Rough Syllabus, Part 6
Linear mappings

1. Definition. Let H and G be vector spaces over the same number field K, with dimension of n, and m, respectively. The mapping T $(H \to G)$ is called a **linear mapping**, if with notations $T(x_1) = y_1$, $T(x_2) = y_2$, $x_1, x_2 \in H$, $y_1, y_2 \in G$

$$T(\alpha x_1 + \beta x_2) = \alpha T(x_1) + \beta T(x_2) = \alpha y_1 + \beta y_2.$$

1. Theorem. In the case of a linear mapping the image of the space H is a subspace of G. It is denoted by $Im(T)$.

2. Theorem. The set of vectors in H for which $T(x) = 0$ form a subspace of H, and it is denoted by $Ker(T)$.

Examples: Projections of vectors in $\mathcal{R}^3$ perpendicularly on a line (constant multiple of a given vector) or on a plane (linear combination of two nonparallel vectors); Definite integral of continuous functions on $[a, b]$; Derivation of functions; For polynomials of degree not more than two $p(x) \mapsto (x + 2)p'(x)$; If $\mathbf{a}$ is a given vector of $\mathcal{R}^3$, then $\mathbf{r} \mapsto \mathbf{a} \times \mathbf{r}$.

If the values of the linear mapping T are given on a basis (fundamental system), then the mapping is defined uniquely on the total space, since the linear combinations are given uniquely:
Let a_1, a_2,...,a_n be a basis in H, and let their images be
$T(a_i) = c_i \in G$. (The vectors c_i can be also linearly dependent, e.g. any or each of them can be the o vector.) If $x = \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n$, then
$T(x) = T(\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n) = \alpha_1 c_1 + \alpha_2 c_2 + \cdots + \alpha_n c_n$

Let denote the dimension of a linear space A by $\dim(A)$. Assume $\dim(T) = n$, $\dim(G) = m$. If there are bases in both spaces, then the vectors in both spaces can be given by coordinates, and the image of vector x is:

$$T(\mathbf{x}) = C\mathbf{x} \quad \in G,$$

where the columns of matrix C are the images of the basis of H given by their coordinates with respect to the basis in G. Vector x is given by its coordinates with respect to the basis in H consisting of vectors $\{a_i\}$.

3. Theorem. The dimension of $Im(T) \subseteq G$ is equal to the rank of its matrix C, i.e. $\dim(Im(T)) = \rho(C)$.

4. Theorem. The dimension of $Ker(T) \subseteq H$ is equal to $(\dim(H) - \rho(C))$, i.e. $\dim(Ker(T)) = \rho(C)$, where C is the matrix of T.

Obviously, when we change the basis in any of the spaces, the matrix of the mapping C will change as well, but its rank stays the same.

If $H = G$, then we apply the same basis, and in this case matrix C is quadratic. How matrix C is changing, when the basis is changed? We know that the new coordinates can be get by the multiplication of matrix B^{-1}, where the columns of B are the vectors of the new basis given by coordinates with respect to the old one. By simple formal change:

$$\begin{aligned} Cx &= y \\ B^{-1}Cx &= B^{-1}y \\ B^{-1}CBB^{-1}x &= B^{-1}y \\ (B^{-1}CB)(B^{-1}x) &= (B^{-1}y). \end{aligned}$$

Since $B^{-1}x$ and $B^{-1}y$ contain the coordinates of the independent variable and of the image with respect to the new basis, the matrix of the mapping with respect to the new basis is $S = B^{-1}CB$ obviously.

In the general case, similarly: If T is a linear mapping $H \to G$ whose matrix is C, then changing the basis in H where the coordinates of the new basis with respect to the old one are given in B_H, then the matrix of the mapping changes into CB_H, while changing the basis in G, where the coordinates of the new basis with respect to the old one are given in B_G, then the matrix of the mapping T changes into $B_G^{-1}C$.

The linear mappings $H \to H$ are called linear transformations.

2. Definition. Let A be a linear transformation. If a subspace $V \subseteq H$ is such that for every $\mathbf{x} \in V \quad A\mathbf{x} \in V$, then V is called an invariant subspace of A.

3. Definition. Let A be a linear transformation. If $A\mathbf{x} = \lambda\mathbf{x}$ for an $\mathbf{x} \neq \mathbf{0}$, then λ is called an **eigenvalue** of transformation A and $\mathbf{x}$ is an **eigenvector** of transformation A associated to the eigenvalue λ.

Examples: Projection in $\mathcal{R}^3$ onto a plane; derivation of e^{ax}, ax^n.

Let fix a basis in H. Consider the matrix of the mapping A.

$A\mathbf{x} = \lambda\mathbf{x}$

$A\mathbf{x} - \lambda\mathbf{x} = (A - \lambda I)\mathbf{x} = \mathbf{0}$.

4. Definition. Let A be a quadratic matrix. If $A\mathbf{x} = \lambda\mathbf{x}$ for an $\mathbf{x} \neq \mathbf{0}$, then λ is called an **eigenvalue** of matrix A and $\mathbf{x}$ is an **eigenvector** of matrix A associated to the eigenvalue λ.

This homogeneous linear system has a nontrivial solution only if the columns of A are linearly dependent, i.e.
$\det(A - lambdaI) = |A - \lambda I| = 0$. This determinant is a polynomial of degree n, so it can not have more different roots than n.

5. Definition. The polynomial $\det(A - lambdaI)$ is called the characteristic polynomial of A.

The roots of the characteristic polynomial are the eigenvalues of A. After the roots λ_1, λ_2, $\ldots \lambda_k$ are determined, the associated eigenvectors can be found from the equations $(A - \lambda_i I)\mathbf{x} = \mathbf{0}$.

5. Theorem. Some statements about eigenvectors:
-Eigenvectors associated to the same eigenvalue form a subspace. These subspaces are invariant subspaces. -To every eigenvalue there is at least one associated eigenvector (i.e. the subspace of the associated eigenvectors has dimension at least one). -Eigenvectors associated to different eigenvalues are linearly independent. -The number of linearly independent eigenvectors associated to eigenvalue λ_i can not exceed the multiplicity of this eigenvalue in the characteristic polynomial, but maybe less.

6. Definition. Matrices A and C are called **similar** if there is an invertable matrix B such that $A = B^{-1}CB$.

Since we have seen above that similar matrices describe the same mapping in different basis, their basic properties are the same. E.g. they have the same eigenvalues with the same number of linearly independent eigenvectors, etc. .

6. Theorem. Similar matrices have the same characteristic polynomial.

Proof: $|B^{-1}AB-\lambda I| = |B^{-1}AB-\lambda B^{-1}B| = |B^{-1}(A-\lambda I)B| = |B^{-1}||A-\lambda I||B| = |A-\lambda I|$
This statement can not be reversed, e.g.:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

For A, every vector of the space is an eigenvector associated to 1, so there are three linearly independent among them, but matrix B has only one linearly independent eigenvector associated to 1.

These two matrices have the same characteristic polynomial but they are not similar, they have a very different structure. The theory of elementary divisors of matrices discusses these questions.

7. Theorem. Cayley-Hamilton
Every matrix satisfies its characteristic polynomial, i.e. $p(A) = 0$.

If we restrict our investigations for real matrices, then it can happen that a matrix does not have any eigenvalue and eigenvectors, since there are real polynomials without any real root. Every real polynomial with leading coefficient 1 can be the characteristic polynomial of a matrix:

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = (-1)^n|A-\lambda I|$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & -a_3 & \dots & -a_{n-2} & -a_{n-1} \end{bmatrix}$$

If we consider matrices as complex matrices, then their characteristic polynomials can be decomposed into a product of n linear factors. I.e. considering the complex case, every matrix has at least one eigenvalue.

There are several theorems and procedures about how to reveal the structure of a matrix, consequently the structure of the corresponding linear transformation, we discuss only a special case of it.

8. Theorem. If a linear transformation has n linearly independent eigenvectors, then the eigenvectors can be chosen as a basis, and in this basis the matrix of this linear transformation is diagonal.

9. Theorem. If A is selfadjoint (in the real case it means symmetric), then all eigenvalues are real, and it has n linearly independent eigenvectors, and they can be chosen orthogonally.

Corollary: Since the length of eigenvectors can be arbitrary (not zero), the eigenvectors of a selfadjoint matrix can be chosen to form an orthonormed basis.

7. Definition. A linear transformation T is called **selfadjoint**, if $\langle \mathbf{v}, T\mathbf{w}\rangle = \langle \mathbf{v}T, \mathbf{w}\rangle$

10. Theorem. The matrix of a selfadjoint transformation is a selfadjoint matrix.

Let B a selfadjoint matrix. Then $(\mathbf{x}^*B\mathbf{x})$ is always real.

8. Definition. -If $\mathbf{x} * B\mathbf{x} > 0$ for every $\mathbf{x} \neq \mathbf{0}$, then B is called **positive definite**;
-If $\mathbf{x} * B\mathbf{x} \geq 0$ for every $\mathbf{x} \neq \mathbf{0}$, then B is called **positive semidefinite**;
-If $\mathbf{x} * B\mathbf{x} < 0$ for every $\mathbf{x} \neq \mathbf{0}$, then B is called **negative definite**;
-If $\mathbf{x} * B\mathbf{x} \leq 0$ for every $\mathbf{x} \neq \mathbf{0}$, then B is called **positive semidefinite**;
-If $\mathbf{x} * B\mathbf{x}$ takes positive and negative values as well, then B is called **indefinite**.

Examples:

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

Observe, the weight matrix discussed at inner product is a positive definite matrix. Furthermore, the eigenvalues of a positive definite matrix are positive numbers, the eigenvalues of a positive semidefinite matrix are positive numbers and zeros, the eigenvalues of a negative definite matrix are negative numbers, the eigenvalues of a negative semidefinite matrix are negative numbers and zeros, an indefinite matrix has positive and also negative eigenvalues.

Positive definite (or semidefinite) matrices have some special properties which can be efficiently used in the case of solving linear system of equations with these type of coefficient matrix. That is why some numerical solution procedures solves the equation $A^*A\mathbf{x} = A^*\mathbf{b}$ instead of $A\mathbf{x} = \mathbf{b}$.

Of course, to find the eigenvalues from the characteristic polynomial of a matrix of order higher then four is not easy. But there is a relatively easy way to define the definiteness of some selfadjoint matrices. Procedure: Let create the sequence of the subdeterminants of the left upper corner. If this sequence has the sequence of signs: $+, +, +, +, , \ldots$, then the matrix is positive definite. If this sequence has the sequence of signs: $-, +, -, +, , \ldots$, then the matrix is negative definite. If there are "misplaced" signs, then the matrix is indefinite.