## 3.1. Classical problems of Linear Algebra

Nonlinear problems, such as finding maxima or inversion of functions, evaluation of areas and volumes, summation of infinite series, etc. are complicated. Differential and Integral Calculus gives us plenty useful hints how to approach such problems, but simple universal recipes among them are rare. To the contrary, Linear Algebra deals with very simple, linear or quadratic functions. Among numerous questions one may ask about such functions there are, roughly speaking, only four basic, similarly formulated problems which Linear Algebra can handle. It is completeness and simplicity of solutions to these problems what makes Linear Algebra efficient in applications. The four model questions and the answers can be described as follows.

*Question* 1. *Given $m$ linear functions in $n$ variables,*

$$\begin{array}{ccc} y_1 & = & a_{11}x_1 + ... + a_{1n}x_n \\ & ... & \\ y_m & = & a_{m1}x_1 + ... + a_{mn}x_n \end{array},$$

*what is the simplest form to which they can be transformed by linear changes of the variables,*

$$\begin{array}{ccc} y_1 & = & b_{11}Y_1 + ... + b_{1m}Y_m \\ ... & & \\ y_m & = & b_{m1}Y_1 + ... + b_{mm}Y_m \end{array}, \quad \begin{array}{ccc} x_1 & = & c_{11}X_1 + ... + c_{1n}X_n \\ ... & & \\ x_n & = & c_{n1}X_1 + ... + c_{nn}X_n \end{array} \quad ?$$

The answer is given by

The Rank Theorem. *Any $m$ linear functions in $n$ variables can be transformed by suitable linear changes of dependent and independent variables to exactly one of the forms:*

$$Y_1 = X_1, ..., Y_r = X_r, Y_{r+1} = 0, ..., Y_m = 0 \quad \text{where } 0 \le r \le m, n.$$

The number $r$ featuring in the answer is called the rank of the set of $m$ linear functions in question.

*Question* 2. *Given a homogeneous quadratic function in $n$ variables,*

$$Q = q_{11}x_1^2 + 2q_{12}x_1x_2 + 2q_{13}x_1x_3 + ... + q_{nn}x_n^2,$$

*what is the simplest form it can be transformed to by a linear change of the variables*

$$\begin{array}{ccc} x_1 & = & c_{11}X_1 + ... + c_{1n}X_n \\ & ... & \\ x_n & = & c_{n1}X_1 + ... + c_{nn}X_n \end{array} \quad ?$$

The Inertia Theorem. *Any homogeneous quadratic function in $n$ variables can be transformed by a suitable linear change of the variables to exactly one of the normal forms:*

$$X_1^2 + ... + X_p^2 - X_{p+1}^2 - ... - X_{p+q}^2 \quad \text{where } 0 \le p + q \le n.$$

The numbers $p$ and $q$ of positive and negative squares in the normal form are called inertia indices of the quadratic function in question. If the quadratic function $Q$ is known to be positive everywhere outside the origin, the Inertia Theorem tells us that in a suitable coordinate system $Q$ assumes the form $X_1^2 + ... + X_n^2$ with the inertia indices $p = n,\ q = 0$.

*Question* 3. *Given two homogeneous quadratic functions $Q(x_1, ..., x_n)$ and $S(x_1, ..., x_n)$ of which the first one is known to be positive everywhere outside the origin, what is the simplest form they can be simultaneously transformed to by a linear change of the variables?*

The Orthogonal Diagonalization Theorem. *Any pair $Q,\ S$ of homogeneous quadratic functions in $n$ variables, of which $Q$ is positive everywhere outside the origin, can be transformed by a linear changes of the variables to exactly one of the normal forms*

$$Q = X_1^2 + ... + X_n^2,\ S = \lambda_1 X_1^2 + ... + \lambda_n X_n^2, \text{ where } \lambda_1 \geq ... \geq \lambda_n.$$

*Question* 4.*Given a constant coefficient system of $n$ linear homogeneous 1-st order ordinary differential equations*

$$\begin{array}{rcl} \dot{x}_1 & = & a_{11}x_1 + ... + a_{1n}x_n \\ & ... & \\ \dot{x}_n & = & a_{n1}x_1 + ... + a_{nn}x_n \end{array},$$

*what is the simplest form to which it can be transformed by a linear change of the phase variables*

$$\begin{array}{rcl} x_1 & = & c_{11}X_1 + ... + c_{1n}X_n \\ & ... & \\ x_n & = & c_{n1}X_1 + ... + c_{nn}X_n \end{array} \quad ?$$

The answer to this question is easier to formulate assuming that the coefficients $a_{ij}$ of the system as well as the coefficients $c_{ij}$ in the change of variables are allowed to be complex numbers.

*Example.* The system of ODEs

$$\begin{array}{lcccccccc} \dot{x}_1 & = & \lambda x_1 & + & x_2 & & & & \\ \dot{x}_2 & = & & & \lambda x_2 & + & x_3 & & \\ & & & & ... & & & & \\ \dot{x}_{m-1} & = & & & & & \lambda x_{m-1} & + & x_m \\ \dot{x}_m & = & & & & & & & \lambda x_m \end{array}$$

is equivalent to the single $m$-th order ODE

$$(\frac{d}{dt} - \lambda)^m y(t) = 0,$$

$$y = x_1,\ \frac{d}{dt}y - \lambda y = x_2,\ (\frac{d}{dt} - \lambda)^2 y = x_3,\ ...,$$

and is called the Jordan cell of size $m$ with the eigenvalue $\lambda$. Let us introduce a Jordan system of several Jordan cells of sizes $m_1, ..., m_r$ with eigenvalues $\lambda_1, ..., \lambda_r$ similarly equivalent to the system

$$(\frac{d}{dt} - \lambda_1)^{m_1} y_1 = 0,\ ...,\ (\frac{d}{dt} - \lambda_r)^{m_r} y_r = 0$$

of $r$ unlinked ODEs of orders $m_1, ..., m_r$.

The Jordan Theorem. *Any constant coefficient system of $n$ linear 1-st order ODEs can be transformed by a complex linear changes of phase variables to exactly one (up to reordering of cells) of the Jordan systems with $m_1 + ... + m_r = n$.*

Note that the classification list in the Jordan Theorem (as well as in the Orthogonal Diagonalization Theorem) is not discrete since Jordan systems depend on the choice of complex numbers $\lambda_1, ..., \lambda_r$. In fact the numbers can be found as the roots of the characteristic polynomial $\det(\lambda I - A)$ of the coefficient matrix $A = [a_{ij}]$ of the original ODE system. In the typical case when all roots are simple all Jordan cells have size 1. Thus we arrive at the following corollary of the Jordan Theorem:

*A typical constant coefficient system of $n$ linear 1-st order ODEs can be transformed by linear changes of phase variables to the form*

$$\dot{X}_1 = \lambda_1 X_1, \ ..., \ \dot{X}_n = \lambda_n X_n.$$

That's about it. One may ask many other similarly looking questions, for instance — about simultaneous classification of triples of quadratic forms or pairs of ODE systems. Such problems are considered unsolvable: Linear Algebra helps to solve only those problems which can be reduced to one of the previous four or to their slightly more general variants. The catch here is not in the word *general* but in the word *reduced*: each of the above theorems has numerous equivalent reformulations and corollaries (we have seen this in the example of the Orthogonal Diagonalization Theorem on the plane), and one needs quite a bit of experience in order to recognize the questions which can be reduced to them and rule out those where Linear Algebra is helpless.

There is however one more basic theorem (or better to say — formula) in Linear Algebra which has no resemblance with the above classifications. It answers the question *which substitutions of the form*

$$\begin{array}{rcl} x_1 & = & c_{11}X_1 + ... + c_{1n}X_n \\ & ... & \\ x_n & = & c_{n1}X_1 + ... + c_{nn}X_n \end{array}$$

*are indeed changes of the variables and therefore allow to express $X_1, ..., X_n$ linearly via $x_1, ..., x_n$.* It turns out that there exists a remarkable function det of $n^2$ variables $c_{11}, ..., c_{nn}$ which vanishes if and only if the square matrix $C = [c_{ij}]$ is not invertible. We begin our study of higher dimensional linear algebra with properties of matrices and determinants.

*Exercises* 3.1.

(a) Formulate The Rank Theorem in the particular case of two linear functions in two variables. Using the theorem classify linear transformations from the $(x_1, x_2)$-plane to $(y_1, y_2)$-plane up to linear changes of coordinates in both planes. Prove The Rank Theorem in the case $m = n = 2$.

(b) Formulate The Inertia Theorem in the particular case $n = 2$ and compare the statement with results of Chapter 1.

(c) Show that $X_1^2 + ... + X_n^2$ is the only one of the normal forms of The Inertia Theorem which is positive everywhere outside the origin.

(d) Prove that the special case $n = 2$ of The Orthogonal Diagonalization Theorem is equivalent to the Orthogonal Diagonalization Theorem of Chapter 1.

(e) Using the binomial formula show that that the Jordan cell of size $m$ with the eigenvalue $\lambda$ can be written as the $m$-th order ODE

$$y^{(m)} - \binom{m}{1}\lambda y^{(m-1)} + \binom{m}{2}\lambda^2 y^{(m-2)} + ... + (-1)^{m-1}\binom{m}{m-1}\lambda^{m-1}y' + (-1)^m y = 0.$$

(f) Show that $y(t) = e^{\lambda t}(c_0 + tc_1 + ... + c_{m-1}t^{m-1})$ is the general solution to the ODE $(\frac{d}{dt} - \lambda)^m y = 0$.

(g) Specialize the formulation of the Jordan theorem to the case of $n = 2$ linear ODEs $\dot{\mathbf{x}} = A\mathbf{x}$.