutions: for each s, t we get a solution. The solutions linear combinations of the three vectors:

$$v_1 = \begin{bmatrix} 1\\ -2\\ 1\\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 2\\ -3\\ 0\\ 1 \end{bmatrix}, \quad v_3 = \begin{bmatrix} -\frac{1}{5}\\ \frac{3}{5}\\ 0\\ 0 \end{bmatrix}$$

Note the structure of the echelon matrix: there are zeros below the pivot elements. In each step of the algorithm we use a pivot element in order to obtain zeros below it. Therefore the pivot elements must not be zero. It is always possible to make the pivots =1 as in the example, but this is not necessary. It is possible to make the elements above the pivots =0. We show this now:

$$\begin{bmatrix} 1 & 2 & 3 & 4 & | & 1 \\ 0 & 1 & 2 & 3 & | & -\frac{3}{5} \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix} \begin{bmatrix} -2 \\ -2 \end{bmatrix}$$

The result is:

$$\begin{bmatrix} 1 & 0 & -1 & -2 & | & -\frac{1}{5} \\ 0 & 1 & 2 & 3 & | & -\frac{3}{5} \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix}$$

An echelon matrix with zeros also above the pivots and with the pivots =1 is said to have row reduced echelon form ("radreducerad trappstegsform").

In MATLAB we have the function **rref()**. So the previous calculation can done as follows.

>> A=[1 2 3 4; 4 3 2 1; -10 -5 0 1] >> b=[1; 1; -1] >> AA = rref([A b])

Note the the extended matrix [A b].

Example. We solve Ax = c with the same matrix A and

$$c = \begin{bmatrix} 1\\1\\1 \end{bmatrix}$$

If we apply Gaussian elimination (in MATLAB: rref([A c])) we get the row reduced echelon matrix:

$$[\hat{A}|\hat{c}] = \begin{bmatrix} 1 & 0 & -1 & -2 & | & 0 \\ 0 & 1 & 2 & 3 & | & 0 \\ 0 & 0 & 0 & 0 & | & 1 \end{bmatrix}$$

and the system $\hat{A}x = \hat{c}$:

$$\begin{cases} x_1 & -x_3 & -2x_4 & =0 \\ & x_2 & +2x_3 & +3x_4 & =0 \\ & & 0 & =1 \end{cases}$$

Clearly, this has no solution, and hence the original system has no solution either.

Example. We can solve the previous systems Ax = b and Ax = c at the same time if we form the extended matrix [A|bc] and use Gauss elimation (in MATLAB: rref([A b c])) to get

$$[\hat{A}|\hat{b}\,\hat{c}] = \begin{bmatrix} 1 & 0 & -1 & -2 & | & -\frac{1}{5} & 0 \\ 0 & 1 & 2 & 3 & | & \frac{3}{5} & 0 \\ 0 & 0 & 0 & 0 & | & 0 & 1 \end{bmatrix}$$

Example.

Here we use the first pivot element 2 to get zeros below it:

$$\begin{bmatrix} 2 & 2 & 4 & 4 \\ 0 & 0 & -8 & -8 \\ 0 & 1 & -2 & -1 \end{bmatrix}$$

Now we must interchange the two last rows in order to get a non-zero pivot in the right place:

$$\begin{bmatrix} 2 & 2 & 4 & 4 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & -8 & -8 \end{bmatrix}$$

This matrix is in echelon form with the pivots marked with boxes. If we want the row reduced echelon form we have to continue with a few more row operations. The result is

$$\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Summary of Gauss elimination. In Gauss elimination we perform three kinds of *row operations*:

- interchange two rows;
- multiply a row by a number;
- add a multiple of one row to another row.

By these operations we can obtain an echelon matrix. The corresponding system has the same solutions as the original system. The echelon system can be solved easily.

Problems

Problem 1. Let

$$x = \begin{bmatrix} 1\\1\\1\\1 \end{bmatrix}, \quad y = \begin{bmatrix} 1\\2\\1\\2 \end{bmatrix}$$

(a) Compute ||x||, ||y||, (x, y). (b) Determine a number t so that z = x + ty is orthogonal to x.

Problem 2. Compute the angle between
$$\begin{bmatrix} -2\\5\\4\\6 \end{bmatrix}$$
 and $\begin{bmatrix} 1\\2\\-4\\2 \end{bmatrix}$

LECTURE 5.1

Problem 3. Prove Pythagoras' theorem: if x, y are orthogonal then $||x + y||^2 = ||x||^2 + ||y||^2$. Problem 4. Solve the following systems of equations.

$$(a) \begin{cases} x_1 - 8x_2 = 3\\ 2x_1 + x_2 = 1\\ 4x_1 + 7x_2 = -4 \end{cases} (b) \begin{cases} 3x_1 - 4x_2 + 5x_3 = 0\\ 7x_1 - 2x_2 - x_3 = 0\\ 2x_1 + x_2 - 3x_3 = 0 \end{cases} (c) \begin{cases} 2x_1 + 3x_2 - x_3 - x_4 = 0\\ x_1 - x_2 - 2x_3 - 4x_4 = 0\\ 3x_1 + x_2 + 3x_3 - 2x_4 = 0 \end{cases}$$
$$(d) \begin{cases} x_1 - 2x_2 + 4x_3 + x_4 = 1\\ 2x_1 + x_2 + 3x_3 - 2x_4 = 2\\ x_1 + 8x_2 - 6x_3 - 7x_4 = 1 \end{cases}$$

Answers

1. (a) ||x|| = 2, $||y|| = \sqrt{10}$, (x, y) = 6. (b) t = -32. $\arccos(4/45)$ 4. (a) no solution.

(b) $x = t \begin{bmatrix} 5 \\ -3 \\ -2 \\ 3 \end{bmatrix}$ (c) $x = t \begin{bmatrix} 7 \\ 19 \\ 11 \end{bmatrix}$ (d) $x = s \begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 3 \\ 4 \\ 0 \\ 5 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

/stig