2. Matrices, systems of linear equations

Tento učební text byl podpořen z Operačního programu Praha - Adaptabilita

Irena Sýkorová

OPPA
PRAHA PRAGUE PRAGA PRAG
EVROPSKÁ UNIE
Evropský sociální fond Praha & EU: Investujeme do vaší budoucnosti

Basic properties of matrices

Let us have a *square* matrix $A = (a_{ij})_{n\times n}$. The matrix A is called **symmetric** if $A = A'$, i.e.

$$a_{ij} = a_{ji} \quad \text{for } i, j = 1, \ldots, n.$$

A matrix A is called **positively definite** if all its main minors are positive, i.e.

$$D_1 > 0, \ldots, D_n > 0, \qquad \text{where} \quad D_i = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1i} \\ a_{21} & a_{22} & \ldots & a_{2i} \\ \ldots & \ldots & \ldots & \ldots \\ a_{i1} & a_{i2} & \ldots & a_{ii} \end{vmatrix}$$

A matrix A is called **diagonally dominant** if the absolute value of the term on the diagonal is grater or equal to the sum of absolute values of the other terms either for all rows or for all columns, i.e.

$$|a_{ii}| \geq \sum_{i\neq j} |a_{ij}| \qquad \text{pro} \quad i = 1, \ldots, n.$$

A matrix is **strictly diagonally dominant** if the inequalities are strict.

Example 1: Determine whether the matrix A is symmetric, positively definite, diagonally dominant,

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

The matrix A *is not* symmetric, because $a_{12} \neq a_{21}$. For determining the definiteness we compute the minors

$$D_1 = 4\ (>0), \qquad D_2 = \begin{vmatrix} 4 & -1 \\ 2 & 3 \end{vmatrix} = 14\ (>0), \qquad D_3 = \begin{vmatrix} 4 & -1 & 0 \\ 2 & 3 & -2 \\ 0 & -2 & 3 \end{vmatrix} = 26\ (>0).$$

All the minors are positive, thus the matrix A *is* positively definite.

For the diagonal dominance we compare the terms on the diagonal and sums of the rests first for the rows:

$$\begin{array}{ccccc} |\,4\,| & > & |-1| & + & |\,0\,| \\ |\,3\,| & < & |\,2\,| & + & |-2| \\ |\,3\,| & > & |-2| & + & |\,0\,| \end{array}$$

the condition is not satisfied for the second row, let us try the columns:

$$\begin{array}{ccccc} |\,4\,| & > & |\,2\,| & + & |\,0\,| \\ |\,3\,| & \geq & |-1\,| & + & |-2| \\ |\,3\,| & > & |-2| & + & |\,0\,| \end{array}$$

The condition holds for all the columns, thus the matrix A *is* diagonally dominant, but it is not strictly diagonally dominant (the inequality in the second column is not strict).

MATLAB

proof that A is symmetric, i.e. $A = A'$ or $A - A' = O$ (zero matrix):

```
>> A-A'
```

or comparison – result: 1 (A is symmetric), 0 (A is not symmetric)

```
>> isequal(A,A')
```

proof that the matrix A is positively definite – all minors are positive

```
>> A(1,1)
>> det(A(1:2,1:2))
>> det(A(1:3,1:3))
>> ...
>> det(A)
```

Eigenvalues of matrices

Let us have a *square* matrix $A = (a_{ij})_{n\times n}$. If for a real number λ (generally complex) and a *non-zero* vector $\bar{x}$ the following condition is satisfied

$$A\bar{x} = \lambda\bar{x}\,, \tag{1}$$

the number λ is called **eigenvalue** of the matrix A and the vector $\bar{x}$ is **eigenvector** of A assigned to λ.

The equation (1) can be transformed, J is the unit (identity) matrix of the order n

$$\begin{aligned} A\bar{x} &= \lambda\bar{x}, \\ A\bar{x} - \lambda\bar{x} &= \bar{o}, \\ (A-\lambda J)\,\bar{x} &= \bar{o}. \end{aligned} \tag{2}$$

This is a matrix form of the homogeneous system of n linear equations with n unknowns with the matrix $A - \lambda J$. We look for the non-zero vector $\bar{x}$, i.e. non-zero solution of the system, which exists iff the matrix of the system is singular, i.e.

$$\det(A - \lambda J) = 0. \tag{3}$$

The matrix $A - \lambda J$ is called **characteristic matrix**, polynomial $p(\lambda) = \det(A - \lambda J)$ is **characteristic polynomial** (of order n) and the equation (3) is **characteristic equation**. We find the roots of the polynomial of n-th order, and, thus there are n (real or complex, maybe multiple) eigenvalues of any matrix A.

Spectral radius of a matrix A is the greatest of absolute values of eigenvalues

$$\rho(A) = \max\{|\lambda_i|,\ \text{where } \lambda_i \text{ is an eigenvalue of } A\}.$$

If we know the eigenvalue λ then its eigenvector is every non-zero solution of the system (2).

Example 2: Find eigenvalues and spectral radius of the matrix

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

First form the characteristic matrix A-$\lambda J = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{pmatrix}$,
then the characteristic polynomial

$$\det(A - \lambda J) = \begin{vmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda) - 6 = \lambda^2 - 3\lambda - 4.$$

Eigenvalues are the roots of the characteristic polynomial $\lambda^2 - 3\lambda - 4 = 0$, thus the eigenvalues of A are $\lambda_1 = 4$, $\lambda_2 = -1$.

The spectral radius is the greatest of absolute values of the eigenvalues:

$$\max\{|\lambda_1|, |\lambda_2|\} = \max\{|\ 4\ |, |-1|\} = \max\{4, 1\} = 4.$$

The spectral radius is $\rho(A) = 4$.

Remark

(1) "a)" If the given matrix is *symmetric*, then all its eigenvalues are real.
(2) "b)" A matrix is *symmetric* and *positively definite* iff all its eigenvalues are positive.

MATLAB

eigenvalues of A:

```
>> eig(A)
```

stectral radius of A (maximum of absolute values of eigenvalues):

```
>> max(abs(eig(A)))
```

Norm of matrix

Norm of matrix is a mapping $\mathbb{R}^{n\times n}$ into $\mathbb{R}$ that satisfies the following conditions:

(1) "(i)" $\|A\| > 0$ for every *non-zero* square matrix and $\|A\| = 0$ only for $A = O$,
(2) "(ii)" $\|c \cdot A\| = |c| \cdot \|A\|$ for every real number c,
(3) "(iii)" $\|A + B\| \leq \|A\| + \|B\|$ for every pair of matrices of the same type,
(4) "(iv)" $\|A \cdot B\| \leq \|A\| \cdot \|B\|$ for every pair of matrices of the same type.

We say that a norm of matrix is **consistent** with a norm of vector if for each matrix A and vector $\bar{x}$ we have

$$\|A\bar{x}\| \leq \|A\| \cdot \|\bar{x}\|.$$

Let $\|\cdot\|$ be a norm of vector, then the norm

$$\|A\| = \sup_{\bar{x}\neq\bar{o}} \frac{\|A\bar{x}\|}{\|\bar{x}\|} = \sup_{\|\bar{x}\|=1} \|A\bar{x}\| \tag{4}$$

is **generated** by $\|\cdot\|$.

A norm of matrix generated by the norm of vector is consistent with this norm.

Examples of matrix norms:

(1) "a)" $\|A\|_\infty = \max\limits_{1\leq i\leq n} \sum_{j=1}^n |a_{ij}|$, **row sum, infinity** norm,
(2) "b)" $\|A\|_1 = \max\limits_{1\leq j\leq n} \sum_{i=1}^n |a_{ij}|$, **column sum, 1-** norm,
(3) "c)" $\|A\|_F = \left(\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2\right)^{\frac{1}{2}}$, **Frobenius, Schur** norm,
(4) "d)" $\|A\|_2 = \sqrt{\rho(A'A)}$, **spectral, Euclidean** norm.

Example 3: Compute all the norms of the matrix $A = \begin{pmatrix} 1 & -5 \\ 2 & 6 \end{pmatrix}$.

a)

$$\|A\|_\infty = \max\{|1| + |-5|, |2| + |6|\} = \max\{6, 8\} = 8.$$

b)

$$\|A\|_1 = \max\{|1| + |2|, |-5| + |6|\} = \max\{3, 11\} = 11.$$

c)

$$\|A\|_F = \sqrt{(1)^2 + (2)^2 + (-5)^2 + (6)^2} = \sqrt{1+4+25+36} = \sqrt{66} = 8,124.$$

d)

$$A'A = \begin{pmatrix} 1 & 2 \\ -5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & -5 \\ 2 & 6 \end{pmatrix} = \begin{pmatrix} 5 & 7 \\ 7 & 61 \end{pmatrix},$$

then find eigenvalues:

$$(5-\lambda)(61-\lambda) - 49 = 0 \quad \Leftrightarrow \quad \lambda^2 - 66\lambda + 256 = 0 \quad \Leftrightarrow \quad \lambda_1 = 61,8617, \ \lambda_2 = 4,1383.$$

Then

$$\rho(A'A) = 61,8617 \quad \Rightarrow \quad \|A\|_2 = \sqrt{61,8617} = 7,8652.$$

MATLAB – norms of a matrix A:

row sum norm:

```
>> norm(A,inf)
>> max(sum(abs(A')))
```

column sum norm:

```
>> norm(A,1)
>> max(sum(abs(A)))
```

Frobenius norm:

```
>> norm(A,'fro')
>> sqrt(sum(diag(A'* A)))
```

spectral norm:

```
>> norm(A,2)
>> norm(A)
```

Remark: Frobenius norm of matrix is not generated by Euclidean norm of vector, in fact for the unit matrix J of order n we have $\|J\|_F = \sqrt{n}$, but from (4) we have $\|J\|_F = 1$. Euclidean norm of vector generates spectral norm of matrix.

If the norm of a matrix is consistent with a norm of vector the following condition holds: $\rho(A) \leq \|A\|$.

Conditionality of matrix

Example 4: Let us have the linear system

$$2x + 6y = 8$$
$$2x + 6,0001y = 8,0001,$$

with the solution $x = 1$, $y = 1$. Consider a small change in coefficients (10^{-4}), and get the system

$$2x + 6y = 8$$
$$2x + 5,9999y = 8,0002$$

with the solution $x = 10$, $y = -2$. Thus a very small change in coefficients and right sides imply a big change in the solution.

A matrix is called **well conditioned** if relatively small changes in coefficients and right sides imply relatively small changes in solution.

A matrix is called **badly conditioned** if relatively small changes in coefficients and right sides imply relatively big changes in solution.

Consider a linear system

$$A\bar{x} = \bar{b} \tag{5}$$

with a regular (non-singular, invertible) matrix A. Denote $\bar{x}^*$ the exact solution of the system $A\bar{x} = \bar{b}$ and $\bar{x}_c$ is the exact solution of the changed system $(A + \delta A)\bar{x} = \bar{b} + \delta\bar{b}$, then the estimation of the relative error of the solution $\frac{\|\bar{x}_c - \bar{x}^*\|}{\|\bar{x}^*\|}$ depends directly on the product

$$K(A) = \|A\| \cdot \|A^{-1}\|,$$

where the norm of matrix is generated by norm of vector. The number $C(A)$ is called **condition number** of the matrix A. The bigger it is the bigger the estimation of the relative error is.

The matrix $A = \begin{pmatrix} 2 & 6 \\ 2 & 6,0001 \end{pmatrix}$ from the foregoing example has a big condition number $C(A) = 4 \cdot 10^5$.

MATLAB

condition number of a matrix A:

```
>> cond(A)
```

Direct methods for solving linear systems

We will solve a linear system

$$A\bar{x} = \bar{b},$$

where $A = (a_{ij})$ is a real *regular* matrix of the order n, $\bar{b} = (b_1, b_2, \ldots, b_n)'$ is the column vector of right sides.

Direct methods: Gaussian elimination , Jordan (Gauss-Jordan) elimination, method with inverse matrix, Cramer's rule,....

Method with inverse matrix:

When we know A^{-1} we have

$$\bar{x} = A^{-1}\bar{b},$$

but sometimes finding A^{-1} is difficult.

Gaussian elimination:

We transform the augmented matrix of the given system $(A|b)$ using elementary row operations to $(U|y)$, where U is upper triangular matrix – **forward process**.

Then we solve the system

$$U\bar{x} = \bar{y},$$

from down we compute the coordinates of $\bar{x}$ – **backward process**.

Example 5: Solve the system by using the Gaussian elimination

$$\begin{aligned} x_1 + 2x_2 + 2x_3 &= 2 \\ 2x_1 + 2x_2 + 3x_3 &= 5 \\ -x_1 - x_2 &= -1. \end{aligned} \tag{6}$$

een: We transform the augmented matrix of the system into a upper triangular form. In the k-th step we multiply the k-th row by a constant add this row to the others to eliminate the terms under the diagonal (constants are on the right-hand side of the rows). If the term on the diagonal is zero we have to exchange the rows.

$$\left(\begin{array}{ccc|c} 1 & 2 & 2 & 2 \\ 2 & 2 & 3 & 5 \\ -1 & -1 & 0 & -1 \end{array}\right) \begin{array}{cc} (-2) & (1) \\ & \\ & \end{array} \sim \left(\begin{array}{ccc|c} 1 & 2 & 2 & 2 \\ 0 & -2 & -1 & 1 \\ 0 & 1 & 2 & 1 \end{array}\right) (\tfrac{1}{2}) \sim \left(\begin{array}{ccc|c} 1 & 2 & 2 & 2 \\ 0 & -2 & -1 & 1 \\ 0 & 0 & \frac{3}{2} & \frac{3}{2} \end{array}\right)$$

From here we easily get $x_3 = 1$, $x_2 = -1$, $x_1 = 2$.

In **Gauss-Jordan elimination** we transform A into a unit matrix and the solution is in the right-hand column.

Elementary row operations can be done by multiplying the given matrix with special regular matrices from the left.

If we add an m-multiple of i-th row to the j-th row in a matrix M, we get the product VM (not MV), where V is the unit matrix with m instead of v_{ji}. If $j > i$ then V is lower triangular.

Example 6: In the matrix M add the 4-multiple of the 1st row to the 3rd row ($i = 1$, $j = 3$, $m = 4$).

$$M = \begin{pmatrix} -1 & 2 & 0 \\ 3 & 2 & 2 \\ 4 & -5 & 2 \end{pmatrix} \quad V = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix}$$

$$V \cdot M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 2 & 0 \\ 3 & 2 & 2 \\ 4 & -5 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 2 & 0 \\ 3 & 2 & 2 \\ 0 & 3 & 2 \end{pmatrix}$$

If we exchange the i-th and j-th rows in M, we get the product PM (not MP), where P is the unit matrix where the i-th and j-th rows are exchanged. The matrix P is called **permutation** matrix.

Example 6: In M exchange the first and second rows ($i = 1$, $j = 2$).

$$M = \begin{pmatrix} -1 & 2 & 0 \\ 3 & 2 & 2 \\ 4 & -5 & 2 \end{pmatrix} \quad P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$P \cdot M = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 2 & 0 \\ 3 & 2 & 2 \\ 4 & -5 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 2 \\ -1 & 2 & 0 \\ 4 & -5 & 2 \end{pmatrix}$$

We will use **LU decomposition, LU factorization** of a matrix, which means that we express a matrix A in the form of product

$$A = LU,$$

where L is lower triangular matrix with ones on diagonal and U is upper triangular matrix. The system (5) will be solved in two steps - two systems with triangular matrices:

$$L\bar{y} = \bar{b}$$

$$U\bar{x} = \bar{y}.$$

Triangular matrices L and U can be found by using the elementary row operations on A, symbolically:

$$(A|J) \sim \ldots \sim (U|L'),$$

where L' is lower triangular with $L'A = U$. We can just add a real multiple of a row to a row *below*. From $L'A = U$ we get

$$A = (L')^{-1}U = LU.$$

We can interpret this process as the (step-by-step) multiplication of A with V_k

$$L'A = V_k \ldots V_1 A = U \quad \Rightarrow \quad A = (L')^{-1}U = (V_k \ldots V_1)^{-1}U = (V_1^{-1} \ldots V_k^{-1})\,U = LU.$$

Every L has ones on the diagonal and coefficients from the Gaussian elimination with $-$ below.

Example 7: Solve the system (6) by using LU decomposition.

The matrix (6) can be decomposed

$$A = LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -\frac{1}{2} & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 2 \\ 0 & -2 & -1 \\ 0 & 0 & \frac{3}{2} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 2 & 3 \\ -1 & -1 & 0 \end{pmatrix}.$$

Then we solve $L\bar{y} = \bar{b}$

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix},$$

now we get from the first row $y_1 = 2$, and after substituting we get $y_2 = 1$ from the second row, and from the third row get $y_3 = \frac{3}{2}$, thus

$$\bar{y} = \begin{pmatrix} 2 \\ 1 \\ \frac{3}{2} \end{pmatrix}.$$

Now solve the system $U\bar{x} = \bar{y}$, which is the backward process of the Gaussian elimination:

$$\begin{pmatrix} 1 & 2 & 2 \\ 0 & -2 & -1 \\ 0 & 0 & \frac{3}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ \frac{3}{2} \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$$

Even if A is regular we can have zeros on the diagonal. We have to exchange rows first, and then decompose the obtained matrix:

$$PA = LU.$$

The solution of such a system:

$$L\bar{y} = P\bar{b}$$
$$U\bar{x} = \bar{y}.$$

MATLAB – direct methods for solving linear system
(A – matrix of szstem, b – *column* vector of right sides):

Matlab command:

```
>> x = A \ b
```

with inverse matrix:

```
>> A1 = inv(A)
>> x = A1 * b
```

with LU decomposition (solve $Ly = Pb$, $Ux = y$, where P is permutation matrix):

```
>> [L, U, P] = lu(A)
>> x = U \ ( L \ (P*b))
```
