Ministry of Education and Science of Ukraine
Oles Honchar Dnipro National University
Faculty of Mechanics and Mathematics
Department of Statistics and Probability Theory
mst-mmf-dnu.dp.ua

S. V. Goncharov

LECTURE NOTES ON

STATISTICAL DECISION THEORY

I

Dnipro
2021

UDC 519.2
G63

G63 Goncharov S.V. Lecture notes on Statistical Decision Theory I [Text] / S.V. Goncharov. — Dnipro: Lira, 2021. — 68 p.

Lecture notes provide basic introduction to Statistical Decision Theory on the following themes: Linear Programming, Elements of Game Theory, and Optimal Statistical Decisions.

Notes are intended for students that study Statistical Decision Theory and anyone who wants to acquaint themselves with the subject.

Recommended for printing
by Scientific Council of Faculty of Mechanics and Mathematics
of Oles Honchar Dnipro National University
protocol #7, 16.02.2021

Навчальне видання

Сергій Валерійович Гончаров

Конспект лекцій
з теорії статистичних рішень I
(англійською мовою)

Друкується за авторською редакцією

Підписано до друку ??.03.2021. Формат 60×84/16. Папір друкарський. Друк плоский. Ум. друк. арк. 3,95. Тираж 35 пр. Зам. № ???.

Друкарня «Ліра», вул. Наукова, 5, м. Дніпро, 49107. Свідоцтво про внесення до Державного реєстру серія ДК №6042 від 26.02.2018 р.

Content

Origin(s)

No speaker has now spoken yet — may one who will speak now speak
and another find what he will speak good!
No one has now spoken yet for a matter spoken afterwards,
as they have done long before. Here is no speaking
what is only planned to be said: this is searching after ruin,
this is falsehood — there is none who will remember his name to others!
I have said these things as I have seen them; from the first generation
until those who come after, they are now like what has passed away.

From "Sayings of Khakheperraseneb", circa 1900 B.C.
`https://www.britishmuseum.org/collection/object/Y_EA5645`

Traduttore, traditore.

Italian proverb

These lecture notes closely follow

М.П. Моклячук, Р.Є. Ямненко: Теорія вибору та прийняття рішень. — ВПЦ «Київський університет», 2013.

up to few examples (numbers in examples have been altered though).

This and other sources are listed in References on p. 68. We advise to read them to get more thorought and complete understanding of the subject. Meanwhile...

Introduction

Statistical Decision Theory (SDT) considers questions and problems that, generally, can be described as: in certain situation, about which some information is known, having certain restrictions and following certain rules, — what is the best/optimal way to act and why?

Of course, this definition is vague, — we can apply it to almost everything (for example, to our lives). To make it more exact, we must define what is a "rule", what changes an "action" causes, what "optimal" means? for whom? etc.

We consider "situations" where, to our actions, the outcome corresponds, which usually is of a numerical kind: it is a scalar — our "winnings" / "profit" / "gain" (in this case we want to make it as large as we can) or our "loss" / "cost" (this number must be minimized).

Themes that SDT aggregates also gather together under the titles "Decision Theory", "Operations Research", and their variations. We have "statistical" in the title because the processes that occur in considered situations are often stochastic, non-determined, and so is the outcome of our actions.

You can unfold "decision" as "to decide how to act optimally".

For instance, when we play (with someone) some game, and for us it has only 2 possible outcomes — we win or we lose — we act in such a way as to maximize the *probability* of winning. Or, when the payoff is a random variable (RV) ξ, whose distribution depends on our actions, we act in such a way as to maximize the mean payoff, $\mathrm{E}\,\xi$.

Let us consider an example of such game, decide how to play it the best way, and make some conclusions.

The game field consists of 3 cells, and the player A — imagine it is you — begins in the central one: | | A | |. The adversary, player B, resides nearly and has a "gun". The game consists of consecutive moves, at each move simultaneously

1) player A moves to the adjacent cell or stays in her/his current cell,

2) player B shoots into any cell.

Edge cells are exits: as soon as A reaches such cell and B misses, A wins (and B loses). If B hits A, then B wins (and A loses).

How to play this game? — what decision to make at each move?

In other words, we are interested in the strategy for player A (for now we do not try to make the notion of "strategy" more precise, its conventional meaning satisfies our needs). Additionally, we suppose that player B knows this strategy, — things become harder for A, — up to the probabilities of possible actions of A at each move.

So, both players can use *stochastic* strategies, when the state after each move determines not the next move, but the (probability) *distribution* of the next move.

The word “move” may seem inappropriate, because on the face of it, clearly, A must move to one of two exits right away and win or lose, that is, there is only 1 “move”. This edge cell-exit must be determined randomly, otherwise B will know for sure where to shoot (e.g. if your strategy as A is to go right, then B shoots there and is guaranteed to win). It is easy to see that the best thing to do is to choose each exit-cell with probability $\frac{1}{2}$, then B shoots into this cell with probability $\frac{1}{2}$ as well (to substantiate it, we “switch to B’s side”, — in her/his place, if you know that A goes left or right with probability $\frac{1}{2}$, it is the best for you to choose target by the same distribution), and by the Total Probability Theorem (TPT) A wins with probability $\frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{2}$.

Consider other strategy of A, e.g. the one by which she/he commits a “folly”: at each move remains motionless with non-zero probability. To be more precise, the distribution of the next A position is $\begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$: all positions are equiprobable.

What if B replies on this strategy of A by the same distribution of the cell where she/he shoots? Now the game can continue in principle as long as one wants (A stands still, B shoots into edge cells). What is the probability of A winning? To stand still when someone shoots at you, — a wrong decision, so it must decrease...

At each, i-th, move A wins (W_i), loses (L_i) or the game continues (C_i).

By TPT

$$P(W_1)=2\cdot\frac{1}{3}\cdot\frac{2}{3}=\frac{4}{9},\ P(L_1)=3\cdot\frac{1}{3}\cdot\frac{1}{3}=\frac{1}{3},\ P(C_1)=1-P(W_1)-P(L_1)=\frac{2}{9}$$

The 2nd move will occur only if C_1 happened, and the situation is repeated “in the world whose probability is $\frac{2}{9}$”. Thus

$$P(W_2)=P(W_2\cap C_1)=\frac{2}{9}\cdot\frac{4}{9},\ P(L_2)=\frac{2}{9}\cdot\frac{1}{3},\ P(C_2)=\left(\frac{2}{9}\right)^2$$

etc. The probability of A winning is $P(W)=\sum_{i=1}^{\infty}P(W_i)=\frac{4}{9}\sum_{i=0}^{\infty}\left(\frac{2}{9}\right)^i=\frac{4}{9}\cdot\frac{1}{1-\frac{2}{9}}=\frac{4}{7}>\frac{1}{2}$ by the 1st strategy.

Now, what if A’s strategy is, as before, $\begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$, but B “pays attention” only to the cells that provide exit to A and shoots by the distribution $\begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$? Then we obtain $P(W)=\frac{1}{2}$ again. But wait, isn’t it now safer for A to stay in cell 2, at least until B changes her/his strategy too?

Consider simpler version with left exit-cell removed: if A goes right with probability 1, B is guaranteed to win by shooting into the same cell, so it is better for A to stand still/go right with probability $\frac{1}{2}$.

One conclusion from this example is doubts about our, “formed by life experience”, ability to make “optimal” decisions: “foolish” strategy may appear to be no less reliable than the “obvious” one.

Another conclusion: we need general approaches to such (and other) problems, which will assist in solving their more complex variations (let the cells form the square 7×7, edge cells are exits, at the beginning A is at the central cell), when, if using the "straightforward" approach, like the one described above with TPT, we quickly become "confused".

What other conclusions can you make from this example?

Is there even better strategy for A? What about B?

Is there at all the best strategy for A, by which the probability of her/his winning is the largest? Maybe, whatever strategy you choose, there is always only a little, but better one?..

With "game-like" language we can describe the problems of Statistics as well. For instance, if we suggest the estimate $\hat{\theta}$ of the unknown parameter θ of some distribution, and then we pay the "penalty" $L = |\hat{\theta} - \theta|$, — absolute error, — our goal is to minimize L, that is, our *loss* (the best loss is 0). Since L is RV, the "minimization" in turn means $\mathrm{E}\, L \to \min$ or $\mathrm{E}\, L^2 \to \min$ etc.

Traditionally, such game is called "Statistician vs Nature": Nature "knows" the actual value of a parameter (say, the mean a of the distribution $N_{a;\sigma^2}$ of the mass of a leaf on the tree), and Statistician tries to "guess" it (measuring the masses of individual leaves and suggesting the estimate $\hat{a}$) as precisely as she/he can, that is, to minimize, in average, the loss from imprecise calculation of the parameter.

Terminology. Considering the situations in which the decisions are made, we will construct the *mathematical models* of these situations; one and the same situation can be described by several models. Such models are defined by *variables*. By means of variables we express the *constraints* of the model (for example, as inequalities) and the *target function*, which we want to maximize or, on the contrary, to minimize, satisfying these constraints.

The set of variables makes the *solution* of the problem. The solution that satisfies all constraints is called an *admissible* one, and if, at that, it maximizes/minimizes (depending on what we need) the target function, it is called an *optimal* one. Also, there are *locally optimal* solutions: such solution is the local extremum of the target function, which can be different from the global one.

The considerable part of the methods to find the optimal solution consists of iterative methods, when we obtain the consecutive approximations of the optimal solution. At that, usually, it is not represented in closed form, as an explicit function of the problem's parameters. Naturally, these methods are oriented towards use of "computation devices", or computers. If the model (and the problem) is so complex that the optimal solution cannot be found even by these methods, sometimes we can use the "heuristic" approach and find not the optimal, but the "suitable" solution.

1. Linear programming

The term *linear programming* (LP) was introduced by G.B. Dantzig. In 1947 he developed the simplex method to solve such problems. It is considered that for the first time the LP problem was formulated in 1939 in L.V. Kantorovich's paper "Mathematical Methods of Organizing and Planning Production", where one method to solve it was suggested. I.I. Dikin, N. Karmarkar, T.C. Koopmans, H.W. Kuhn, J. von Neumann, A.W. Tucker developed this area of researches.

Consider few "economic" problems where similar mathematical models appear:

Resource distribution problem. We have $b_i \geqslant 0$ units of *resource* R_i, $i = \overline{1, m}$ (e.g. 100 tons of iron ore, 20 tons of carbon, 10000 m^3 of water; units of measurement may be different).

Using these (and other) resources, we manufacture the *productions* P_j, $j = \overline{1, n}$ (e.g. steel, cast iron).

To manufacture one unit of the production P_j we spend $a_{ij} \geqslant 0$ units of resource R_i (e.g. 1 ton of steel requires 900 kg of ore, 50 kg of carbon etc.)

1 unit of production P_j has the "cost" $c_j \geqslant 0$ (e.g. 1 ton of steel is sold for 3000 "coins").

We manufacture $x_j \geqslant 0$ units of P_j, $j = \overline{1, n}$. To do that, we need $\sum\limits_{j=1}^{n} a_{ij} x_j$ units of resource R_i, such amount cannot exceed the total number of units of this resource, i.e. $\sum\limits_{j=1}^{n} a_{ij} x_j \leqslant b_i$, $i = \overline{1, m}$.

Our goal is to manufacture so many units of each production from limited amount of resources, as to get the largest profit: $\sum\limits_{j=1}^{n} c_j x_j \to \max$.

Transport problem. We have the single *product* that is manufactured at the points P_i, $i = \overline{1, m}$ and consumed at the points Q_j, $j = \overline{1, n}$: during certain time the point P_i manufactures a_i units of the product, and the point Q_j consumes b_j units, the *balance condition* $\sum\limits_{i=1}^{m} a_i = \sum\limits_{j=1}^{n} b_j$ holding (there is neither surplus, nor shortage of the product). The cost of the transportation of 1 unit of the product from P_i to Q_j is c_{ij}.

We must make up the plan of transportation $\{x_{ij}\}_{1 \leqslant i \leqslant m,\, 1 \leqslant j \leqslant n}$, where $x_{ij} \geqslant 0$ is the number of units of the product that is transported from P_i to Q_j. At that the needs of manufacturers and consumers must be satisfied completely, that is, $\sum\limits_{j=1}^{n} x_{ij} = a_i$ and $\sum\limits_{i=1}^{m} x_{ij} = b_j$.

Our goal is to minimize the total cost of transportation: $\sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} c_{ij} x_{ij} \to \min$.

Food ration problem. The ration consists of the products P_j, $j=\overline{1,n}$ (e.g. bread, canned food).

1 unit of the product P_j contains a_{ij} units of the "nutritious substance" S_i, $i=\overline{1,m}$ (e.g. proteins, fats, carbohydrates) and has the "cost" c_j (e.g. 250 grams of canned food costs 20 "coins").

The whole ration must contain not less than b_i units of the substance S_i (e.g. not less than 50 grams of fats).

Thus, if the ration includes x_j units of the product P_j, $j=\overline{1,n}$, we need $\sum\limits_{j=1}^{m} a_{ij}x_j \geqslant b_i$, $i=\overline{1,m}$.

Our goal is to minimize the expenses on the ration: $\sum\limits_{j=1}^{m} c_j x_j \to \min$.

In these problems the left hand side of the constraints and the target function are *linear (forms)*. We transform $\geqslant$-inequalities into $\leqslant$-ones multiplying them by -1 (in particular, $x_j \geqslant 0 \to -x_j \leqslant 0$) and write

LP problem (LPP) in general form. Find the point $x=(x_1;...;x_n)\in\mathbb{R}^n$ that satisfies the constraints

$$\sum_{j=1}^{n} a_{ij}x_j \leqslant b_i,\ i=\overline{1,m}$$

$$\sum_{j=1}^{n} p_{ij}x_j = q_i,\ i=\overline{1,r}$$

and maximizes/minimizes the target function

$$f(x)=\langle c;x\rangle=\sum_{j=1}^{n} c_j x_j$$

DEF. The set of points satisfying all constraints is called *admissible (domain)*.

DEF. If at least 1 admissible point exists, the constraint set is called *consistent*, otherwise *inconsistent*.

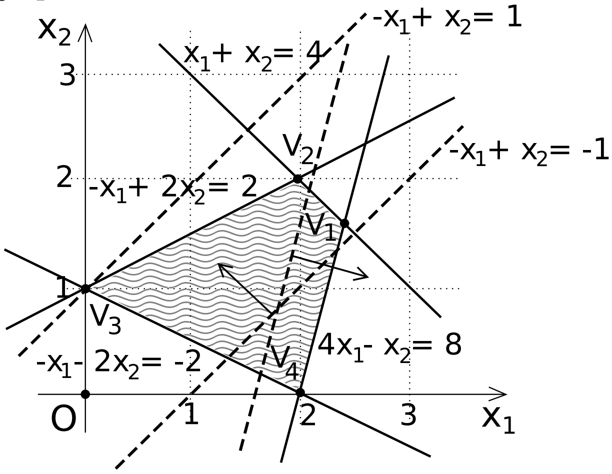
The substitution of $-c=(-c_1;...;-c_n)$ for $c=(c_1;...;c_n)$ turns the "$\to\max$" problem to "$\to\min$" problem and vice versa.

We can solve LP problems with 2 variables by *graphic* method, using the geometric interpretation. Let us consider two such problems, they illustrate the properties of admissible domains (AD), target functions (TF), solutions etc. that take place in general case too.

EX. 1. $f(x)=-x_1+x_2\to\max$ under constraints

$$\begin{cases} -x_1 + 2x_2 \leqslant 2, \\ -x_1 - 2x_2 \leqslant -2, \\ 4x_1 - x_2 \leqslant 8, \\ x_1 + x_2 \leqslant 4, \\ x_1 \geqslant 0, \ x_2 \geqslant 0. \end{cases}$$

On the Cartesian plane $(x_1; x_2)$ the equation $Ax_1 + Bx_2 + C = 0$ determines straight line, and corresponding inequalities determine semiplanes this line divides the plane into. Hence AD is the intersection of some semiplanes, in this case AD is the polygon $V_1V_2V_3V_4$. Note that it is convex.

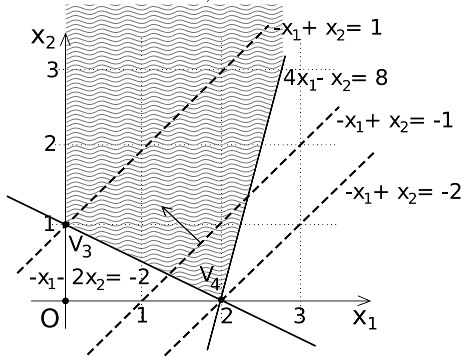


Analogously, the set of points in which $f(x) = C$ is the straight line $-x_1 + x_2 - C = 0$. When C increases, this line moves in the direction $(-1; 1)$. Admissible solutions x such that $f(x) = C$ exist **iff** this line intersects with given polygon. The optimal solution, that is, the point of polygon in which $f(x)$ attains max, is $V_3 = (0; 1)$, at that $f(V_3) = 1$.

Let TF $g(x) = 8x_1 - 2x_2$. The optimal solutions make the segment V_1V_4 — edge of the polygon; at any p. of this edge, in particular at ends-vertices, $g(x) = 16$.

<u>EX. 2.</u> $f(x) = -x_1 + x_2 \to \max$ under constraints $\begin{cases} -x_1 - 2x_2 \leqslant -2, \\ 4x_1 - x_2 \leqslant 8, \\ x_1 \geqslant 0, \ x_2 \geqslant 0 \end{cases}$

(Ex. 1 without 1st and 4th constraints).



This problem does not have a solution, because the straight line $-x_1+x_2=C$ intersects the unbounded AD for any, as large as one wants, C. Meanwhile the problem $f(x)\to\min$ has unique solution V_4.

If there are 3 variables, the geometric interpretation leads to 3D polyhedrons, perhaps unbounded; TF = const on planes. LP problem in $\mathbb{R}^3$ can be solved graphically as well, — moving such plane until it intersects the "edge" points of a polyhedron, — but the "clearness" deteriorates; moreover it applies to the "real" problems, where there can be dozens, hundreds (, millions, ...) of variables.

Therefore we need general, non-graphic methods, "programmable" ones at that, and they should work — find the solution — fast enough. These general methods rely on general properties of AD, TF on them etc.

Taking into account that the AD in a LP problem is the convex set of a certain kind, —

1.1. Convex sets and elements of convex analysis

DEF. Let $x=(x_1;...;x_n)$, $y=(y_1;...;y_n)\in\mathbb{R}^n$. The *segment* with the ends at these points is

$$[x;y]=\big\{\lambda x+(1-\lambda)y \mid \lambda\in[0;1]\big\}$$

where $\lambda x+(1-\lambda)y=\big(\lambda x_1+(1-\lambda)y_1;...;\lambda x_n+(1-\lambda)y_n\big)$.

DEF. $A\subseteq\mathbb{R}^n$ is called *convex* if $\forall x,y\in A$: $[x;y]\subseteq A$.

Name and draw examples of convex sets in $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^n$.

PROP. 1. Let $a\in\mathbb{R}^n$, $\beta\in\mathbb{R}$. Semispace $H_-=\big\{x\in\mathbb{R}^n \mid \langle a;x\rangle=\sum\limits_{i=1}^n a_ix_i\leqslant\beta\big\}$ is convex.

◀ $\forall x,y\in H_-$, $\forall\lambda\in[0;1]$: for $z=\lambda x+(1-\lambda)y$

$$\langle a;z\rangle=\lambda\langle a;x\rangle+(1-\lambda)\langle a;y\rangle\leqslant\lambda\beta+(1-\lambda)\beta=\beta\Rightarrow z\in H_- \text{ ▶}$$

Similarly, semispace $H_+=\big\{x \mid \langle a;x\rangle\geqslant\beta\big\}$ and (hyper)plane $H=\big\{x \mid \langle a;x\rangle=\beta\big\}$ are convex.

PROP. 2. If $\{A_i\}_{i\in I}$, $A_i\subseteq\mathbb{R}^n$ are convex, then $A=\bigcap\limits_{i\in I}A_i$ is convex.

◀ $\forall x,y\in A$, by def. of $\cap$, $\forall i\in I$: $x,y\in A_i$. Since A_i is convex, we have $[x;y]\subseteq A_i\Rightarrow[x;y]\subseteq A$. ▶

DEF. The intersection of a finite number of semispaces is called a *polyhedral set*. If this set is bounded, it is called a *polyhedron*.

COR. A polyhedral set is convex.

PROP. 3. Let $\{A_i\}_{i=1}^k$, $A_i \subseteq \mathbb{R}^n$ be convex and $a_1, ..., a_k \in \mathbb{R}$. Then the *linear combination* of $\{A_i\}$ is convex:

$$A = \sum_{i=1}^k a_i A_i = \{\sum_{i=1}^k a_i x^{(i)} \mid x^{(i)} \in A_i, i = \overline{1, k}\}$$

◀ $\forall x, y \in A$: $x = \sum_{i=1}^k a_i x^{(i)}$, $y = \sum_{i=1}^k a_i y^{(i)}$, where $x^{(i)}, y^{(i)} \in A_i$. $\forall \lambda \in [0; 1]$:

$$\lambda x + (1-\lambda)y = \sum_{i=1}^k \left[\lambda a_i x^{(i)} + (1-\lambda) a_i y^{(i)}\right] = \sum_{i=1}^k a_i \underbrace{\left[\lambda x^{(i)} + (1-\lambda) y^{(i)}\right]}_{\in A_i} \in A$$

in other words, $[x; y] \subseteq A$. ▶

DEF. Let $x^{(1)}, ..., x^{(k)} \in \mathbb{R}^n$. A *convex combination* of these points is

$$\sum_{i=1}^k \lambda_i x^{(i)}, \text{ where } \lambda_i \geqslant 0 \text{ and } \sum_{i=1}^k \lambda_i = 1$$

PROP. 4. A convex set contains all convex combinations of its points.

◀ We denote this set by A, and let $x^{(1)}, ..., x^{(k)} \in A$. We claim that $\forall \{\lambda_i\}_{i=1}^k$, $\lambda_i \geqslant 0$, $\sum_{i=1}^k \lambda_i = 1$, we have $x = \sum_{i=1}^k \lambda_i x^{(i)} \in A$. Rewrite

$$x = \lambda_k x^{(k)} + (1-\lambda_k) \sum_{i=1}^{k-1} \underbrace{\frac{\lambda_i}{1-\lambda_k}}_{\lambda_{1;i}} x^{(i)} =$$

$$= \lambda_x x^{(k)} + (1-\lambda_k)\left[\lambda_{1;k-1} x^{(k-1)} + (1-\lambda_{1;k-1}) \sum_{i=1}^{k-2} \underbrace{\frac{\lambda_{1;i}}{1-\lambda_{1;k-1}}}_{\lambda_{2;i}} x^{(i)}\right] =$$

$$= \ldots = \lambda_k x^{(k)} + (1-\lambda_k)\Big[\lambda_{1;k-1} x^{(k-1)} + (1-\lambda_{1;k-1})\big[\lambda_{2;k-2} x^{(k-2)} + \ldots$$
$$\ldots \{\lambda_{k-2;2} x^{(2)} + (1-\lambda_{k-2;2}) x^{(1)}\} ...\big]\Big]$$

(we remark that $\sum_{i=1}^{k-m-1} \lambda_{m+1;i} = \sum_{i=1}^{k-m-1} \frac{\lambda_{m;i}}{1-\lambda_{m;k-m}} = 1$). In the deepest brackets $\{...\}$, — conv.comb. of 2 points from A, $x^{(2)}$ and $x^{(1)}$, which, by def. of convexity, $\in A$. Thus analogously $\lambda_{k-3;3} x^{(3)} + (1-\lambda_{k-3;3})\{...\} \in A$ and so forth ("surface from depth"); finally we obtain $x \in A$.

If $1 - \lambda_{m;k-m} = 0$, stop "sink" at this m and start to "surface". ▶

DEF. Let $A \subseteq \mathbb{R}^n$. The *convex hull* of A is

$$\operatorname{conv} A = \bigcap_{\mathbb{R}^n \supseteq B \supseteq A,\ B \text{ is convex}} B$$

Draw examples of convex hulls in $\mathbb{R}^1$, $\mathbb{R}^2$.

PROP. 5. Elementary properties of conv:
- $\operatorname{conv} A$ is convex;
- $\operatorname{conv} A$ is the smallest, by inclusion, convex set that contains A;
- A is convex **iff** $A = \operatorname{conv} A$;
- $A \subseteq B \Rightarrow \operatorname{conv} A \subseteq \operatorname{conv} B$.

PROP. 6. $\operatorname{conv} A$ coincides with the set of all conv.comb. of A points.

◀ We denote the set of all such combinations by C. $\forall x, y \in C$

$$x = \sum_{i=1}^{k} \alpha_i x^{(i)},\ y = \sum_{i=1}^{m} \beta_i y^{(i)},\ \text{where } \alpha_i, \beta_i \geqslant 0,\ \sum_{i=1}^{k} \alpha_i = 1,\ \sum_{i=1}^{m} \beta_i = 1;\ x^{(i)}, y^{(i)} \in A$$

Then $\forall \lambda \in [0;1]$ the point $\lambda x + (1-\lambda)y$ is the convex combination of A points, because $\lambda\alpha_i, (1-\lambda)\beta_i \geqslant 0$ and $\sum\limits_{i=1}^{k} \lambda\alpha_i + \sum\limits_{i=1}^{m}(1-\lambda)\beta_i = \lambda + 1 - \lambda = 1$. In other words, $[x;y] \subseteq C$, thus C is convex.

Obviously, $A \subseteq C$, so $\operatorname{conv} A \subseteq C$. On the other hand, take any convex $B \supseteq A$. $\forall x \in C$ is the conv.comb. of A points, and therefore of B points; by Prop. 4, $x \in B$. Hence $C \subseteq B$, and it follows from arbitrariness of B that $C \subseteq \operatorname{conv} A$. Thus $\operatorname{conv} A = C$. ▶

In the representation of $\operatorname{conv} A$ point by a conv.comb. of some A points (always possible due to this Prop., but not unique in general case) we naturally strive to decrease the number of "combined" points, to simplify the "combination". The following important Theor. of finite-dimensional convex analysis provides the upper bound on necessary number of points: this number depends only on dimensionality of the space.

THEOR. (Caratheodory) Let $A \subseteq \mathbb{R}^n$. Any $\operatorname{conv} A$ point is the conv.comb. of no more than $n+1$ points of A:

$$x \in \operatorname{conv} A \Rightarrow x = \sum_{i=1}^{k} \lambda_i x^{(i)},\ \text{where } x^{(i)} \in A,\ \lambda_i \geqslant 0,\ \sum_{i=1}^{k} \lambda_i = 1,\ \underline{\text{and } k \leqslant n+1}$$

◀ $\forall x \in \operatorname{conv} A$ by Prop. 6 can be represented as $x = \sum\limits_{i=1}^{k} \lambda_i x^{(i)}$, where $x^{(i)} \in A$, $\underline{\lambda_i > 0}$, $\sum\limits_{i=1}^{k} \lambda_i = 1$. If $k \leqslant n+1$, Theor. is proved. Suppose that $k > n+1$. We are going to show that this number can be decreased by at least 1, without adding other points to $\{x^{(i)}\}_{i=1}^{k}$.

To do it, consider the vectors $y^{(i)} = (x_1^{(i)};...;x_n^{(i)};1) \in \mathbb{R}^{n+1}$, $i = \overline{1,k}$. Since $k > n+1$, this collection of vectors is lin.dep., and therefore $\exists\{\alpha_i\}_{i=1}^k$, $\alpha_i \in \mathbb{R}$: $\sum\limits_{i=1}^k \alpha_i^2 \neq 0$, $\sum\limits_{i=1}^k \alpha_i y^{(i)} = \theta_{n+1}$ (here $\theta_m = (0;0;...;0)$ is the zero of LVS $\mathbb{R}^m$). All coord. except for $(n+1)$-th give $\sum\limits_{i=1}^k \alpha_i x^{(i)} = \theta_n$, while for that one $\sum\limits_{i=1}^k \alpha_i = 0$, and since not all $\alpha_i = 0$, there are positive ones among them. Let

$$\varepsilon = \min\left\{\tfrac{\lambda_i}{\alpha_i} \mid \alpha_i > 0\right\}$$

and let min be attained for $i = i_0$. Consider

$$\lambda_i' = \lambda_i - \varepsilon\alpha_i,\ i = \overline{1,k}$$

If $\alpha_i \leqslant 0$, $\lambda_i' \geqslant \lambda_i > 0$, if $\alpha_i > 0$, $\lambda_i' = \alpha_i(\frac{\lambda_i}{\alpha_i} - \varepsilon) \geqslant 0$ by definition of ε. Particularly, $\lambda_{i_0}' = \lambda_{i_0} - \lambda_{i_0} = 0$.

$$\sum_{i=1}^k \lambda_i' = \sum_{i=1}^k \alpha_i - \varepsilon \sum_{i=1}^k \alpha_i = 1 - 0 = 1$$

$$\sum_{i=1}^k \lambda_i' x^{(i)} = \sum_{i=1}^k \lambda_i x^{(i)} - \varepsilon \sum_{i=1}^k \alpha_i x^{(i)} = x - \theta_n = x$$

and we can remove from this sum (at least) one, i_0-th member, because $\lambda_{i_0}' = 0$.

Thus decrease the number of combined points until it becomes $k \leqslant n+1$. ►

DEF. Let $x \in \mathbb{R}^n$, $A \subseteq \mathbb{R}^n$. The *projection of x into A*, $\pi_A(x)$, is the nearest to x point of A:

$$\pi_A(x) \in A,\ \rho(x;\pi_A(x)) = \inf_{u\in A} \rho(x;u)$$

(here $\rho(x;y) = \|x-y\| = \left(\sum\limits_{i=1}^n (x_i-y_i)^2\right)^{\frac{1}{2}}$ is usual Euclidean distance). Obviously, $x \in A \Rightarrow \pi_A(x) = x$.

$\pi_A(x)$ is also called the *best approximant of x in A*. In general case, $\pi_A(x)$ may not exist or be non-unique.

PROP. 7. Let closed and convex $A \subset \mathbb{R}^n$ and $x \notin A$.

Then $\exists!\pi_A(x)$ and $\forall u \in A$:

$$\langle \pi_A(x) - x; u - \pi_A(x)\rangle \geqslant 0 \qquad (*)$$

$$\langle \pi_A(x) - x; u - x\rangle \geqslant \|\pi_A(x) - x\|^2 > 0 \qquad (**)$$

◄ We take $\forall y \in A$ and let $D := \{u \in A\colon \|x-u\| \leqslant \|x-y\|\}$ ($\ni y$). Due to A being closed and the continuity of metric/norm this subset of $\mathbb{R}^n$ is closed and bounded, therefore it is compact. Hence the continuous function $f(u) = \|u-x\|$ on D is bounded from below and attains its inf, which is equal to d, at some point $u_0 \in A$; since it attains inf on the whole A at the same point, u_0 is the projection of x into A.

We remark that $d = \|x - u_0\| > 0$, otherwise $x \in A$, which contradicts the conditions.

Suppose that such projection is not unique, $\exists v_0 \in A$: $v_0 \neq u_0$, $\|x - v_0\| = \|x - u_0\|$. Then $w = \frac{1}{2}(u_0 + v_0)$, on the one hand, $\in A$, because A is convex; on the other hand, due to strict convexity of Euclidean norm,

$$\|w - x\| = \|\tfrac{1}{2}(u_0 - x) + \tfrac{1}{2}(v_0 - x)\| < d$$

— contradiction with $f(u)$ on A not taking values smaller than d.

Thus $\exists! \pi_A(x) = u_0$. $\forall u \in A$ and $\forall \lambda \in (0; 1]$ convexity of A implies $u' = \lambda u + (1 - \lambda) u_0 \in A$. Therefore $d^2 = \|u_0 - x\|^2 \leqslant \|u' - x\|^2 =$

$$= \|u_0 - x + \lambda(u - u_0)\|^2 = \|u_0 - x\|^2 + \lambda^2 \|u - u_0\|^2 + 2\lambda \langle u_0 - x; u - u_0 \rangle$$

which implies $\langle u_0 - x; u - u_0 \rangle \geqslant -\frac{1}{2}\lambda \|u - u_0\|^2$. Let $\lambda \to 0$ to obtain $(*)$.

In $(*)$, we add and subtract x to the 2nd multiplier of $\langle \cdot; \cdot \rangle$, and get $\langle \pi_A(x) - x; u - x \rangle - \langle \pi_A(x) - x; \pi_A(x) - x \rangle \geqslant 0 \Leftrightarrow \langle \pi_A(x) - x; u - x \rangle \geqslant d^2 > 0$ — $(**)$ ►

$(*)$ and $(**)$ mean, accordingly, that the angle between $\pi_A(x) - x$ and $u - \pi_A(x)$ is always non-obtuse, and the angle between $\pi_A(x) - x$ and $u - x$ is acute.

DEF. We say that $A, B \subseteq \mathbb{R}^n$ are

- *separated*, if $\exists p \in \mathbb{R}^n$, $p \neq \theta$, and $\exists \alpha \in \mathbb{R}$ such that

$$\forall x \in A,\ \forall y \in B\colon \langle p; x \rangle \geqslant \alpha \geqslant \langle p; y \rangle$$

- *properly separated*, if they are separated and for the corresponding p

$$\exists x' \in A,\ \exists y' \in B\colon \langle p; x' \rangle > \langle p; y' \rangle$$

- *strongly separated*, if $\exists p \in \mathbb{R}^n$, $p \neq \theta$, and $\exists \alpha \in \mathbb{R}$ such that

$$\inf_{x \in A} \langle p; x \rangle > \alpha > \sup_{y \in B} \langle p; y \rangle$$

Vector p and scalar α define, in $\mathbb{R}^n$, the (hyper)plane $H = \{x\colon \langle p; x \rangle = \alpha\}$, which divides this space into 2 semispaces: $H_+ = \{x\colon \langle p; x \rangle \geqslant \alpha\}$ and $H_- = \{x\colon \langle p; x \rangle \leqslant \alpha\}$. We say that H *separates* the sets A and B, which get into different semispaces; H is called *separating* at that.

Proper separation means that A and B do not lie entirely in the separating plane (otherwise it would be rather “uniting”), and strong separation means that they reside at positive distances from it in “their” semispaces; clearly, when the sets are strongly separated, they are separated properly as well.

THEOR. (Minkowski, on strong separation of point and set) Let closed and convex $A \subseteq \mathbb{R}^n$ and $x \notin A$. Then $\exists p \in \mathbb{R}^n$, $p \neq \theta$, and $\exists \alpha \in \mathbb{R}$ such that

$$\inf_{u \in A} \langle p; u \rangle > \alpha > \langle p; x \rangle$$

◄ By Prop. 7, $\exists! \pi_A(x) \in A$, at that $d = \|x - \pi_A(x)\| > 0$. Let $p = \pi_A(x) - x$. $\forall u \in A$ by $(**)$

$$\langle p; u - x \rangle = \langle p; u \rangle - \langle p; x \rangle \geqslant d^2$$

Thus, for $\alpha_0 = \langle p; x \rangle$, we have $\langle p; u \rangle \geqslant d^2 + \alpha_0$. What is left is to take arbitrary α from $(\alpha_0; \alpha_0 + d^2)$ — for the sake of definiteness, let $\alpha = \alpha_0 + \frac{1}{2}d^2$ — then

$$\inf_{u \in A} \langle p; u \rangle > \alpha > \langle p; x \rangle$$ ►

Recall some topological concepts and their connection with convexity in $\mathbb{R}^n$.

DEF. The *closure* of $A \subseteq \mathbb{R}^n$ is

$$\overline{A} = \{x \in \mathbb{R}^n \mid \exists \{x^{(k)}\}_{k=1}^{\infty} \subseteq A \colon x^{(k)} \xrightarrow[k\to\infty]{} x\} = \\ = \{x \in \mathbb{R}^n \mid \forall \delta > 0 \ \exists y \in A \colon \|y - x\| < \delta\}$$

(Thus here $\overline{A}$ means closure, not complement.)

DEF. Point $x \in A$ is called an *interior* point of the set $A \subseteq \mathbb{R}^n$ if x is contained in A with some δ-neighbourhood ($\delta > 0$): $\rho(y; x) < \delta \Rightarrow y \in A$.

DEF. The *interior* $\operatorname{Int} A$ of the set $A \subseteq \mathbb{R}^n$ is the set of its interior points.

DEF. The *boundary* of the set $A \subseteq \mathbb{R}^n$, denoted by ∂A, is $\overline{A} \setminus \operatorname{Int} A$.
The boundary of the segment I in $\mathbb{R}^1$ is its ends; in $\mathbb{R}^n$, $n \geqslant 2$ — the entire I.

PROP. 8. If A is convex, then $\overline{A}$ is convex too.
◀ $\forall x, y \in \overline{A} \ \exists \{x^{(k)}\}_{k=1}^{\infty} \subseteq A, \{y^{(k)}\}_{k=1}^{\infty} \subseteq A \colon x^{(k)} \xrightarrow[k\to\infty]{} x$, $y^{(k)} \xrightarrow[k\to\infty]{} y$. So $\forall \lambda \in [0;1] \colon A \ni \lambda x^{(k)} + (1-\lambda) y^{(k)} \xrightarrow[k\to\infty]{} \lambda x + (1-\lambda) y \Rightarrow \lambda x + (1-\lambda) y \in \overline{A}$ ▶

....

LEMMA 1. Let $\{v^{(i)} = (s_1; ...; s_n) \mid s_j = \pm 1, j = \overline{1,n}\}$ be the vertices of the "symmetric unit" cube in $\mathbb{R}^n$, $i = \overline{1, 2^n}$. $\forall \delta < 1$ let the points $\{x^{(i)}\}_{i=1}^{2^n} = M$ be such that $\|x^{(i)} - v^{(i)}\| < \delta$. Then $\theta \in \operatorname{conv} M$.

◀ Since Prop. 6, it suffices to show that θ is the conv.comb. of M points.

$|x_j^{(i)} - v_j^{(i)}| \leqslant \big(\sum\limits_{j=1}^{n} |x_j^{(i)} - v_j^{(i)}|^2\big)^{\frac{1}{2}} < \delta < 1$ implies $\operatorname{sign} x_j^{(i)} = \operatorname{sign} v_j^{(i)}$ $(= \pm 1)$.

We apply induction by $n = \dim \mathbb{R}^n$.

Induction base, $n = 1$: we have 2 p. $x^{(1)} > 0$ and $x^{(2)} < 0$. For $\alpha_1 = |x^{(2)}|$ and $\alpha_2 = |x^{(1)}|$: $\alpha_1 x^{(1)} + \alpha_2 x^{(2)} = 0$, thus, taking $\lambda_1 = \frac{\alpha_1}{\alpha_1 + \alpha_2}$ and $\lambda_2 = \frac{\alpha_2}{\alpha_1 + \alpha) 2}$, we have $\lambda_1 x^{(1)} + \lambda_2 x^{(2)} = 0$ and $\lambda_i \geqslant 0$, $\lambda_1 + \lambda_2 = 1$.

Ind. supposition: Prop. holds in $\mathbb{R}^n$.

Ind. step: consider 2^{n+1} points $x^{(i)}$ in $\mathbb{R}^{n+1}$. Divide them into 2 collections, 2^n points each, reindexing: $A_- = \{x^{(-i)}\}$ such that $x_{n+1}^{(-i)} < 0$, and $A_+ = \{x^{(+i)}\}$ such that $x_{n+1}^{(+i)} > 0$, — "under" and "above" the plane $x_{n+1} = 0$. Project these p. into this plane and obtain p. $h^{(\pm i)} \in \mathbb{R}^n$, $h_j^{(\pm i)} = x_j^{(\pm i)}$ for $j = \overline{1,n}$. It is easy to see that both collections of projections — p. of A_- and p. of A_+ — satisfy the conditions of ind. supp., because the distance between each of them and the corresponding vertex of symm. unit cube, at the projection from $\mathbb{R}^{n+1}$ into $\mathbb{R}^n$, has not increased (it suffices that the signs of coordinates have not changed). Therefore

$$\exists\{\gamma_i^-\}_{i=1}^{2^n},\ \gamma_i^-\geqslant 0,\ \sum_{i=1}^{2^n}\gamma_i^-=1\colon \sum_{i=1}^{2^n}\gamma_i^- h^{(-i)}=\theta_n \Leftrightarrow \sum_{i=1}^{2^n}\gamma_i^- x_j^{(-i)}=0,\ j=\overline{1,n}$$
$$\exists\{\gamma_i^+\}_{i=1}^{2^n},\ \gamma_i^+\geqslant 0,\ \sum_{i=1}^{2^n}\gamma_i^+=1\colon \sum_{i=1}^{2^n}\gamma_i^+ h^{(+i)}=\theta_n \Leftrightarrow \sum_{i=1}^{2^n}\gamma_i^+ x_j^{(+i)}=0,\ j=\overline{1,n}$$
— in other words, p. $\sum\limits_i \gamma_i^- x^{(-i)}$ and $\sum\limits_i \gamma_i^+ x^{(+i)}$ belong to the axis Ox_{n+1} in $\mathbb{R}^{n+1}$. Let $s_\pm = \sum\limits_i \gamma_i^\pm x_{n+1}^{(\pm i)}$ be the coord. of these 2 p. on this axis, then $s_- < 0$, $s_+ > 0$. Analogously to the reasoning in the base, let $\mu_\pm = \frac{|s_\mp|}{|s_-|+|s_+|}$, then $\mu_\pm \geqslant 0$, $\mu_- + \mu_+ = 1$ and $\mu_- s_- + \mu_+ s_+ = 0$, and so, denoting $\lambda_i^\pm = \mu_\pm \gamma_i^\pm$, we have
$$\sum_i \lambda_i^- x^{(-i)} + \sum_i \lambda_i^+ x^{(+i)} = \mu_- \sum_i \gamma_i^- x^{(-i)} + \mu_+ \sum_i \gamma_i^+ x^{(+i)} = \theta_{n+1}$$
at that $\lambda_i^\pm \geqslant 0$ and $\sum\limits_i (\lambda_i^- + \lambda_i^+) = 1$; the statement holds in $\mathbb{R}^{n+1}$.

By MIP it holds for $\forall n \in \mathbb{N}$. ▶

<u>REC.</u> *Open ball with center x of radius r*, $B(x;r) := \{y \in \mathbb{R}^n \colon \|y - x\| < r\}$.

<u>PROP. 9.</u> Let $A \subseteq \mathbb{R}^n$ be convex and $x \notin A$. Then
$$\forall B(x;r)\ \exists B(y;\varepsilon) \subseteq B(x;r)\colon B(y;\varepsilon) \cap A = \varnothing$$
◀ Assume the contrary — $\exists B(x;r)\colon \forall B(y;\varepsilon) \subseteq B(x;r)\colon B(y;\varepsilon) \cap A \neq \varnothing$. Inscribe into $B(x;\frac{r}{2})$ the scaled symm. unit cube of $\mathbb{R}^n$ (with x as its center) and use L. 1 for $\delta < 1$ small enough for $B(x;r)$ to hold all 2^n balls of radius δ with centers at its vertices, $B(y^{(i)};\delta)$. Then by assumption $\exists x^{(i)} \in B(y^{(i)};\delta) \cap A$, which implies, by L. 1, $x \in A$ — $\bigotimes$. ▶

<u>PROP. 10.</u> If A is convex, then $\operatorname{Int} A = \operatorname{Int} \overline{A}$.

◀ $A \subseteq \overline{A}$, thus $\operatorname{Int} A \subseteq \operatorname{Int} \overline{A}$. Take $\forall x \in \operatorname{Int} \overline{A}$, by def. $\exists \delta > 0\colon B(x;\delta) \subseteq \overline{A}$. Then $\forall y \in B(x, \frac{\delta}{2})\colon B(y;\frac{\delta}{2}) \subseteq B(x;\delta) \subseteq \overline{A}$. Assume that $y \notin A$. By Prop. 9 $\exists B(z;\varepsilon) \subseteq B(y;\frac{\delta}{2})\colon B(z;\varepsilon) \cap A = \varnothing$. But then $z \notin \overline{A}$ — $\bigotimes$. Therefore all such $y \in A$ (in particular, x), implying $x \in \operatorname{Int} A$. Hence $\operatorname{Int} \overline{A} \subseteq \operatorname{Int} A$. ▶

. . . .

<u>COR.</u> If A is convex, then $\partial \overline{A} = \overline{\overline{A}} \setminus \operatorname{Int} \overline{A} = \overline{A} \setminus \operatorname{Int} A = \partial A$.

<u>DEF.</u> A plane H is called *supportive for $A \subseteq \mathbb{R}^n$ at (boundary) point $x \in \partial A$* if this plane contains x and one of semispaces — H_+ or H_- — contains A:
$$\forall u \in A\ \langle p;u\rangle \geqslant \alpha = \langle p;x\rangle \text{ or } \forall u \in A\ \langle p;u\rangle \leqslant \alpha = \langle p;x\rangle$$
(Substitution $p \leftarrow -p$ inverts the inequality sign.)

<u>DEF.</u> A plane H is called *properly supportive* for A at $x \in \partial A$ if H is accordingly supportive and $A \not\subseteq H$, that is, $\exists u' \in A\colon \langle p;u'\rangle > \alpha = \langle p;x\rangle$.

DEF. If a semispace S determined by (properly) supportive plane is such that $A \subseteq S$, then S is called *(properly) supportive for A at x* as well.

PROP. 11. At any boundary point $x \in \partial A$ of a convex set $A \subseteq \mathbb{R}^n$ a supportive plane exists.

◀ By the Cor. above $\partial\overline{A} = \partial A$, thus $x \in \partial\overline{A}$. By properties of $\partial(\cdot)$ $\exists\{x^{(k)}\}_{k=1}^{\infty}$, $x^{(k)} \in \mathbb{R}^n \backslash \overline{A}$: $x^{(k)} \xrightarrow[k\to\infty]{} x$. Due to Prop. 7, $\|\pi_{\overline{A}}(x^{(k)}) - x^{(k)}\| > 0$, hence we define $y^{(k)} = \frac{\pi_{\overline{A}}(x^{(k)}) - x^{(k)}}{\|\pi_{\overline{A}}(x^{(k)}) - x^{(k)}\|}$. They belong to the unit sphere $S = \{s\colon \|s\| = 1\}$, which is a closed and bounded set in $\mathbb{R}^n$, and therefore compact. Thus we can extract, from $\{y^{(k)}\}$, the subsequence that converges to $y \in S$; without restrictions on generality we assume that $y^{(k)} \xrightarrow[k\to\infty]{} y$ from the start. $\forall u \in A \subseteq \overline{A}$, by $(**)$ from Prop. 7, $\langle \pi_{\overline{A}}(x^{(k)}) - x^{(k)}; u - x^{(k)}\rangle \geqslant \|\pi_{\overline{A}}(x^{(k)}) - x^{(k)}\|^2 \Leftrightarrow \langle y^{(k)}; u - x^{(k)}\rangle \geqslant \|\pi_{\overline{A}}(x^{(k)}) - x^{(k)}\| > 0 \Leftrightarrow \langle y^{(k)}; u\rangle > \langle y^{(k)}; x^{(k)}\rangle$.

As $k \to \infty$, we get $\langle y; u\rangle \geqslant \langle y; x\rangle$, that is, sought plane H is determined by the parameters $p = y$ and $\alpha = \langle y; x\rangle$. ▶

REC. A *subspace/linear manifold* in $\mathbb{R}^n$ is $L \subseteq \mathbb{R}^n$, such that $\forall x, y \in L$, $\forall \alpha, \beta \in \mathbb{R}$: $\alpha x + \beta y \in L$. Number of elements of a base of L is called the *dimensionality* of subspace L, $\dim L$. $\forall z \in \mathbb{R}^n$ the set $z + L = \{z + x \mid x \in L\}$ is a *shifted* subspace; $\dim(z + L) := \dim L$.

The *dimensionality* of $A \subseteq \mathbb{R}^n$, $\dim A$, is the minimal subset of shifted subspace $z + L = \{z + x \mid x \in L\}$ that contains A (one of such subspaces is the entire $\mathbb{R}^n$).

Note that we can choose, as z, any p. of A, and choose elements of L base, $\{e^{(i)}\}$, in such a way as to $z + e^{(i)} \in A$ as well (in other words, draw the vectors of "shifted" base $z + L$ from z to some p. of A).

REM. If, under conditions of Prop. 11, $\dim A = n$, then a corresponding supportive plane is proper, $A \not\subseteq H$, because $\dim H = n - 1 < \dim A$.

DEF. Let $A \subseteq \mathbb{R}^n$ be convex. P. $x \in A$ is called *edge/extreme* if it is not an inner point of a segment with ends in A, that is, x cannot be represented as

$$x = \lambda y + (1 - \lambda)z, \text{ where } y, z \in A,\ y \neq z, \text{ and } 0 < \lambda < 1$$

DEF. The set of all edge points of A is denoted by $E(A)$.

On the plane E(convex polygon) is its vertices, E(disk) is the entire circle that is the disk's boundary.

REM. In general case $E(A) \neq \partial A$ (*boundary of convex polygon consists of its edges*), but $E(A) \subseteq \partial A$.

LEMMA. (On supportive plane) Let $A \subseteq \mathbb{R}^n$ be closed and convex, $x \in \partial A$, $H = H_{p;\alpha}$ is the properly supportive plane for A at x. Let $\widehat{A} = A \cap H$. Then

1) $E(\widehat{A}) \subseteq E(A)$, 2) $\dim \widehat{A} < \dim A$.

◀ By conditions, $\forall u \in A$: $\langle p; u\rangle \geqslant \alpha = \langle p; x\rangle$, at that $\exists u' \in A$: $\langle p; u'\rangle > \alpha$. Def. of $E(\widehat{A})$ is proper, — $\widehat{A}$ is convex by Prop. 2 as $\cap$ of convex A and H.

1) Take $\forall u \in E(\widehat{A}) \subseteq A \cap H$ and assume that $u \notin E(A)$. Therefore, $u = \lambda v + (1-\lambda)w$, where $v, w \in A$, $v \neq w$ and $0 < \lambda < 1$. It follows that

$$\alpha = \langle p; u\rangle = \lambda\langle p; v\rangle + (1-\lambda)\langle p; w\rangle \geqslant \lambda\alpha + (1-\lambda)\alpha = \alpha$$

and, for the inequality to be equality, it is necessary that $\langle p; v\rangle = \langle p; w\rangle = \alpha$. However, it means that $v, w \in H \cap A = \widehat{A}$, implying $u \notin E(\widehat{A})$ — $\bigotimes$. Hence $u \in E(A)$ and $E(\widehat{A}) \subseteq E(A)$.

2) Let the shifted subspaces $x + L$ and $x + \widehat{L}$ contain A and $\widehat{A}$ accordingly, at that $\dim L = \dim A$ and $\dim \widehat{L} = \dim \widehat{A}$. Since we can draw the vectors of shifted base $x + \widehat{L}$, denote it by B, from $x \in \widehat{A}$ to the p. of $\widehat{A}$, and $\widehat{A} \subseteq H$, we have $B \subseteq \widehat{A} \subseteq H$, thus $x + \widehat{L} \subseteq H$ as well. At that $\dim \widehat{A} = \dim(x + \widehat{L}) = |B|$.

Elements of B are lin.indep., and remain lin.indep. in $x + L \supseteq A \supseteq \widehat{A}$. Thus $\dim(x+L) \geqslant |B|$. Since $\langle p; u'\rangle > \alpha \Rightarrow u' \notin H \Rightarrow u' \notin (x+\widehat{L})$ (i.e. $u' \in A \subseteq x+L$ is outside of subspace spanned by the shifted base B), $B' = B \cup \{u'\}$ is lin.indep. system of shifted vectors too, so the shifted subspace it spans is contained in $x+L$, hence $\dim A = \dim(x + L) \geqslant |B'| > |B| = \dim \widehat{L}$. ▶

THEOR. (Minkowski, on convex compact set) Let $A \subset \mathbb{R}^n$ be a convex compact set. Then $A = \operatorname{conv} E(A)$.

◀ The inclusion $\operatorname{conv} E(A) \subseteq A$ follows from $E(A) \subseteq A$ and convexity of A (see Prop. 5); what is left is to prove the converse. Compact sets in $\mathbb{R}^n$ are closed and bounded, thus $A = \overline{A} = \operatorname{Int} A \cup \partial A$.

We apply induction by $\dim A = m$. Base, $m = 1$, holds: A is a segment then, so $A \subseteq \operatorname{conv} E(A)$ is obvious. Assume that the statement holds when $\dim A < m$, and let $\dim A = m$. For the sake of simplicity we suppose that $m = n$.

a) Take $\forall x \in \partial A$, let H be a properly supportive plane for A at x (such plane exists due to Prop. 11 and remark to it), and let $\widehat{A} = A \cap H$. Then $\widehat{A}$ is convex and compact, thus by Lemma on supportive plane $\dim \widehat{A} < \dim A = m$. By ind. supposition $\widehat{A} \subseteq \operatorname{conv} E(\widehat{A})$. It follows that $x \in \operatorname{conv} E(\widehat{A})$, by the same Lemma $E(\widehat{A}) \subseteq E(A) \Rightarrow x \in \operatorname{conv} E(A)$. Therefore $\partial A \subseteq \operatorname{conv} E(A)$.

b) Now take $\forall x \in \operatorname{Int} A$. For $\forall y \in \partial A$ we have $y \neq x$; consider the straight line $L = \{x + t(y-x) \mid t \in \mathbb{R}\}$. $L \cap A$ is the segment with the ends $y, z \in \partial A$; this segment is $\operatorname{conv}\{y; z\}$. Hence $x \in \operatorname{conv}\{y; z\} \subseteq \operatorname{conv} \partial A \subseteq \operatorname{conv}\big(\operatorname{conv} E(A)\big) = \operatorname{conv} E(A)$, therefore $\operatorname{Int} A \subseteq \operatorname{conv} E(A)$.

(a) & (b) imply $A \subseteq \operatorname{conv} E(A)$. By MIP the statement holds for $\forall \dim A$. ▶

1.2. Behaviour of target function on admissible domain

DEF. An edge point of a polyhedral set is called a *vertex* of this set.

PROP. 12. The set of vertices of a polyhedral set D is finite.

◀ By def. $D = \bigcap_{i=1}^{m} H^+_{p_i;\alpha_i}$. Let the collection of corresponding planes be $\mathcal{H} = \{H_{p_i;\alpha_i}\}_{i=1}^{m}$. We denote by V the set of vertices of D and consider arbitrary $v \in V$. $\langle p_i; v\rangle \geqslant \alpha_i$, $i = \overline{1,m}$; we divide $\mathcal{H}$ into 2 parts: $\mathcal{H}_=$, where inequality turns into equality, and $\mathcal{H}_>$ is the rest. In other words, v is contained in all planes of $\mathcal{H}_=$ and in the interiors of all $+$-semispaces determined by the planes of $\mathcal{H}_>$.

$\mathcal{H}_= \neq \varnothing$, otherwise we could choose p. u', u'', close enough to v and equidistant from it, that belong to Int of $+$-semispaces determined by planes of $\mathcal{H}_> = \mathcal{H}$. That is, $v = \frac{1}{2}u' + \frac{1}{2}u''$ would be an inner p. of a segment whose ends are in D, which contradicts $v \in V$.

Assume that $\exists x \in \mathbb{R}^n$: $x \in A = \bigcap_{H \in \mathcal{H}_=} H$ and $x \neq v$. Then $L = \{v + t(x - v) \mid t \in \mathbb{R}\}$ is the straight line that goes through v and x. $\forall y \in L$ and $\forall H_{p_j;\alpha_j} \in \mathcal{H}_=$: $y = tx + (1-t)v$, hence

$$\langle p_j; y\rangle = t\langle p_j; x\rangle + (1-t)\langle p_j; v\rangle = (t + 1 - t)\alpha_j = \alpha_j$$

that is, $y \in A$; so, $L \subseteq A$. Therefore we can choose, on L, points u' and u'', equidistant from v (so that $v = \frac{1}{2}u' + \frac{1}{2}u''$) and close enough to it, such points to belong to interior of all $+$-semispaces determined by planes of $\mathcal{H}_>$. But then v is an inner p. of a segment whose ends $u', u'' \in D$ — $\bigotimes$. The assumption was wrong, — besides v, there is no other p., particularly in V, that belongs to A.

So the function $f\colon V \to \mathfrak{B}(\mathcal{H})$ mapping the vertex v to the set of planes from $\mathcal{H}$ such that v belongs to them is injective. Therefore $|V| \leqslant |\mathfrak{B}(\mathcal{H})| = 2^m$. ▶ (This is a rough estimate, $|V| \leqslant C_m^n$ at most.)

COR. The set of vertices of a polyhedron is finite.

THEOR. Let the LP problem in $\mathbb{R}^n$

$$\sum_{j=1}^{n} a_{ij}x_j \leqslant b_i,\ i = \overline{1,m};\ f(x) = \langle c; x\rangle = \sum_{j=1}^{n} c_j x_j \to \max$$

be such that AD D is a polyhedron (thus bounded). Then

1) $f(x) \to \max$ at vertex of D;

2) if $f(\widehat{x}^{(k)}) = \max_{x\in D} f(x)$, $k = \overline{1,r}$, then at $\forall$ conv.comb. $x' = \sum_{k=1}^{r} \lambda_k \widehat{x}^{(k)}$, where $\lambda_k \geqslant 0$, $\sum_{k=1}^{r} \lambda_k = 1$, we have $f(x') = \max_{x\in D} f(x)$ as well.

◀ Let the set of vertices of D be $V = \{x^{(l)}\}_{l=1}^{s}$ ($s \in \mathbb{N}$ due to Prop. 12). D is bounded and closed, so D is compact, on which continuous $f(x) \to \max$ at some p. $x^* \in D$. If $x^* \in V$, Theor. is proved. If not, then by Minkowski Theor. on convex compact set $D = \operatorname{conv} V$ and, by Prop. 6, $x^* = \sum_{l=1}^{s} \lambda_l x^{(l)}$, $\lambda_l \geqslant 0$, $\sum_{l=1}^{s} \lambda_l = 1$. It follows that $\langle c; x^* \rangle = \sum_{l=1}^{s} \lambda_l \langle c; x^{(l)} \rangle$. $\forall l$: $\langle c; x^{(l)} \rangle \leqslant \langle c; x^* \rangle$, and it is easy to see that $\forall \lambda_l > 0$: $\langle c; x^{(l)} \rangle = \langle c; x^* \rangle$, otherwise $\langle c; x^* \rangle < \sum_{l:\, \lambda_l > 0} \lambda_l \langle c; x^* \rangle = \langle c; x^* \rangle$ — $\bigotimes$. Take any $\lambda_l > 0$, then corresponding $x^{(l)}$ is the sought vertex, which satisfies (1).

(2) follows analogously: $f(x') = \langle c; \sum_{k=1}^{r} \lambda_k \widehat{x}^{(k)} \rangle = \sum_{k=1}^{r} \lambda_k f(\widehat{x}^{(k)}) = \max_{x \in D} f(x)$. ▶

COR. To solve LP problem with AD being a polyhedron, it suffices to search through all its vertices.

Here we could stop... were we not interested in *speed* of finding the solution. To search through all vertices of n-dimensional polyhedron, in arbitrary order, — simply calculating the TF at each vertex and neglecting of the *connection* between TF values at different vertices, — is a "long" procedure; can we accelerate the search? In particular, we want to be able to

1) find vertices of AD faster than "search through all combinations of n planes from given m and verify, for each, if these planes intersect at a single point that satisfies other constraints";

2) move from current vertex of AD not to the "randomly" chosen next one, but to the vertex where TF is as larger as it can be than at the current vertex, and do not return to vertices we have already considered.

To find a vertex of AD where TF attains its extremal value is the optimization; here we speak about how to do it faster — the "optimization of optimization".

1.3. Standard form of LPP and base solutions

From the *general* form of LPP we move to the *standard* one, where all constraints except for $x_j \geqslant 0$ are equalities. In order to do that, we transform each such constraint, $\sum_{j=1}^{n} a_{ij} x_j \leqslant b_i$, into $\sum_{j=1}^{n} a_{ij} x_j + x_{n+i} = b_i$, where the *fictitious* (fict.) variable $x_{n+i} \geqslant 0$. This way the dimensionality of the problem increases from n to $n + m$. We remark that TF contains fict. variables with coefficients $c_j = 0$.

REM. In more general case, when some x_j can take negative values (i.e. the constraint $x_j \geqslant 0$ is absent), they are "split": $x_j = x_j^+ - x_j^-$, where $x_j^+ \geqslant 0$, $x_j^- \geqslant 0$.

We re-denote $n \leftarrow n+m$ (so $n>m$), then the standard form of LPP is to find $x \in \mathbb{R}^n$ such that

$$x_j \geqslant 0,\ j=\overline{1,n};\ \sum_{j=1}^{n} a_{ij}x_j = b_i,\ i=\overline{1,m};\ n>m;\ f(x)=\sum_{j=1}^{n} c_j x_j \to \max$$

DEF. $A_j = (a_{1j}; a_{2j}; ...; a_{mj})^T$, $j=\overline{1,n}$, and $b=(b_1; ...; b_m)^T$ are called *condition vectors* (CV).

We rewrite LPP in standard form with CVs:

$$x_j \geqslant 0,\ j=\overline{1,n};\ \sum_{j=1}^{n} x_j A_j = b;\ f(x)=\langle c; x\rangle \to \max$$

DEF. An admissible solution $x=(x_1; ...; x_n)$ of LPP in standard form is called *base/supportive* if the set of CVs $\{A_j \mid x_j>0\}$ is lin.indep. (in $\mathbb{R}^m$).

THEOR. Let x be an admissible solution of LPP with AD D.

x is a vertex of D **iff** x is base.

◀ **Sufficiency.** Let x be a base sol. We renumerate coordinates to have $x=(x_1; ...; x_k; 0; ...; 0)$, where $x_j>0$, $j=\overline{1,k}$. By def. of adm.sol. $\sum_{j=1}^{k} x_j A_j = \sum_{j=1}^{n} x_j A_j - \theta_m = b$, by def. of base sol. $\{A_j\}_{j=1}^k$ are lin.indep.

Assume that x is not a vertex of D: $x=\lambda x^{(1)}+(1-\lambda)x^{(2)}$, where $x^{(1)}, x^{(2)} \in D$, $x^{(1)} \neq x^{(2)}$ and $0<\lambda<1$.

From the last condition and $x_j \geqslant 0$ it follows that $x^{(i)}=(x_1^{(i)}; ...; x_k^{(i)}; 0; ...; 0)$, $i=1,2$. At that $\sum_{j=1}^{k} x_j^{(i)} A_j = b$, $i=1,2$, hence $\sum_{j=1}^{k}(x_j^{(1)}-x_j^{(2)})A_j=\theta_m$. And since $\{A_j\}_{j=1}^k$ are lin.indep., we have $x_j^{(1)}-x_j^{(2)}=0$, $j=\overline{1,k}$, thus $x^{(1)}=x^{(2)}$ — $\bigotimes$.

Necessity. Let $x=(x_1; ...; x_k; 0; ...; 0)$, where $x_j>0$ for $j=\overline{1,k}$, be a vertex of D. Assume that the corresponding CV set $\{A_j\}_{j=1}^k$ is lin.dep., that is, $\exists \{\alpha_j\}_{j=1}^k$, $\alpha_j \in \mathbb{R}$: $\sum_{j=1}^{k} \alpha_j^2 \neq 0$, $\sum_{j=1}^{k} \alpha_j A_j = \theta_m$. Then $\forall \varepsilon>0$ we have

$$\sum_{j=1}^{k}(x_j \pm \varepsilon\alpha_j)A_j = \sum_{j=1}^{k} x_j A_j \pm \varepsilon \sum_{j=1}^{k} \alpha_j A_j = b \pm \theta_m = b$$

For small enough $\varepsilon>0$ all $x_j \pm \varepsilon\alpha_j > 0$ and both p. $x^{\pm}=(x_1 \pm \varepsilon\alpha_1; ...; x_k \pm \varepsilon\alpha_k; 0; ...; 0) \in D$, at that $x^+ \neq x^-$. From the equalities provided above it follows that $x=\frac{1}{2}x^+ + \frac{1}{2}x^-$ is the inner p. of a segment whose ends belong to D, thus $x \notin E(D)$; we have got $\bigotimes$. Hence $\{A_j\}_{j=1}^k$ are lin.indep. and x is base. ▶

Instead of searching through combinations of planes to find vertices we could search through lin.indep. subsets of $\{A_j\}_{j=1}^n$ and solve corresponding equation sets $\sum\limits_{j=1}^{k} x_j A_j = \theta_m$, then let $x_{k+1} = ... = x_n = 0$, verify $x_j \geqslant 0$ for $j = \overline{1,k}$ and calculate TF if these constraints hold. However, this method is "ineffective" as well from speed point of view, — we spend time on too many vertices.

1.4. Canonical form of LPP and Jordan-Gauss exclusion method

DEF. LPP in standard form is written in *canonical* form if $\forall i = \overline{1,m}$:

1) $b_i \geqslant 0$, 2) $\exists j_i$: $a_{i,j_i} = 1$ and $\forall k \neq i$ $a_{k,j_i} = 0$.

(2) means that each equation has a variable that is present only in this equation with coefficient 1.

M method. To turn standard form into a canonical one, consider each of m equations of the former, multiply by (-1) if $b_i \leqslant 0$ to make right hand side $\geqslant 0$, and introduce into the left hand side the *artificial* (art.) variable $x_{n+i} \geqslant 0$ (in addition to the fictitious ones at transition from general to standard form) with coef. $+1$. To ensure that these new variables do not "distort" the sought solution of the initial problem, add them to TF as well, but with coef. $-M < 0$ whose absolute value is large enough rather than with zero coef.: $f(x) \leftarrow f(x_1; ...; x_n) - M \sum\limits_{i=1}^{m} x_{n+i}$. For such M $f(x) \to \max$ only if $x_{n+1} = x_{n+2} = ... = x_{n+m} = 0$.

Again, we re-denote $n \leftarrow n+m$, renumerate variables such that art. ones have indices $1, ..., m$, and write canonical form:

$$x_j \geqslant 0,\ j = \overline{1,n};\ x_i + \sum_{j=m+1}^{n} \alpha_{ij} x_j = \beta_i \geqslant 0,\ i = \overline{1,m};\ f(x) = \langle c; x \rangle \to \max$$

where $x_1, ..., x_m$ are *base* and $x_{m+1}, ..., x_n$ are *free* variables. We can rewrite these equations in more general way $\sum\limits_{j=1}^{n} \alpha_{ij} x_j = \beta_j$, the coef. α_{ij} being δ_{ij} for $j \leqslant m$.

From equations of canonical form, base variables are expressed by free ones:

$$x_i = \beta_i - \sum_{j=m+1}^{n} \alpha_{ij} x_j,\ i = \overline{1,m}$$

For clearness we provide the matrix of constraints' coefficients:

$$C = \begin{pmatrix} 1 & 0 & \cdots & 0 & \alpha_{1,m+1} & \alpha_{1,m+2} & \cdots & \alpha_{1,n} \\ 0 & 1 & \cdots & 0 & \alpha_{2,m+1} & \alpha_{2,m+2} & \cdots & \alpha_{2,n} \\ \cdots & \cdots & \ddots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & \alpha_{m,m+1} & \alpha_{m,m+2} & \cdots & \alpha_{m,n} \end{pmatrix}$$

For LPP in this form we have (coordinates of) one vertex of AD right away: $x^{(0)} = (\beta_1; ...; \beta_m; 0; ...; 0)$. Indeed, $x = x^{(0)}$ satisfies all equations-constraints,

$\beta_j \geqslant 0$ and CVs A_j that correspond to $\beta_j > 0$ look like $(0; ...; 0; 1; 0; ...; 0)$ with 1 at j-th place, thus the system $\{A_j \mid \beta_j > 0\}$ is a subset of the base of $\mathbb{R}^m$, therefore this system is lin.indep., hence by the Theor. above $x^{(0)}$ is a vertex.

How to move from current vertex to the next one? In particular, if the vertex is obtained in this way from current canonical form, how to move to another canonical form, which will determine the next vertex?

Jordan-Gauss method. We choose the free variable x_k $(k \geqslant m+1)$ and exclude it from all equations except for l-th, where $\alpha_{lk} \neq 0$. In order to do that, we subtract from i-th equation $(i = 1, ..., l-1, l+1, ..., m)$ the l-th one, multiplied by $\frac{\alpha_{ik}}{\alpha_{lk}}$. Coef. at x_i in i-th equation remains 1, because other equations do not have x_i. Then we divide l-th equation by α_{lk} to make the coef. at x_k in it equal to 1. So, new coefficients become

$$\alpha'_{ij} = \begin{cases} \frac{\alpha_{lj}}{\alpha_{lk}}, & i = l, \\ \alpha_{ij} - \frac{\alpha_{ik}}{\alpha_{lk}}\alpha_{lj}, & i \neq l, \end{cases} \qquad \beta'_i = \begin{cases} \frac{\beta_l}{\alpha_{lk}}, & i = l, \\ \beta_i - \frac{\alpha_{ik}}{\alpha_{lk}}\beta_l, & i \neq l \end{cases}$$

(in particular, $\alpha'_{ik} = 0$ when $i \neq l$, and $\alpha'_{lk} = 1$). And we obtain the equation set

$$\begin{cases} \alpha'_{ll}x_l + x_k + \sum\limits_{m+1 \leqslant j \leqslant n,\, j \neq k} \alpha'_{lj}x_j = \beta'_l, & i = l, \\ x_i + \alpha'_{il}x_l + \sum\limits_{m+1 \leqslant j \leqslant n,\, j \neq k} \alpha'_{ij}x_j = \beta'_i, & i = 1, ..., l-1, l+1, ..., m. \end{cases} \tag{\#}$$

This transition is defined by the pair $(k; l)$, where $m+1 \leqslant k \leqslant n$, $1 \leqslant l \leqslant n$, $\alpha_{lk} \neq 0$; however, in general case these restrictions are not enough for the new set (#) to be a canonical form as well, — we also need $\beta'_i \geqslant 0$. From this constraint for $i = l$, $\frac{\beta_l}{\alpha_{lk}} \geqslant 0$, and the fact that $\beta_l \geqslant 0$, it follows that $\alpha_{lk} > 0$; at least 1 component of CV A_k, which corresponds to x_k, must be positive.

Consider $i \neq l$. If $\alpha_{ik} \leqslant 0$, the condition $\beta'_i \geqslant 0$ holds, because $\beta'_i \geqslant \beta_i \geqslant 0$, and if $\alpha_{ik} > 0$, we need

$$\beta_i - \frac{\alpha_{ik}}{\alpha_{lk}}\beta_l \geqslant 0 \Leftrightarrow \frac{\beta_i}{\alpha_{ik}} \geqslant \frac{\beta_l}{\alpha_{lk}},\ i = 1, ..., l-1, l+1, ..., m$$

Therefore we must choose l to satisfy $\beta'_l = \frac{\beta_l}{\alpha_{lk}} = \theta := \min\limits_{i:\, \alpha_{ik} > 0} \frac{\beta_i}{\alpha_{ik}}$. Rewrite $\beta'_i = \beta_i - \theta\alpha_{ik}$ for $i \neq l$.

This way we have moved to another canonical form that determines the next vertex of AD

$$x' = (\beta'_1; ...; \beta'_{l-1}; 0; \beta'_{l+1}; ...; \beta'_m; 0; ...; 0; \beta'_l; 0; ...; 0), \quad \text{where } \beta'_l = \theta \text{ is } x'_k$$

1.5. Simplex Method, Optimality Criterion, Unboundedness Indicator

In the way we have described it is natural to move not to arbitrary vertex x' of AD, but to the one at which TF is larger than at the initial $x^{(0)}$ ("closer" to the vertex-solution of $f(x) = \langle c; x \rangle \to \max$ problem).

$$f(x') = \langle c; x' \rangle = \sum_{j=\overline{1,m},\, j\neq l} c_j\beta'_j + c_k\beta'_l =$$

$$= \Big[\underbrace{\beta_l - \theta\alpha_{lk}}_{0}\Big] + \sum_{j=\overline{1,m},\, j\neq l} c_j(\beta_j - \theta\alpha_{jk}) + c_k\theta =$$

$$= \sum_{j=1}^{m} c_j(\beta_j - \theta\alpha_{jk}) + c_k\theta = \langle c; x^{(0)} \rangle + \theta\big(c_k - \sum_{j=1}^{m} c_j\alpha_{jk}\big) = f(x^{(0)}) + \theta(c_k - z_k)$$

where $z_k = \sum\limits_{j=1}^{m} c_j\alpha_{jk}$. Aiming at $f(x') > f(x^{(0)})$ and taking into account $\theta \geqslant 0$, we need, clearly, $c_k - z_k > 0$.

DEF. $\Delta_k = c_k - z_k$ is called a *relative estimate* (RE).
(For $k \leqslant m$: $z_k = \sum\limits_{j=1}^{m} c_j\delta_{jk} = c_k$, $\Delta_k = 0$.)

Simplex Method (SM) consists of the following steps, which correspond to consecutive transition between AD vertices with TF increasing:

[0] Write LPP in canonical form that determines the initial vertex $x^{(r)} = (\beta_1; ...; \beta_m; 0; ...; 0)$, $r = 0$.
[1] For each free variable x_k, where $k = \overline{m+1, n}$, calculate Δ_k.
[2] $\forall\Delta_k \leqslant 0$? If yes — halt: current vertex is the sought one. If no — continue:
[3] Does x_k exist such that $\Delta_k > 0$ and all $\alpha_{ik} \leqslant 0$, $i = \overline{1,m}$? If yes — halt: TF is unbounded, LPP does not have a solution. If no — continue:
[4] Choose x_k such that $\Delta_k > 0$ and $\exists\alpha_{ik} > 0$.
[5] Calculate $\theta = \min\limits_{i:\, \alpha_{ik}>0} \frac{\beta_i}{\alpha_{ik}}$; suppose the min is attained in l-th equation.
[6] By Jordan-Gauss method exclude x_k from all equations except for l-th and obtain the next canonical form that determines the next vertex $x^{(r+1)}$.
[7] Increment $r := r + 1$ and go to step (1).

The process is finite, because at each (1)–(7) iteration TF increases, and it takes a finite set of values on a finite set of AD vertices. The outcome is either the conclusion that LPP does not have a solution (we will ground such conclusion below), or the vertex where all RE $\Delta_k \leqslant 0$. Does really $f(x) \to \max$ there?

PROP. 13. Let, in stand. form of given LPP, CVs be A_j, $j = \overline{1,n}$, and in canon. form of this LPP, which (form) corresponds to the vertex x, the coefs. at x_j in equalities be α_{ij}. Then $A_j = \sum\limits_{i=1}^{m} \alpha_{ij}A_i$.

◀ Column CVs A_j form the matrix $A = (A_1, ..., A_m, A_{m+1}, ..., A_n)$. Matrix of canon. form $C = \|\alpha_{ij}\|_{\substack{1\leqslant i\leqslant m,\\ 1\leqslant j\leqslant n}}$ is obtained by writing to the right of the $m \times m$

identity matrix the $m \times (n-m)$ matrix with coefs. α_{ij}, where $i = \overline{1,m}$, $j = \overline{m+1,n}$. Since x is a base sol., $\{A_i\}_{i=1}^m$ are lin.indep.: it is a base of $\mathbb{R}^m$. Hence, having made the matrix B from column vectors A_i, $B = (A_1, ..., A_m)$, we represent the connection between the coordinates of corresponding vectors: $A = BC$. That is, j-th column of A, $A_j = BC_j = \sum\limits_{i=1}^m A_i \alpha_{ij} = \sum\limits_{i=1}^m \alpha_{ij} A_i$. ►

THEOR. (Optimality Criterion) Let the vertex/base sol. $x = (x_1; ...; x_n)$ be such that $\Delta_k \leqslant 0$, $k = \overline{1,n}$.

Then x is a solution of LPP with AD D: $\forall y \in D \ f(y) \leqslant f(x)$.

◄ Let $y = (y_1; ...; y_n)$ be arbitrary admiss.sol., that is, $y_j \geqslant 0$, $j = \overline{1,n}$, and $\sum\limits_{j=1}^n y_j A_j = b$. Substitute the representation from Prop. 13 into this equality: $b = \sum\limits_{j=1}^n y_j A_j = \sum\limits_{j=1}^n y_j \big(\sum\limits_{i=1}^m \alpha_{ij} A_i\big) = \sum\limits_{i=1}^m \big(\sum\limits_{j=1}^n \alpha_{ij} y_j\big) A_i$.

But $b = \sum\limits_{i=1}^m x_i A_i$, because x is also an admiss.sol. So $\sum\limits_{i=1}^m \big[x_i - \sum\limits_{j=1}^n \alpha_{ij} y_j\big] A_i = \theta_m$. From that and lin.indep. of CVs $A_1, ..., A_m$ we derive $\sum\limits_{j=1}^n \alpha_{ij} y_j = x_i$, $i = \overline{1,m}$. Since $\Delta_k \leqslant 0 \Leftrightarrow c_k \leqslant z_k$, $f(y) = \sum\limits_{j=1}^n c_j y_j \leqslant \sum\limits_{j=1}^n z_j y_j =$

$$= \sum_{j=1}^n \big(\sum_{i=1}^m c_i \alpha_{ij}\big) y_j = \sum_{i=1}^m \big(\sum_{j=1}^n \alpha_{ij} y_j\big) c_i = \sum_{i=1}^m c_i x_i = f(x) \ ►$$

THEOR. (Unboundedness Indicator) Let the vertex (or, equivalently, base sol.) $x = (x_1; ...; x_m; 0; ...; 0)$ be such that for some free variable x_k we have $\Delta_k > 0$ and $\alpha_{ik} \leqslant 0$, $i = \overline{1,m}$. Then TF is not bounded from above on AD D (this implies the unboundedness of D as well).

◄ We claim that $\forall \theta > 0$: $x' = (x_1 - \theta\alpha_{1k}; ...; x_m - \theta\alpha_{mk}; 0; ...; 0; \theta; 0; ...; 0)$, where θ stands at k-th position, is an admiss.sol. Since $x_j \geqslant 0$ and $\alpha_{jk} \leqslant 0$, we have $x'_j \geqslant 0$ too, $j = \overline{1,m}$. Prop. 13 implies $\sum\limits_{j=1}^n x'_j A_j =$

$$= \sum_{j=1}^m (x_j - \theta\alpha_{jk}) A_j + \theta A_k = \sum_{j=1}^m x_j A_j - \theta \sum_{j=1}^m \alpha_{jk} A_j + \theta A_k = b - \theta A_k + \theta A_k = b$$

Thus x' is an admiss.sol. And $f(x') = \langle c; x' \rangle = \sum\limits_{j=1}^m c_j (x_j - \theta\alpha_{jk}) + c_k \theta = \langle c; x \rangle + \theta(c_k - z_k) = f(x) + \theta\Delta_k$. As $\theta \to +\infty$, we have $f(x') \to +\infty$ as well. ►

REM. The former of these 2 theorems is a criterion, because we can reverse its statement. Indeed, let the vertex x be a solution of LPP. Assume that $\exists \Delta_k > 0$. If $\forall i$: $\alpha_{ik} \leqslant 0$, then, by Unboundedness Indicator, there is no solution of LPP, and if $\exists\alpha_{ik} > 0$, then we can move to another vertex with larger value of TF, thus x is not a solution; either way, we get $\bigotimes$. Therefore $\forall\Delta_k \leqslant 0$.

Simplex Tables (ST). Coefficients and free members of equation sets and linear forms that appear at each SM iteration of LPP solving are written as tables for convenience. The simplest version is as follows:

	x_1	...	x_m	x_{m+1}	...	x_k	...	x_n	
x_1	1	...	0	$\alpha_{1,m+1}$	...	$\alpha_{1,k}$	...	$\alpha_{1,n}$	β_1
...	...	...	...	...	...	...	...	...	...
x_l	0	...	0	$\alpha_{l,m+1}$	...	$\alpha_{l,k}$	...	$\alpha_{l,n}$	β_l
...	...	...	...	...	...	...	...	...	...
x_m	0	...	1	$\alpha_{m,m+1}$	...	$\alpha_{m,k}$	...	$\alpha_{m,n}$	β_m
f	c_1	...	c_m	c_{m+1}	...	c_k	...	c_n	c_0
z		...		z_{m+1}	...	z_k	...	z_n	
Δ		...		Δ_{m+1}	...	Δ_k	...	Δ_n	

REM. In general case, at each SM iteration the corresponding ST changes; in particular, the collections of base and free variables "exchange" one variable each (base variables get k-th instead of l-th, and vice versa for free variables). Thus the collections of indices are not constant: the denotations "x_1" or "x_{15}" in the table above at different iterations are *names* of different input variables: initial ones as well as fictitious or artificial ones.

First m rows of ST correspond to the equations that represent the base variables by the free ones: $x_j = \beta_j - \sum\limits_{r=m+1}^{n} \alpha_{jr}x_r$ for $j = \overline{1,m}$. The last column on the right contains the free members.

DEF. The coef. of ST α_{lk} such that $\Delta_k > 0$ and $\frac{\beta_l}{\alpha_{lk}} = \min\limits_{i:\,\alpha_{ik}>0} \frac{\beta_i}{\alpha_{ik}}$ is called a *general element.*

Now we rewrite TF $f(x) = c_0 + \sum\limits_{j=1}^{n} c_jx_j$, expressing the base variables by the free ones: $f(x) = c_0 + \sum\limits_{j=1}^{m} c_j\big[\beta_j - \sum\limits_{r=m+1}^{n} \alpha_{jr}x_r\big] + \sum\limits_{j=m+1}^{n} c_jx_j =$

$$= \underbrace{\Big[c_0 + \sum_{j=1}^{m} c_j\beta_j\Big]}_{c_0'} + \sum_{j=m+1}^{n} \underbrace{\Big[c_j - \overbrace{\sum_{r=1}^{m} c_r\alpha_{rj}}^{z_j}\Big]}_{c_j'} x_j$$

Clearly, $f(x) \to \max$ **iff** $g(x) = \sum\limits_{j=m+1}^{n} c'_j x_j \to \max$. Being represented this way, TF depends only on free variables and $c'_j = c_j - z_j = \Delta_j$ "plays part" of Δ_j in the initial representation of TF. We can remove 2 lowest rows from the table,

	x_1	...	x_m	x_{m+1}	...	x_k	...	x_n	
x_1	1	...	0	$\alpha_{1,m+1}$	...	$\alpha_{1,k}$	...	$\alpha_{1,n}$	β_1
...	...	...	...	...	...	...	...	...	...
x_l	0	...	0	$\alpha_{l,m+1}$	...	$\alpha_{l,k}$	...	$\alpha_{l,n}$	β_l
...	...	...	...	...	...	...	...	...	...
x_m	0	...	1	$\alpha_{m,m+1}$	...	$\alpha_{m,k}$	...	$\alpha_{m,n}$	β_m
g	0	...	0	c'_{m+1}	...	c'_k	...	c'_n	

and apply SM, assigning $\Delta_k = c_k$ after such transition at the beginning of each iteration. In other words, the step (1) is split into

1.1) Re-express TF by free variables, $\{c_j\} \leftarrow \{c'_j\}$.

1.2) For each free variable x_k, where $k = \overline{m+1, n}$, let $\Delta_k = c_k$.

. . . .

EX. 3. Let us solve LPP from Ex. 1 (see p. 9): $f(x) = -x_1 + x_2 \to \max$ under constraints $x_1 \geqslant 0$, $x_2 \geqslant 0$,

$$-x_1 + 2x_2 \leqslant 2, \qquad -x_1 - 2x_2 \leqslant -2, \qquad 4x_1 - x_2 \leqslant 8, \qquad x_1 + x_2 \leqslant 4,$$

by means of SM using ST. First of all, move to standard form:

$-x_1 + 2x_2 + x_3 = 2$, $-x_1 - 2x_2 + x_4 = -2$, $4x_1 - x_2 + x_5 = 8$, $x_1 + x_2 + x_6 = 4$, where fictitious variables $x_3, ..., x_6 \geqslant 0$. Then multiply 2nd equation by -1 and add the artificial variable $x_7 \geqslant 0$ to it; also add the same variable to TF with coef. $-M$, where M is "large enough":

$-x_1+2x_2+x_3 = 2$, $x_1+2x_2-x_4+x_7 = 2$, $4x_1-x_2+x_5 = 8$, $x_1+x_2+x_6 = 4$, $f(x) = -x_1 + x_2 - Mx_7 \to \max$. For the sake of definiteness, let $M = 100$.

As base variables, we take x_3, x_7, x_5, and x_6, — each of them appears in exactly one equation with coef. $+1$, — then x_1, x_2, and x_4 are free. ST for the corresponding canonical form looks as follows:

	x_3	x_7	x_5	x_6	x_1	x_2	x_4	
x_3	1	0	0	0	-1	2	0	2
x_7	0	1	0	0	1	2	-1	2
x_5	0	0	1	0	4	-1	0	8
x_6	0	0	0	1	1	1	0	4
f	0	-100	0	0	-1	1	0	

The initial vertex $x^{(0)} = (0; 0; 2; 0; 8; 4; 2)$. The step **(0)** is done.

1.1) Express TF by free variables:

$$f(x) = -x_1 + x_2 - 100(2 - x_1 - 2x_2 + x_4) = -200 + 99x_1 + 201x_2 - 100x_4$$

1.2) $\Delta_1 = c_1 \leftarrow c'_1 = 99$, $\Delta_2 = c_2 \leftarrow c'_2 = 201$, $\Delta_4 = c_4 \leftarrow c'_4 = -100$; ST

	x_3	x_7	x_5	x_6	x_1	x_2	x_4	
x_3	1	0	0	0	-1	2	0	2
x_7	0	1	0	0	1	2	-1	2
x_5	0	0	1	0	4	-1	0	8
x_6	0	0	0	1	1	1	0	4
g	0	0	0	0	99	201	-100	

2) Are all $\Delta_k \leqslant 0$? No, continue.

3) Does $\Delta_k > 0$ exist such that all $\alpha_{ik} \leqslant 0$? No, continue.

4) Make x_2 the new base variable; it is possible, since $\Delta_2 = 201 > 0$ and $\alpha_{3,2} = 2 > 0$.

5) $\min\{\frac{\beta_i}{\alpha_{i,2}} \mid \alpha_{i,2} > 0\} = \min\{\frac{2}{2}; \frac{2}{2}; \frac{4}{1}\} = 1$ is attained for $l = i = 3$; base variables will include x_2 instead of x_3.

6) By Jordan-Gauss method exclude x_2 from all equations except for $i = 3$, and divide that one by $\alpha_{3,2} = 2$. We obtain new canonical form with ST

	x_2	x_7	x_5	x_6	x_1	x_3	x_4	
x_2	1	0	0	0	$-\frac{1}{2}$	$\frac{1}{2}$	0	1
x_7	0	1	0	0	2	-1	-1	0
x_5	0	0	1	0	$\frac{7}{2}$	$\frac{1}{2}$	0	9
x_6	0	0	0	1	$\frac{3}{2}$	$-\frac{1}{2}$	0	3
f	201	0	0	0	99	0	-100	

and corresponding vertex $x^{(1)} = (0; 1; 0; 0; 9; 3; 0)$.

7) Go to the next iteration.

TF $f(x) = 201(1 + \frac{1}{2}x_1 - \frac{1}{2}x_3) + 99x_1 - 100x_4 = 201 + \frac{399}{2}x_1 - \frac{201}{2}x_3 - 100x_4$. We rewrite ST:

	x_2	x_7	x_5	x_6	x_1	x_3	x_4	
...	...	...	...	...	...	...	...	...
g	0	0	0	0	$\frac{399}{2}$	$-\frac{201}{2}$	-100	

It is easy to see that we can make only x_1 a new base variable, and $\min\{\frac{\beta_i}{\alpha_{i,1}} \mid \alpha_{i,1} > 0\} = 0$ is attained for $l = i = 7$. The Jordan-Gauss method results in the new canonical form with ST

	x_2	x_1	x_5	x_6	x_7	x_3	x_4	
x_2	1	0	0	0	$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{4}$	1
x_1	0	1	0	0	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	0
x_5	0	0	1	0	$-\frac{7}{4}$	$\frac{9}{4}$	$\frac{7}{4}$	9
x_6	0	0	0	1	$-\frac{3}{4}$	$\frac{1}{4}$	$\frac{3}{4}$	3
f	0	$\frac{399}{2}$	0	0	0	$-\frac{201}{2}$	-100	

and the corresponding vertex $x^{(2)} = (0; 1; 0; 0; 9; 3; 0)$.

At the next iteration the TF

$f(x) = \frac{399}{2}(-\frac{1}{2}x_7 + \frac{1}{2}x_3 + \frac{1}{2}x_4) - \frac{201}{2}x_3 - 100x_4 = -\frac{399}{4}x_7 - \frac{3}{4}x_3 - \frac{1}{4}x_4$
— in the rewritten ST all $\Delta_k \leqslant 0$, so at step **(2)** the process halts and the vertex $x^{(2)}$ is the sought one. At this point the initial coordinates $x_1^{(2)} = 0$, $x_2^{(2)} = 1$, so the solution of LPP is $x^* = (0; 1)$, where $f(x^*) = -0 + 1 = 1$.

The same solution has been obtained by graphic method.

1.6. Dual (adjoint) LPP

Consider some LPP that we also call *direct*, in standard form:

$$x_j \geqslant 0,\ j = \overline{1, n};\ \sum_{j=1}^{n} a_{ij}x_j = b_i,\ i = \overline{1, m};\ f(x) = \langle c; x\rangle \to \max \qquad (\#)$$

DEF. The *dual* (*adjoint*) LPP to this LPP is the following:

$$u_i \geqslant 0,\ i = \overline{1, m};\ \sum_{i=1}^{m} a_{ij}u_i \geqslant c_j,\ j = \overline{1, n};\ \langle b; u\rangle = \sum_{i=1}^{m} b_i u_i \to \min \qquad (\#\#)$$

Using the denotations introduced earlier, CVs $A_j = (a_{1j}; ...; a_{mj})^T$, $j = \overline{1, n}$, $b = (b_1; ...; b_m)^T$ and the matrix $A = (A_1; ...; A_n)$ made from $\{A_j\}$, also, letting $u = (u_1; ...; u_m)^T$, we rewrite these problems (inequality sign between vectors means that this inequality holds for each coordinate):

Direct: $x \in \mathbb{R}^n_+$, $Ax = b$, $\langle c; x\rangle \to \max$.
Dual: $u \in \mathbb{R}^m_+$, $u^T A \geqslant c$, $\langle b; u\rangle \to \min$.

DEF. The variables $\{u_i\}_{i=1}^m$ are called *dual*, or *Lagrangian multipliers.*

DEF. The problems of the (#)–(##) pair are called *symmetric dual LPPs.*

Since $Ax - b = \theta_m$, we have $\langle c; x\rangle = \langle c; x\rangle + \langle b - Ax; u\rangle = \langle b; u\rangle + \langle c - u^T A; x\rangle$. If we choose u such that $u^T A \geqslant c$, then, taking into account $x \geqslant \theta_n$, we have $\langle c - u^T A; x\rangle \leqslant 0$, implying $\langle b; u\rangle \geqslant \langle c; x\rangle$.

Moreover, we can presume that $\min\limits_u \langle b; u\rangle = \max\limits_x \langle c; x\rangle$. And it is true, as the following shows:

THEOR. (First Duality Theorem) **1)** If one of symmetric dual LPPs has an optimal solution (OS), then another LPP has OS as well, at that the optimal values of target linear forms are equal.

2) If, in one of symmetric dual LPPs, the TF is unbounded on AD, then another LPP does not have admissible solutions.

(Proof is omitted.)

<u>THEOR.</u> (Dual Optimality Criterion) An admiss.sol. x^* of LPP (#) is optimal **iff** $\exists \{u_i\}_{i=1}^m \in \mathbb{R}_+^m$:

$$\sum_{i=1}^m a_{ij} u_i \begin{cases} = c_j, & x_j^* > 0, \\ \geqslant c_j, & x_j^* = 0. \end{cases}$$

(Proof is omitted.)

<u>COR.</u> Simplex-multipliers in the TF that corresponds to OS of the direct LPP are the solution of the dual LPP.

Economic interpretation of dual LPP. Consider the problem of distribution of m resources at manufacturing of n productions, in general form:

$$x_j \geqslant 0,\ j = \overline{1,n};\ \sum_{j=1}^n a_{ij} x_j \leqslant b_i,\ i = \overline{1,m};\ f(x) = \langle c; x \rangle \to \max$$

We can interpret the dual problem

$$u_i \geqslant 0,\ i = \overline{1,m};\ \sum_{i=1}^m a_{ij} u_i \geqslant c_j,\ j = \overline{1,n};\ \langle b; u \rangle = \sum_{i=1}^m b_i u_i \to \min$$

as follows: let u_i be the specific value of i-th resource, expressed by the same units as the profit is measured in ("the cost of 1 unit of i-th resource"), then $\sum_{i=1}^m a_{ij} u_i$ is the value of total expenses of manufacturing 1 unit of j-th production (it is also called "the manufacturing with unit intensity"), since u_i are the specific values, and the constraints of the dual problem arise because, concerning each production, the profit/loss as the outcome of the manufacturing must not exceed the expenses. TF of dual LPP is the total value of all resources. Thus, we seek to find such specific values of the resources that their total cost is minimal.

In this context, First Duality Theorem means that the existence of the optimal manufacturing plan implies the existence of the optimal resource values, and the maximal possible profit is equal to the minimal total value of the resources.

2. Elements of Game Theory

In general, a *game* is a situation where participants — *players* — with different interests interact. Particularly, when the interests are opposite, it is called a *conflict*. The outcome of every interaction is some change of each player's state.

From the variety of game types we pick out few "typical", "classical" ones, which we formalize and analyze for the purpose of to elucidate how the player must act in some game, — what decisions she/he is to make, — so that the outcome for her/him is "the best". To be more precise, we consider the games

• with 2 players: first P1 and second P2;

• with a single "move", at which each player chooses one of the actions available to her/him;

• with the number of such actions being finite for each player: $m > 1$ to P1 (enumerated $1, ..., m$) and $n > 1$ to P2 ($1, ..., n$);

• the pair of chosen actions then determines the outcome $(w_1; w_2)$, $w_i \in \mathbb{R}$ being the payoff of P-i;

• $w_1 + w_2 = 0$ (*zero-sum*, or *antagonistic* game), so it suffices to know one of them, let it be w_1 for the sake of definiteness. In the context of "capital games" it means that the constant total capital is only redistributed between players;

• each player knows w_1 for each pair of their actions (*full information* game).

Imagine yourself being P1, while P2 is the adversary. What is your goal, what do you seek?

We can call a *strategy* of a given player either the actions she/he chooses, or the way she/he chooses them; in particular, below we consider the so-called mixed strategies, where action is a RV with distribution on the set of possible actions.

Anyway, our goal is to determine the optimal strategy (OS) of a given, for the sake of definiteness let it be the 1st, player, but the meaning of "optimality" depends on what we call a strategy.

The conditions just given do not determine the situation completely. For example, do players make their decision simultaneously? If not, does the player who decides later know the choice of the other one (if so, it seems obvious that the one who decides later has an "advantage"). Does each player really aim to act optimally for her/himself, and does the other player know about it? Different answers to these questions can lead to different player strategies. We suppose that

• any player, at the moment of her/his own choice, does not know what choice the other player made/makes/will make. However, a player can "put her/himself in the other player's place", that is, know the strategy of the other player.

Point out other ambiguities that retain under these conditions.

2.1. Matrix games and saddle points

Such game is described by a *game matrix* (GM) $A = \|a_{ij}\|_{\substack{1\leqslant i\leqslant m,\\ 1\leqslant j\leqslant n}}$, where a_{ij} is the payoff of P1 if the choices of the players are $(i; j)$ (P1 chose her/his i-th action and P2 chose her/his j-th); at that, obviously, the payoff of P2 is $-a_{ij}$ (or, her/his loss is a_{ij}). In other words, P1 chooses the row of A, and P2 chooses the column; the number at their intersection is the gain of P1 and the loss of P2.

EX. 1. Consider two games: $A_1 = \begin{pmatrix} +3 & +1 \\ 0 & -1 \end{pmatrix}$ and $A_2 = \begin{pmatrix} +1 & -1 \\ -1 & +1 \end{pmatrix}$.

1) Here P1 can reason as follows: "whatever choice-column P2 does, I gain more by choosing the 1st row"; to be more precise, P1 can guarantee to her/himself the payoff no less than 1. In turn, P2 (from whose point of view the GM is $-A_i$) thinks: "whatever choice-row P1 does, I gain more by choosing the 2nd column", that is, P2 can guarantee to her/himself the payoff no less than -1/loss no more than 1. Therefore we suggest that the optimal strategy for P1 is 1 (1st row), and for P2 it is 2 (2nd column).

Does some player have a strategy that, in certain sense, is better than these?

2) Here similar reasonings do not work, because for each player the choice of "better" action depends on what the other player does. Besides, unlike (1), it is not advantageous for any player to choose certain action, because then the adversary, "having put her/himself in this player's place", knows this action and makes a choice guaranteeing to this player the loss $+1$.

However, each player must choose something. How e.g. P1 should act?

This Ex. indicates that, in general case, a given player may not have the single "right" action. The question arises, when such action — or, precisely, the pair of actions for P1 and P2 — exists.

REM. In matrix games where all a_{ij} have the same sign, the player whose payoffs are all negative reasons: "I lose for sure, at least I try to minimize my loss".

Having chosen i, P1 gets payoff $\geqslant \min\limits_j a_{ij}$. Knowing it, P1 chooses i to make this value maximal, that is, to get payoff $\geqslant \alpha = \max\limits_i \min\limits_j a_{ij}$. For this purpose she/he chooses i_0 at which $\max\limits_i$ is attained (if there are several such i, any of them). At that P2 gets payoff no more than $-\alpha$, that is, loses no less than α.

Reasoning analogously, P2 can guarantee to her/himself the payoff no less than $\max\limits_j \min\limits_i(-a_{ij}) = \max\limits_j\bigl(-\max\limits_i a_{ij}\bigr) = -\min\limits_j \max\limits_i a_{ij} = -\beta$, choosing j_0 where $\min\limits_j$ is attained. That is, P2 can guarantee to her/himself the loss $\leqslant \beta$.

DEF. α is *lower value/maximin* of the game; β is *upper value/minimax* of the game.

Action i_0 of P1 is called a *maximin strategy*, and the action j_0 of P2 — a *minimax strategy* (when a strategy is an action chosen by a player).

Suppose that

$$\max_i \min_j a_{ij} = \alpha \boxed{=} \beta = \min_j \max_i a_{ij} \qquad (*)$$

Then P1 knows, firstly, that she/he can gain no less than α, and secondly, that P2 can prevent her/him from gaining more than α. Adding the assumption that each player aims to act in the best, for her/himself, way, the best action for P1 is the action i_0. Analogously it is the best for P2 to choose the action j_0.

$(*)$ may not hold, — see, for instance, the matrix A_2 from Ex. 1: $\alpha = -1$, $\beta = 1$. Thus all matrix games are split into 2 types: those where $(*)$ holds, — then the optimal strategies for P1 and P2 are $(i_0; j_0)$ described above, — and those where it does not hold, — we shall consider such games later.

We consider the notions and properties related to $(*)$ in the more general case of 2-variable functions.

PROP. 1. Let $f\colon A \times B \to \mathbb{R}$, $s = \sup\limits_{x\in A} \inf\limits_{y\in B} f(x;y)$, $S = \inf\limits_{y\in B} \sup\limits_{x\in A} f(x;y)$. Then $s \leqslant S$.

◄ $\forall (x;y) \in A \times B$: $g(x) = \inf\limits_{y\in B} f(x;y) \leqslant f(x;y) \leqslant \sup\limits_{x\in A} f(x;y) = h(y)$. So, $s = \sup\limits_{x\in A} g(x) \leqslant h(y)$, thus $s \leqslant \inf\limits_{y\in B} h(y) = S$. ►

COR. 1. If, under cond. of Prop. 1, $\exists \alpha = \max\limits_{x\in A} \min\limits_{y\in B} f(x;y)$ and $\exists \beta = \min\limits_{y\in B} \max\limits_{x\in A} f(x;y)$, then $\alpha \leqslant \beta$.

COR. 2. Since, in GM A, coefs. $a_{ij} = f(i;j) \in \mathbb{R}$, where $i \in \{1; ...; m\}$, $j \in \{1; ...; n\}$, — domains of definition of both arguments are finite, — conditions of Cor. 1 hold, hence $\max\limits_i \min\limits_j a_{ij} \leqslant \min\limits_j \max\limits_i a_{ij}$.

DEF. Let $f\colon A \times B \to \mathbb{R}$. P. $(x_0; y_0) \in A \times B$ is called a *saddle point* of $f(x;y)$ if $\forall (x;y) \in A \times B$: $f(x;y_0) \leqslant f(x_0;y_0) \leqslant f(x_0;y)$.

The «classical» example is $f\colon \mathbb{R}^2 \to \mathbb{R}$: $f(x;y) = y^2 - x^2$ with saddle p. $(0;0)$, because $\forall x, y \in \mathbb{R}$: $-x^2 = f(x;0) \leqslant f(0;0) \leqslant f(0;y) = y^2$. We remark that $f_2(x;y) = x^2 - y^2$ does not have a saddle point, although «from differential geometry's POV» the surface $z = f_2(x;y)$ is a hyperbolic paraboloid as well.

THEOR. (on saddle p.) Let $f\colon A \times B \to \mathbb{R}$ and $\exists \alpha = \max\limits_{x\in A} \min\limits_{y\in B} f(x;y)$, $\exists \beta = \min\limits_{y\in B} \max\limits_{x\in A} f(x;y)$.

$\alpha = \beta$ **iff** $f(\cdot)$ has a saddle p. $(x_0; y_0)$, at that $\alpha = f(x_0; y_0) = \beta$.

◀ **Sufficiency.** Let $(x_0; y_0)$ be a saddle p. of $f(\cdot)$. We have the succession of inequalities $\beta = \min\limits_{y\in B} \max\limits_{x\in A} f(x;y) \leqslant \max\limits_{x\in A} f(x;y_0) \leqslant f(x_0;y_0) \leqslant$

$$\leqslant \min_{y\in B} f(x_0;y) \leqslant \max_{x\in A} \min_{y\in B} f(x;y) = \alpha$$

and due to Cor. 1 from Prop. 1, $\alpha \leqslant \beta$. Thus $\alpha = \beta$ $(= f(x_0; y_0))$.

Necessity. Let $\alpha = \beta$. We take $x_0 \in A$ and $y_0 \in B$ such that $\min\limits_{y\in B} f(x_0;y) = \alpha$ and $\max\limits_{x\in A} f(x;y_0) = \beta$ (it is possible, because corresponding $\max\limits_{x}$ for α and $\min\limits_{y}$ for β are attained). Hence $\min\limits_{y\in B} f(x_0;y) = \max\limits_{x\in A} f(x;y_0)$.

It follows that $\min\limits_{y\in B} f(x_0;y) \geqslant f(x_0;y_0)$, so $\forall y \in B$: $f(x_0;y_0) \leqslant f(x_0;y)$; analogously $\max\limits_{x\in A} f(x;y_0) \leqslant f(x_0;y_0) \Rightarrow \forall x \in A$: $f(x;y_0) \leqslant f(x_0;y_0)$. By def. $(x_0;y_0)$ is a saddle p. of $f(\cdot)$.

Since $\alpha = \min\limits_{y\in B} f(x_0;y) \leqslant f(x_0;y_0) \leqslant \max\limits_{x\in A} f(x;y_0) = \beta$ and $\alpha = \beta$, we have $f(x_0;y_0) = \alpha = \beta$. ▶

DEF. A *saddle p. of m-x* $A = \|a_{ij}\|$ is a pair $(i_0; j_0)$ such that a_{i_0,j_0} is
1) the smallest value in i_0-th row, $a_{i_0,j_0} \leqslant a_{i_0,j}$, and
2) the largest value in j_0-th column, $a_{i,j_0} \leqslant a_{i_0,j_0}$.

COR. MG A satisfies $(*)$ **iff** GM A has a saddle p. $(i_0; j_0)$, at that

$$\alpha = \beta = a_{i_0,j_0}$$

2.2. Mixed strategies

EX. 2. Recall the game $A_2 = \begin{pmatrix} +1 & -1 \\ -1 & +1 \end{pmatrix}$ from Ex. 1. There is no saddle p. here, so there is no single optimal action for any player as well. Perhaps it exists after all, but is determined in some way that is unfathomable to us, that is, at least one player is able, after all, to choose some action, the "best" to her/him?

However, each player can think as follows: "if I choose a concrete action, it will be my strategy, and since the adversary can put her/himself in my place, she/he will know this strategy, my action, and will choose her/his to win 1, then I will lose 1". Thus it is not profitable, for any player in such game, to act in a single predictable way, determinably. Therefore, they can act *non*-determinably, randomly. At that their payoffs become random variables as well.

Suppose that P1 chooses an action accordingly to the distribution $\begin{pmatrix} 1 & 2 \\ p & 1-p \end{pmatrix}$, and P2 knows this distribution, i.e. $p \in [0;1]$. Consider possible cases:

1) $p < \frac{1}{2}$. It is more probable that P1 chooses 2, thus it is more profitable for P2 to choose 1. Payoff of P1 is the RV $\xi \sim \begin{pmatrix} -1 & 1 \\ 1-p & p \end{pmatrix}$. Mean payoff $\mathrm{E}\,\xi = p - 1 + p = 2p - 1 < 0$, — on the average P1 loses.

2) At $p > \frac{1}{2}$ P2 chooses 2, $\xi \sim \begin{pmatrix} -1 & 1 \\ p & 1-p \end{pmatrix}$, $\mathrm{E}\,\xi = 1 - 2p < 0$ again.

3) If $p = \frac{1}{2}$, P2 can choose either 1 or 2; the probability of her/him gaining 1 will be $\frac{1}{2}$. P2 can also choose her/his action randomly with the same distribution, then the payoff of P1 is $\xi \sim \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$, and $\mathrm{E}\,\xi = 0$.

So P1, aiming to maximize her/his mean payoff, must act accordingly to (3).

To distinguish individual actions and random choice of the individual action in the MG $A = \|a_{ij}\|_{\substack{1 \leqslant i \leqslant m, \\ 1 \leqslant j \leqslant n,}}$

DEF. The actions available to a given player are her/his *pure strategies* (PS).

DEF. A *mixed strategy* (MS) of a given player is a probability distribution on the set of her/his PSs.

Let the set of MSs of P1 be $S_1 = \left\{(x_1; ...; x_m) \mid x_i \geqslant 0, \sum\limits_{i=1}^{m} x_i = 1\right\}$, analogously the set of MSs of P2 be $S_2 = \left\{(y_1; ...; y_n) \mid y_j \geqslant 0, \sum\limits_{j=1}^{n} y_j = 1\right\}$. We also denote $X = (x_1; ...; x_m)$, $Y = (y_1; ...; y_n)$.

PS is a "degenerate" case of MS: PS i of P1 is a MS with $x_k = \delta_{ik}$, and PS j of P2 is a MS with $y_k = \delta_{jk}$.

In addition to preceding assumptions we assume that

- each player knows adversarial MS, but does not know what PS adversary chooses, realizing her/his MS;
- RVs that are realizations of MSs of both players are independent.

If P1 applies the MS $X = \{x_i\}$ and P2 applies the MS $Y = \{y_j\}$, then the payoff of P1 is the RV ξ, $P\{\xi = a_{ij}\} = x_i y_j$. Therefore the mean payoff of P1 is

$$W(X,Y) := \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} y_j x_i = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij} x_i y_j$$

DEF. If $\exists X^* \in S_1$ and $\exists Y^* \in S_2$ such that

$$\forall X \in S_1,\ \forall Y \in S_2\colon W(X,Y^*) \leqslant W(X^*,Y^*) \leqslant W(X^*,Y)$$

then X^* and Y^* are called *optimal mixed strategies* (OMSs) of P1 and P2 accordingly, or *game solution* (this pair solves the problem "how to play optimally?"), and $w = W(X^*,Y^*)$ is called *game value*.

If P1 follows X^* and P2 follows Y^*, then, on the average, each of them gains no less than if she/he alone plays accordingly with some different MS.

Such pair is a particular case of a saddle p. with $f(\cdot) = W(\cdot)$ ($(X^*;Y^*)$ is a "saddle strategy of mean payoff"). Shall we split all MGs (or at least MGs without "usual" saddle p.) into 1) the ones where such pair of MSs exists, — then it is the best for the players to act accordingly with them, — and 2) the rest, where it does not, — then the players would have to act "even more complicatedly"? It appears that "complications" end at the games with the optimal pair of MSs.

PROP. 2. If $A \subset \mathbb{R}^n$ is finite, then $\operatorname{conv} A$ is closed.

. . . .

◀ Let $A = \{x^{(k)}\}_{k=1}^m$ and $C = \operatorname{conv} A$. By Prop. 6 from Ch. 1 $C = \Big\{\sum\limits_{k=1}^m \lambda_k x^{(k)} \mid \lambda_k \geqslant 0,\ \sum\limits_{k=1}^m \lambda_k = 1\Big\}$. Take $\forall y \notin C$. The set $S = \Big\{\lambda \in \mathbb{R}_+^m \mid \sum\limits_{k=1}^m \lambda_k = 1\Big\}$ — the surface of the unit simplex of $\mathbb{R}^m$ — is compact, implying that the continuous function $d\colon S \to \mathbb{R}_+\colon d(\lambda) = \|y - \sum\limits_{k=1}^m \lambda_k x^{(k)}\|$ attains on S its min value $\mu = \rho(y;C)$ for some $\lambda^{(0)}$. $\mu > 0$, otherwise $y = \sum\limits_{k=1}^m \lambda_k^{(0)} x^{(k)} \in C$, which contradicts the choice of y. Hence $\rho(y;C) > 0$, and for some $\delta > 0$, e.g. $\delta = \rho(y;C)$: $B(y;\delta) \cap C = \varnothing$, which means $y \in \operatorname{Int}(\mathbb{R}^n \backslash C)$.

In that way, $\mathbb{R}^n \backslash C = \operatorname{Int}(\mathbb{R}^n \backslash C)$ is open, thus C is closed. ▶

. . . .

LEMMA. Let $A = \|a_{ij}\|_{\substack{1\leqslant i\leqslant m,\\ 1\leqslant j\leqslant n,}}\ a_{ij} \in \mathbb{R}$.

Then at least one of the following statements holds:

1) $\exists X \in S_1\colon \sum\limits_{i=1}^m a_{ij}x_i \geqslant 0,\ j = \overline{1,n}$ ($\Leftrightarrow XA \geqslant \theta_n$);

2) $\exists Y \in S_2\colon \sum\limits_{j=1}^n a_{ij}y_j \leqslant 0,\ i = \overline{1,m}$ ($\Leftrightarrow AY^T \leqslant \theta_m$).

◀ Let $\{e^{(i)}\}_{i=1}^m$ be a standard base of $\mathbb{R}^m$, that is, $e_k^{(i)} = \delta_{ik}$, and $\{a^{(j)}\}_{j=1}^n = \{(a_{1j};...;a_{mj})\}_{j=1}^n$ — $\mathbb{R}^m$ points, whose coords. are coefs. of j-th column of A, so $a_k^{(j)} = a_{kj}$. We denote $C = \operatorname{conv}\{e^{(1)},...,e^{(m)},a^{(1)},...,a^{(n)}\}$.

Consider 2 possible cases:

a) $\theta_m \in C$. Then, by Prop. 6 from Ch. 1,

$$\theta_m = \sum_{i=1}^{m} u_i e^{(i)} + \sum_{j=1}^{n} v_j a^{(j)};\ u_i, v_j \geqslant 0,\ \sum_{i=1}^{m} u_i + \sum_{j=1}^{n} v_j = 1$$

Per coordinate

$$\sum_{i=1}^{m} u_i \delta_{ik} + \sum_{j=1}^{n} v_j a_{kj} = 0,\ k = \overline{1,m} \Leftrightarrow \sum_{j=1}^{n} v_j a_{kj} = -u_k,\ k = \overline{1,m}$$

Assume that all $v_j = 0$. Then these equalities imply that all $u_k = 0$ as well, and this contradicts $\sum_i u_i + \sum_j v_j = 1$. Thus there are positive ones among v_j, therefore $v = \sum_j v_j > 0$. Divide obtained equalities by v and let $y_j = \frac{v_j}{v}$. Obviously, $y_j \geqslant 0$ and $\sum_{j=1}^{n} y_j = \frac{v}{v} = 1$, so the divided equalities provide (2):

$$\sum_{j=1}^{n} a_{kj} y_j = -\frac{u_k}{v} \leqslant 0,\ k = \overline{1,m}$$

b) $\theta_m \notin C$. Using Minkowski Theor., we strongly separate θ_m and closed (by Prop. 2) convex C by some hyperplane in $\mathbb{R}^m$. Shift this plane in the direction of θ_m (or take $\alpha = \alpha_0$ instead of $\alpha = \alpha_0 + \frac{1}{2}d^2$ from the proof of Minkowski Theor.), and we have the plane H such that $\theta_m \in H$ and C is contained in the interior of one of semispaces H divides $\mathbb{R}^m$ into. Let the equation of this plane for $z \in \mathbb{R}^m$ be $\sum_{k=1}^{m} b_k z_k = b_{m+1}$ and $C \subseteq \operatorname{Int} H_+$ (if $C \subseteq \operatorname{Int} H_-$, we take, instead of $p = (b_1; ...; b_m)$, $-p$). Since $\theta_m \in H$, we have $b_{m+1} = 0$. Therefore, $\forall z \in C$: $\sum_{k=1}^{m} b_k z_k > 0$, in particular,

– for $z = e^{(i)}$, $i = \overline{1,m}$: $\sum_{k=1}^{m} b_k \delta_{ik} > 0 \Leftrightarrow b_i > 0$, $i = \overline{1,m}$. Then, obviously, $b = \sum_i b_i > 0$ as well;

– for $z = a^{(j)}$, $j = \overline{1,n}$: $\sum_{i=1}^{m} b_i a_{ij} > 0$. Divide these inequalities by b and let $x_i = \frac{b_i}{b}$; then $x_i \geqslant 0$, $\sum_{i=1}^{m} x_i = \frac{b}{b} = 1$, so we come to (slightly strengthened, $\geqslant$ becomes $>$) (1): $\sum_{i=1}^{m} a_{ij} x_i > 0$, $j = \overline{1,n}$. ▶

<u>THEOR.</u> (Main theorem of MG theory) Any MG has a pair of optimal MSs.

◀ Apply Saddle P. Theor. to $A = S_1$, $B = S_2$, and $f\colon S_1 \times S_2 \to \mathbb{R}$: $f(X;Y) = W(X,Y)$. We then have the criterion of existence of such p., that is, the optimal pair of MSs:

$$\exists\alpha = \max_{X \in S_1} \min_{Y \in S_2} W(X,Y), \qquad \exists\beta = \min_{Y \in S_2} \max_{X \in S_1} W(X,Y), \qquad \alpha = \beta$$

For a fixed $Y \in S_2$ the function $W(X,Y)$ is continuous (and linear) in X on the compact $S_1 \subset \mathbb{R}^m$, thus by Weierstrass Theor. $\forall Y \in S_2\ \exists \max\limits_{X \in S_1} W(X,Y)$. Besides, this $\max\limits_{X\in S_1} W(X,Y)$ is also a continuous (and piecewise linear) function in Y on the compact $S_2 \subset \mathbb{R}^n$. It follows that $\exists \beta$. Analogously $\exists \alpha$.

Take $\forall s \in \mathbb{R}$ and consider "MG with shifted payoffs" $A_s = \|a_{ij}^{(s)}\|_{i,j} = A - s = \|a_{ij} - s\|_{i,j}$; also, let $W_s(X,Y)$ be the mean payoff of P1 under MSs X, Y of MG A_s. It is easy to see that $W_s(X,Y) = \sum\limits_{i,j}(a_{ij} - s)x_i y_j = W(X,Y) - s$ and, by preceding reasonings,

$$\exists \alpha_s = \max_{X\in S_1} \min_{Y \in S_2} W_s(X,Y) = \alpha - s, \quad \exists \beta_s = \min_{Y\in S_2} \max_{X\in S_1} W_s(X,Y) = \beta - s$$

If (1) from the lemma holds, then, for the corresponding $X' \in S_1$, we have $\sum\limits_{i=1}^m a_{ij}^{(s)} x_i' \geqslant 0$, $j = \overline{1,n}$, hence $\forall Y \in S_2$: $W_s(X',Y) = \sum\limits_{j=1}^n \big[\sum\limits_{i=1}^m a_{ij}^{(s)} x_i'\big] y_j \geqslant 0 \Rightarrow g(X') = \min\limits_{Y\in S_2} W_s(X',Y) \geqslant 0 \Rightarrow \alpha_s = \max\limits_{X \in S_1} g(X) \geqslant 0$.

If (2) from the lemma holds, then similarly $\beta_s \leqslant \max\limits_{X\in S_1} W_s(X,Y') \leqslant 0$.

Now apply the lemma to A_s: at least one of (1)–(2) holds, so at least one of the inequalities $\alpha_s \geqslant 0$, $\beta_s \leqslant 0$ takes place. Therefore, $\alpha_s < 0 < \beta_s$ is impossible $\Leftrightarrow \alpha < s < \beta$ is impossible.

Thus, due to arbitrariness of s, we have $\alpha \geqslant \beta$. But, by Cor. 1 from Prop. 1, $\alpha \leqslant \beta$, so $\alpha = \beta$. ▶

2.3. Properties of optimal strategies

$W(i,Y) = \sum\limits_{j=1}^n a_{ij} y_j$ is the mean payoff of P1 if she/he uses PS i and P2 uses MS Y; $W(X,j) = \sum\limits_{i=1}^m a_{ij} x_i$ is the mean payoff of P1 if she/he uses MS X and P2 uses PS j. At that $W(X,Y) = \sum\limits_{i=1}^m W(i,Y) x_i = \sum\limits_{j=1}^n W(X,j) y_j$.

THEOR. 1. Let $w \in \mathbb{R}$, $X^* \in S_1$, $Y^* \in S_2$. w is the game value and X^*, Y^* are OSs of P1 and P2 accordingly **iff** $\forall (i;j)$: $W(i,Y^*) \leqslant w \leqslant W(X^*, j)$.

◀ **Necessity** follows from def. of a saddle strategy for X with $x_k = \delta_{ik}$ and Y with $y_k = \delta_{jk}$.

Sufficiency. $\forall X \in S_1$: $W(X,Y^*) = \sum\limits_{i=1}^m W(i,Y^*) x_i \leqslant \sum\limits_{i=1}^m w x_i = w$, similarly $\forall Y \in S_2$: $W(X^*,Y) \geqslant w$. When $X = X^*$ and $Y = Y^*$ we obtain $W(X^*,Y^*) \leqslant w \leqslant W(X^*,Y^*)$, thus $W(X^*,Y^*) = w$. So

$$\forall (X;Y) \in S_1 \times S_2\colon W(X,Y^*) \leqslant W(X^*,Y^*) \leqslant W(X^*,Y)$$

— $(X^*;Y^*)$ is a saddle point, and w is the game value. ▶

COR. If GM A has the saddle p. $(i_0; j_0)$, then PS i_0 is an optimal MS of P1 and j_0 is an OMS of P2.

◀ $\forall(i;j)$: $W(i,j_0) = a_{i,j_0} \leqslant a_{i_0,j_0} \leqslant a_{i_0,j} = W(i_0,j)$. Let $w = a_{i_0,j_0}$, — conditions of Theor. 1 hold. ▶

THEOR. 2. Let $(X^*;Y^*)$ be a solution of MG A. Then $\max\limits_{1\leqslant i\leqslant m} W(i,Y^*) = W(X^*,Y^*) = \min\limits_{1\leqslant j\leqslant n} W(X^*,j)$.

◀ We prove the 2nd equality, the 1st one is proved analogously. Let $w = W(X^*,Y^*)$. By Theor. 1, $\forall j$: $w \leqslant W(X^*,j)$. Therefore, $w \leqslant \min\limits_{1\leqslant j\leqslant n} W(X^*,j)$. If $w < \min\limits_{1\leqslant j\leqslant n} W(X^*,j)$, then $\forall j$: $w < W(X^*,j)$, implying

$$W(X^*,Y^*) = \sum_{j=1}^{n} W(X^*,j)y_j^* > w$$

(because $\exists y_j^* > 0$) — ⨂ Theor. 1. Thus $w = \min\limits_{1\leqslant j\leqslant m} W(X^*,j)$. ▶

THEOR. 3. Using the same denotations: $W(i,Y^*) < w \Rightarrow x_i^* = 0$ and $W(X^*,j) > w \Rightarrow y_j^* = 0$.

◀ Let us prove, for instance, the 1st implication. Let $W(i_0,Y^*) < w$. Assume that $x_{i_0}^* > 0$. Then $W(i_0,Y^*)x_{i_0}^* < wx_{i_0}^*$. For all other i, by Theor. 1, $W(i,Y^*) \leqslant w \Rightarrow W(i,Y^*)x_i^* \leqslant wx_i^*$. Therefore

$$W(X^*,Y^*) = \sum_{i=1}^{m} W(i,Y^*)x_i^* < \sum_{i=1}^{m} wx_i = w$$

— ⨂ Theor. 1. So the assumption was wrong and $x_{i_0}^* = 0$. ▶

2.4. Simplification of matrix games

DEF. Let $a = (a_1;...;a_n) \in \mathbb{R}^n$, $b = (b_1;...;b_n) \in \mathbb{R}^n$. We say that *a dominates* b, $a \geqslant b$, if $a_i \geqslant b_i$, $i = \overline{1,n}$. If all per-coordinate inequalities are strict, we say that *a strictly dominates* b, $a > b$.

DEF. Let $X = (x_1;...;x_m)$ be a MS of P1 in MG A. The *extension of X at i-th position* is $X' = (...;x_{i-1};0;x_i;...)$. The extension of MS $Y = (y_1;...;y_n)$ of P2 at j-th position is $Y' = (...;y_{j-1};0;y_j;...)$.

THEOR. Let $A = \|a_{ij}\|_{\substack{1\leqslant i\leqslant m,\\ 1\leqslant j\leqslant n}}$ be a GM where i-th row is dominated by some conv.comb. of other rows, and A' be the GM obtained by removing i-th row from A. Then

- the value of A' is equal to the value of A;
- an OS of P2 in A' is an OS of P2 in A;

• if X' if an OS of P1 in A', then its extension at i-th position X^* is an OS of P1 in A;

• if i-th row is *strictly* dominated by some conv.comb. of other rows, then *every* solution of A can be obtained in this way from some solution of A'.

....

◀ Let, for instance, the lowest, m-th row, be a dominated one. Thus $\exists\{\lambda_i\}_{i=1}^{m-1}$: $\lambda_i \geqslant 0$, $\sum\limits_{i=1}^{m} \lambda_i = 1$ and $\sum\limits_{i=1}^{m-1} \lambda_i a_{ij} \geqslant a_{mj}$, $j = \overline{1,n}$ (#)

Let in A': w be the value, $X = (x_1; ...; x_{m-1})$ be OS of P1, $Y = (y_1; ...; y_n)$ be OS of P2. By Theor. 1, for $i = \overline{1, m-1}$ and $j = \overline{1,n}$

$$\sum_{j=1}^{n} a_{ij}y_j = W'(i,Y) \leqslant w \leqslant W'(X,j) = \sum_{i=1}^{m-1} a_{ij}x_i \qquad (\#\#)$$

Let us show that in A for w, $X^* = (x_1; ...; x_{m-1}; 0)$, and Y sufficient conditions of Theor. 1 hold.

$$W(X^*,j) = \sum_{i=1}^{m} a_{ij}x_i^* = \sum_{i=1}^{m-1} a_{ij}x_i + a_{mj}\cdot 0 = W'(X,j) \geqslant w, \quad j = \overline{1,n}$$

is obviously contained in (##). Do inequalities $W(i,Y) = \sum\limits_{j=1}^{n} a_{ij}y_j \leqslant w$, $i = \overline{1,m}$, hold? For $i = \overline{1, m-1}$ they do, because $W(i,Y) = W'(i,Y) \leqslant w$ by (##) then; for $i = m$

$$W(m,Y) = \sum_{j=1}^{n} a_{mj}y_j \overset{(\#)}{\leqslant} \sum_{j=1}^{n}\Big[\sum_{i=1}^{m-1} \lambda_i a_{ij}\Big]y_j = \sum_{i=1}^{m-1}\Big[\sum_{j=1}^{n} a_{ij}y_j\Big]\lambda_i \overset{(\#\#)}{\leqslant} \sum_{i=1}^{m-1} w\lambda_i = w$$

Therefore, by Theor. 1, in A w is the value, X^* is an OS of P1, Y is an OS of P2 — first 3 statements have been proved. Now, let m-th row be strictly dominated, i.e. all $\geqslant \to >$ in (#). Then $\exists y_j > 0$ implies

$$W(m,Y) = \sum_{j=1}^{n} a_{mj}y_j \; < \; \sum_{j=1}^{n}\Big[\sum_{i=1}^{m-1} \lambda_i a_{ij}\Big]y_j = ... \leqslant w$$

so by Theor. 3 in A $\forall$ OS X^* of P1 has $x_m^* = 0$, i.e. X^* is the extension of some MS X of P1 in A'. ▶

....

<u>THEOR.</u> Let A be a GM where j-th column dominat<u>es</u> some conv.comb. of other columns, and A' be the GM obtained by removing j-th column from A. Then

• the value of A' is equal to the value of A;

• an OS of P1 in A' is an OS of P1 in A;

• if Y' is an OS of P2 in A', then its extension at j-th position Y^* is an OS of P2 in A;

• if j-th column *strictly* dominates some conv.comb. of other columns, then *every* solution of A can be obtained in this way from some solution of A'.

(The proof is analogous.)

2.5. Matrix Games and Linear Programming

It is evident (see the proof of Main Theor. of MG theory) that the addition of some constant C to all coefs. of GM A increases the game value w by C, and OSs of P1 and P2 do not change (all mean payoffs increase by C, the inequalities between them are retained). Take $C = 1 - \min\limits_{i,j} a_{ij}$, then the coefs. of new GM $A' = A + C$ are positive and the value of the game A' is positive. Thus, any GM can be transformed into GM A with such coefs. and such value w. Consider this transformed GM, we denote it A right away.

By Theor. 1, for $X = (x_1; ...; x_m) \in \mathbb{R}_+^m$ to be an OS of P1, it is necessary that

$$X \in S_1 \Rightarrow \sum_{i=1}^{m} x_i = 1; \quad W(X, j) = \sum_{i=1}^{m} a_{ij} x_i \geqslant w,\ j = \overline{1, n}$$

and, for $Y = (y_1; ...; y_n) \in \mathbb{R}_+^n$ to be an OS of P2, it is necessary that

$$Y \in S_2 \Rightarrow \sum_{j=1}^{n} y_j = 1; \quad W(i, Y) = \sum_{j=1}^{n} a_{ij} y_j \leqslant w,\ i = \overline{1, m}$$

Divide these conditions by w and denote $x_i' = x_i/w$, $y_j' = y_j/w$ to obtain

$$\sum_{i=1}^{m} x_i' = \tfrac{1}{w}; \quad \sum_{i=1}^{m} a_{ij} x_i' \geqslant 1,\ j = \overline{1, n}, \qquad \sum_{j=1}^{n} y_j' = \tfrac{1}{w}; \quad \sum_{j=1}^{n} a_{ij} y_j' \leqslant 1,\ i = \overline{1, m}$$

The goal of P1 is the choice of x_i, and thus x_i', such that her/his mean payoff is maximal. Since, for the pair of OSs, this payoff is equal to w, it can be interpreted as $w \to \max \Leftrightarrow \frac{1}{w} \to \min$. On the other hand, P2 aims to choose y_j' such that $w \to \min \Leftrightarrow \frac{1}{w} \to \max$. Hence we can rewrite these problems:

$$x_i' \geqslant 0,\ i = \overline{1, m}; \quad \sum_{i=1}^{m} a_{ij} x_i' \geqslant 1,\ j = \overline{1, n}; \quad \sum_{i=1}^{m} x_i' \to \min$$

$$y_j' \geqslant 0,\ j = \overline{1, n}; \quad \sum_{i=1}^{m} a_{ij} y_j' \leqslant 1,\ i = \overline{1, m}; \quad \sum_{j=1}^{n} y_j' \to \max$$

— obviously, these are LPPs, and symmetric dual at that. Solve them (for example, firstly one by means of SM, and the other as dual to it) to obtain the vectors $X' = \{x_i'\}$ and $Y' = \{y_j'\}$. After that determine the value of the game $w = \left[\sum_{i=1}^{m} x_i'\right]^{-1} = \left[\sum_{j=1}^{n} y_j'\right]^{-1}$ and the sought OSs of P1 and P2: $X = wX'$, $Y = wY'$.

<u>REM.</u> This approach to transforming MG problem to LPP is not unique.

2.6. Structure of solutions of a matrix game

We consider GM $A = \|a_{ij}\|_{\substack{1 \leqslant i \leqslant m, \\ 1 \leqslant j \leqslant n.}}$ We denote the set of OSs of P1 by $T_1 = T_1(A) \subset \mathbb{R}^m$, the set of OSs of P2 by $T_2 = T_2(A) \subset \mathbb{R}^n$. It turns out that every T_k is the convex hull of some finite set of points.

THEOR. 1. T_1 and T_2 are non-empty, closed, bounded, and convex sets.

COR. $T_k = \operatorname{conv} E(T_k)$. (By Minkowski's Theor. on convex compact set.)

THEOR. 2. Let the game value be w, $X \in T_1$, $Y \in T_2$ (so $w = W(X, Y)$). $X \in E(T_1)$ and $Y \in E(T_2)$ **iff** a nonsingular $r \times r$ submatrix B of matrix A exists such that for $q = I_r B^{-1} I_r^T$:

$$w = \frac{1}{q}, \qquad \widetilde{X} = \frac{1}{q} I_r B^{-1}, \qquad \widetilde{Y} = \frac{1}{q} I_r (B^{-1})^T$$

where $I_r = (1; 1; ...; 1) \in \mathbb{R}^r$, $\widetilde{X}$ is obtained by excluding from X the elements corresponding to the rows of A that are not in B, $\widetilde{Y}$ is obtained by excluding from Y the elements corresponding to the columns of A that are not in B.

COR. $E(T_1)$ and $E(T_2)$ are finite.

This and Cor. from Theor. 1 imply that T_k is a convex polyhedron whose vertices are $E(T_k)$.

These theorems and corollaries substantiate the following algorithm to find all solutions of a given MG:

1 Search through all square submatrices B of matrix A.

2 For each B: do conditions of Theor. 2 hold ($\det B \neq 0$)? No: reject this B.

3 Yes: calculate w, $\widetilde{X}$, $\widetilde{Y}$, and check: do conditions $\widetilde{x}_i \geqslant 0$, $\sum\limits_i \widetilde{x}_i = 1$, and $\widetilde{y}_j \geqslant 0$, $\sum\limits_j \widetilde{y}_j = 1$ hold? No: reject B.

4 Yes: complement $\widetilde{X}$ to $X \in S_1$ with zeros at positions of the rows removed during $A \to B$ transition and $\widetilde{Y}$ to $Y \in S_2$ with zeros at positions of the columns removed during the same transition. Are $X \in T_1$ and $Y \in T_2$? No: reject B.

5 Yes: besides $(X; Y)$ being a solution of MG at hand, by Theor. 2 we have $X \in E(T_1)$ and $Y \in E(T_2)$.

Search through all such matrices B to find $E(T_1)$ and $E(T_2)$. T_k consists of all conv.comb. of $E(T_k)$ points.

REM. We can omit 1×1 matrices B because they lead to solution only when GM has "usual" saddle point.

2.7. Bimatrix games

Many games, in particular those that are met with "in practice", are not zero-sum games. Such games are called *bimatrix* (BG) and, naturally, are described by a pair of matrices $G = (A; B) = (\|a_{ij}\|; \|b_{ij}\|)$ of the same dimensionality $m \times n$, which determine the payoffs of P1 and P2 accordingly: if P1 carries out i-th action (from available $1, ..., m$) and P2 carries out j-th action (from available $1, ..., n$), then P1's payoff is a_{ij} and P2's payoff is b_{ij}. "Usual" MGs are a particular case of BG, they are described by pairs $(A; -A)$.

At first let a "strategy" of each player mean the individual action that she/he carries out determinably.

As before, each player can know the strategy of the other one and aims to maximize her/his own payoff.

In BGs the analogue of a saddle point is the notion of an equilibrium:

DEF. $(i_0; j_0)$ is called an *equilibrium (situation)* (ES) or a *Nash equilibrium* (NE) of BG G, and strategies i_0, j_0 are called *equilibrial*, if

$$a_{i_0,j_0} = \max_{1 \leqslant i \leqslant m} a_{i,j_0}, \quad b_{i_0,j_0} = \max_{1 \leqslant j \leqslant n} b_{i_0,j}$$

When both players are at equilibrium, — each player acts accordingly to her/his own equilibrial strategy — it is not advantageous for *one* player to deviate to another strategy (her/his payoff will not increase).

If $B = -A$, i.e. BG is MG, then the def. of NE is equal to the def. of saddle point. The following examples demonstrate their common and different properties.

EX. 3. Let $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $B = -A = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$. Search through 4 possible $(i_0; j_0)$ to make sure that the game $(A; B)$ does not have NE. Thus, there are BGs without NE for individual actions, similar to MGs without saddle p.

EX. 4. Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$. There are 2 NEs in this BG: $(1; 1)$ (at that P1 gains 1, P2 gains 2) and $(2; 2)$ (P1 gains 2, P2 gains 1). But if players do not act in concord, then, aiming to $\rightarrow$ max her/his payoff, P1 most likely chooses action 2, and P2 chooses 1, so both "gain" 0. While in a MG, even if saddle p. is not unique, payoffs of player 1 are equal at all saddle points, $\alpha = \beta = a_{i_0,j_0}$.

BGs with several NEs are called *coordination games.* In such games players can act as follows: in the beginning they agree on the common NE (coordinate), then they stop any negotiations and act "autonomously", possibly deviating from equilibrial strategy (*non-coalitional games*). It is not advantageous for each player to deviate from the agreed equilibrial strategy if only she/he deviates.

EX. 5. The classical *Prisoner's Dilemma*: [criminals] Alice (A) and Bob (B) have been arrested on suspicion of bank-robbery and are interrogated separately. If they both confess, each will be sentenced to 5 years in prison. If neither confesses, they will be imprisoned too, but for shorter term, say, 1 year for illegal arms keeping. If one confesses and the other does not, then the one who confesses will be released for cooperation with the police and the other will be sentenced to 10 years in prison. Let "do not confess" be action 1, "confess" be action 2, then the payoff matrices are $A = \begin{pmatrix} -1 & -10 \\ 0 & -5 \end{pmatrix}$, $B = \begin{pmatrix} -1 & 0 \\ -10 & -5 \end{pmatrix}$.

This BG has unique NE $(2;2)$ — they both confess and spend 5 years in prison, although it is more advantageous for both to not confess; on the other hand, each one may "apprehend" that she/he does not confess while the other does, regardless of whatever preceding agreements they had.

DEF. The behaviour $(i_0; j_0)$ of the players in BG $(A; B)$ is called *Pareto efficient* if there is no behaviour $(i; j)$ such that $a_{ij} \geqslant a_{i_0,j_0}$, $b_{ij} \geqslant b_{i_0,j_0}$, and at least one inequality is strict.

In Ex. 5, NE $(2;2)$ is not Pareto efficient, because we can take $(1;1)$ as $(i;j)$.

EX. 6. Let $A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$.

The unique NE is $(1;1)$, then each player gains 1.

In a MG with saddle p., a player can act by the maximin strategy. In this BG P1 has $\max\limits_i \min\limits_j a_{ij} = \max\{-1; -2\} = -1$ attained at $i = 1$, but P2 has $\max\limits_j \min\limits_i b_{ij} = \max\{0; 1; 0\} = 1$ attained at $j = 2$. Hence, even if they have agreed on NE $(1;1)$, it is more "advantageous" for ill-meaning P2 to break the agreement, switch to her/his maximin strategy and do 2, then, if P1 does not break the agreement and does 1, P2 gains 1 as before... but P1 loses 1.

Thus we can point out the following drawbacks of NE notion in BG: it

• may not exist, • may be non-unique, • may be not the "best" for players.

Nevertheless, it is one of principal and practically useful notions that allow to make decisions in conflicts or explain why someone (players) made such decisions.

We formulate the conditions of NE's existence for payoff *functions* instead of payoff *matrices*; this generalization corresponds to transition from PSs to MSs.

THEOR. (Brauer, on fixed point) Let $A \subset \mathbb{R}^n$ be a convex compact set and $f\colon A \to A$ be continuous. Then $\exists x \in A$: $f(x) = x$, i.e. f has a fixed p.

DEF. Function $f\colon \mathbb{R}^n \to \mathbb{R}$ is called

$$\begin{Bmatrix} \textit{convex} \\ \textit{concave} \end{Bmatrix} \text{ if } \forall x, y \in \mathbb{R}^n,\ \forall \lambda \in [0;1]\colon f(\lambda x + (1-\lambda)y) \begin{Bmatrix} \leqslant \\ \geqslant \end{Bmatrix} \lambda f(x) + (1-\lambda) f(y)$$

THEOR. (On NE's existence) Let $K_1 \subset \mathbb{R}^m$ and $K_2 \subset \mathbb{R}^n$ be convex compact sets and bifunction game G be defined by payoff funcs $(f_1; f_2)$, where $f_k\colon K_1 \times$

$K_2 \to \mathbb{R}$ are cont., at that $f_1(x;y)$ for any fixed y is concave in x, and $f_2(x;y)$ for any fixed x is concave in y. Then G has NE, i.e. $\exists(x^*;y^*) \in K_1 \times K_2$:

$$\forall(x;y) \in K_1 \times K_2\colon\ f_1(x;y^*) \leqslant f_1(x^*;y^*),\ f_2(x^*;y) \leqslant f_2(x^*;y^*)$$

....

◀ For $f\colon A \to \mathbb{R}$ we denote $\arg\max\limits_{x\in A} f(x) := \{y \in A \mid f(y) = \max\limits_{x\in A} f(x)\}$.

1) Let, at first, the concavities of f_1 and f_2 be strict. Then $\forall(x;y) \in K_1 \times K_2$ the sets $X(y) = \arg\max\limits_{x\in K_1} f_1(x;y)$ and $Y(x) = \arg\max\limits_{y\in K_2} f_2(x;y)$ consist of 1 element each, $\widehat{x}(y)$ and $\widehat{y}(x)$, at that the functions $\widehat{x}(\cdot)$ and $\widehat{y}(\cdot)$ are cont. in y and x accordingly. They are called the *best response functions* (BRF) of P1 and P2.

We define $f\colon K_1 \times K_2$: $f(x;y) = (\widehat{x}(y);\widehat{y}(x))$. Since $K_1 \times K_2 \subset \mathbb{R}^m \times \mathbb{R}^n$ is a convex compact set as well, by Brauer Theor. $\exists(x^*;y^*)$: $f(x^*;y^*) = (x^*;y^*)$, i.e. $\widehat{x}(y^*) = x^*$, $\widehat{y}(x^*) = y^*$. This means that $\forall x \in K_1$: $f_1(x^*;y^*) \geqslant f_1(x;y^*)$ and $\forall y \in K_2$: $f_2(x^*;y^*) \geqslant f_2(x^*;y)$; by def., $(x^*;y^*)$ is NE.

2) Now we consider the general case. For $\forall\varepsilon > 0$ the functions

$$f_1^{(\varepsilon)}(x;y) := f_1(x;y) - \varepsilon\|x\|^2, \qquad f_2^{(\varepsilon)}(x;y) := f_2(x;y) - \varepsilon\|y\|^2$$

are cont. (in x, y, and ε), $f_1^{(\varepsilon)}$ is strictly concave in x and $f_2^{(\varepsilon)}$ is strictly concave in y. In BG $(f_1^{(\varepsilon)};f_2^{(\varepsilon)})$ due to (1) $\exists$NE $(x^{(\varepsilon)};y^{(\varepsilon)})$. Take $\{\varepsilon_k\}_{k\in\mathbb{N}}$ such that $\varepsilon_k \xrightarrow[k\to\infty]{} 0$ and $\exists \lim\limits_{k\to\infty}(x^{(\varepsilon_k)};y^{(\varepsilon_k)}) = (x';y')$. In the inequalities

$$\forall x \in K_1\colon\ f_1^{(\varepsilon_k)}(x^{(\varepsilon_k)};y^{(\varepsilon_k)}) \geqslant f_1^{(\varepsilon_k)}(x;y^{(\varepsilon_k)})$$
$$\forall y \in K_2\colon\ f_2^{(\varepsilon_k)}(x^{(\varepsilon_k)};y^{(\varepsilon_k)}) \geqslant f_2^{(\varepsilon_k)}(x^{(\varepsilon_k)};y)$$

as $k \to \infty$, we obtain, due to continuity, that $(x';y')$ is NE in BG $(f_1;f_2)$. ▶

....

Now we consider mixed strategies in BG that is described by the pair of $m\times n$ matrices $(A;B)$. As in MG case, MS of a given player is some prob. distrib. on the set of her/his PSs, which may be known to the other player. The collections of all MSs are denoted, as before, S_1 for P1 and S_2 for P2.

If P1 applies MS $X = (x_1;...;x_m)$ and, independently, P2 applies MS $Y = (y_1;...;y_n)$, then the mean payoff of P1 is $W_1(X,Y) = \sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n} a_{ij}x_iy_j$, and the mean payoff of P2 is $W_2(X,Y) = \sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n} b_{ij}x_iy_j$.

(This BG is determined either by the pair $(A;B)$, or by the pair $(W_1;W_2)$.)

$S_1 \subset \mathbb{R}^m$ and $S_2 \subset \mathbb{R}^n$ are convex compact sets, and the functions W_1, W_2 are bilinear, thus continuous and concave in x for any fixed y and in y for any fixed x accordingly, hence, by Theor., this BG has NE $(X^*;Y^*)$ in relation to mean payoffs — $\forall(X;Y) \in S_1 \times S_2$: $W_1(X,Y^*) \leqslant W_1(X^*,Y^*)$, $W_2(X^*,Y) \leqslant W_2(X^*,Y^*)$.

To be more precise, this NE is called *NE in MSs*; thus we have the analogue of the Main Theor. of MG theory — every BG has NE in MSs.

The following properties are similar to the properties of OMS in MG.

THEOR. 1. The pair of MSs $(X^*; Y^*)$ in BG G is NE **iff**
$\forall (i; j)$: $W_1(i, Y^*) \leqslant W_1(X^*, Y^*)$, $W_2(X^*, j) \leqslant W_2(X^*, Y^*)$

THEOR. 2. (*Complementary non-rigidity* property) Let $(X^*; Y^*)$ be NE in MSs in BG G. Then
$x_i^* > 0 \Rightarrow W_1(i, Y^*) = W_1(X^*, Y^*)$ and $y_j^* > 0 \Rightarrow W_2(X^*, j) = W_2(X^*, Y^*)$

COR. Under these conditions, $W_1(i, Y^*) < W_1(X^*, Y^*) \Rightarrow x_i^* = 0$ and $W_2(X^*, j) < W_2(X^*, Y^*) \Rightarrow y_j^* = 0$.

THEOR. 3. $(X^*; Y^*)$ is NE in MSs in BG G **iff** $\exists I \subseteq \{1; ...; m\}, \exists J \subseteq \{1; ...; n\}, \exists w_1 \in \mathbb{R}, \exists w_2 \in \mathbb{R}$:

$$\sum_{j \in J} y_j^* = 1; \quad \sum_{j \in J} a_{ij} y_j^* = w_1,\ i \in I; \quad \sum_{j \in J} a_{ij} y_j^* \leqslant w_1,\ i \notin I \qquad (\#)$$

$$\sum_{i \in I} x_i^* = 1; \quad \sum_{i \in I} b_{ij} x_i^* = w_2,\ j \in J; \quad \sum_{i \in I} b_{ij} x_i^* \leqslant w_2,\ j \notin J \qquad (\#\#)$$

This Theor. substantiates the following algorithm to find NE in MSs:

Search through all $I \subseteq \{1; ...; m\}$, $J \subseteq \{1; ...; n\}$ such that $|I| = |J|$ and consider corresponding square submatrices $A' = \|a_{ij}\|_{i \in I, j \in J}$, $B' = \|b_{ij}\|_{i \in I, j \in J}$. For each $(A'; B')$ solve the set of equations from (#)–(##) (w_k are unknowns as well) and verify whether the solutions satisfy the inequalities. If so, add to these solutions $x_i^* = 0$ for $i \notin I$, $y_j^* = 0$ for $j \notin J$, and obtain NE $(X^*; Y^*)$.

DEF. NE is *completely mixed* if, in its MSs, all PSs have probabilities > 0.

2.8. Positional games

We have looked into the games with a single "move" that the players make "simultaneously" and that determines the final outcome of the game, their payoffs. Consider the *positional games* (PG), where players make several moves and get payoffs in the end. Each move changes the situation in which the players are and which is called a *position*. Let there be n players; $n = 2$ is a particular case.

To get a mathematical model of such game, we describe positions and transitions between them by a finite directed tree-like graph (into each vertex, except for the root, exactly 1 edge-arrow leads, no edges lead into the root), a *game tree*, where vertices denote positions and edges-arrows denote players' moves. *k-th level* consists of all vertices such that the length of the path from the root to them is k; since there is a unique path from the root to each vertex, this definition is proper.

The PG tree (PGT) must satisfy the following constraints:

• The root of the tree is the initial position of the game.

• End/exit vertices (leaves) are end positions of the game. To each end vertex, the set $(w_1; ...; w_n)$ corresponds, where w_i is the payoff of i-th player.

• Every path from the root to end vertex describes some concrete possible game, its "history".

• All non-end vertices are split into *priority sets* (PSs) S_1, ..., S_n, and S_0, where S_i, for $i = \overline{1, n}$, consists of positions where i-th player makes a move, and S_0 consists of positions where the move is made randomly, that is, is chosen accordingly to some known distribution from possible ones.

• Each PS S_i is split into *information sets* (ISs) S_{ij}, where S_{ij} is the set of positions, including every one's past, that i-th player can not distinguish (she/he does not have information to distinguish them).

• To each IS S_{ij}, the set I_{ij} corresponds, consisting of the indices ("names") of moves that i-th player can make in each position from S_{ij} (we remark that moves that have the same index can lead to different positions of the next level from different positions of a given IS).

REM. We model PG by trees, because, in a given position, the player's move and its consequences depend, in general case, not only on this position, but also on its "past", that is, all preceding moves of all players. Any move makes impossible the movement along other branches that grow from current position.

DEF. PG is called a *full information game* if every IS consists of 1 position.

E.g. tic-tac-toe, checkers, chess. A card game where any player does not know precisely what cards other players have is a PG without full information.

DEF. A *strategy* of i-th player is a function that maps each IS S_{ij} to certain index from I_{ij} — the move that she/he makes in any position from S_{ij}. We denote the collection of all her/his strategies by $\mathfrak{S}_i$.

REM. We assume that every player completely determined her/his strategy before the game, i.e. decided beforehand what move to make in every IS S_{ij}. Then, during the game itself, players act "automatically", and the only sources of outcome indeterminacy are random initial position and random moves.

DEF. The *normal form* (NF) of a PG is the table that determines the function $\pi\colon \mathfrak{S}_1 \times ... \times \mathfrak{S}_n \to \mathbb{R}^n$, which maps every set $(\mathfrak{s}_1; ...; \mathfrak{s}_n)$ of players' strategies to the set $(\overline{w}_1; ...; \overline{w}_n)$ of their *mean* payoffs if they apply these strategies.

When $n = 2$, we can write NF as two tables of P1's and P2's payoffs, where rows are strategies of P1 and columns are strategies of P2. That is, in this case we represent PG as BG; if always $\overline{w}_2 = -\overline{w}_1$, then as MG.

3. Optimal Statistical Decisions

Let the model of some random experiment (RE) be a probability space (PS) $\{\Omega; \mathfrak{F}; P\}$. A performance of this RE determines its outcome, identified to $\omega \in \Omega$, which is also called a *state* (of system/nature).

Statistician does not know the outcome ω and makes the decision/applies the action $d \in D$, the consequence U of it being a scalar (or a vector), which depends also on ω, i.e. $U = U(\omega; d)$. U is called a *utility function* / *profit* / *reward* / *payoff* etc. At first we consider scalar payoffs, $U\colon \Omega \times D \to \mathbb{R}$.

Statistician aims to make the payoff as large as possible, or, to be more precise, due to U as a function of ω being RandVar, — to make a decision d that maximizes the mean payoff $\mathrm{E}\, U = \int\limits_{\Omega} U(\omega; d) P(d\omega)$.

In Decision Making Problems (DMP) we usually do not consider a profit U, instead we consider a *loss* $L(\omega; d) = -U(\omega; d)$, the mean of which we aim to minimize, correspondingly. As a rule, $L \geqslant 0$, and it increases as the difference between ω and d increases.

For these definitions to be proper, we require the following condition to hold:
• $\forall d \in D$: $L(\omega; d)\colon \Omega \to \mathbb{R}$ is measurable w.r.t. $\mathfrak{F}$ / relative to $\mathfrak{F}$ (that is, $L^{-1}(\mathfrak{B}^1) \subseteq \mathfrak{F}$) and is integrable w.r.t. $P(\cdot)$ / $P(\cdot)$-integrable.

DEF. $\rho(P; d) = \mathrm{E}\, L = \int\limits_{\Omega} L(\omega; d) P(d\omega)$ is called a *risk*.

A solution of a DMP is a decision/action that minimizes the risk — an *optimal decision* (OD).

3.1. Bayes Risk, Bayes Decisions

DEF. $\rho^*(P) = \inf\limits_{d \in D} \rho(P; d)$ is called the *Bayes Risk* (BR).

DEF. $d^* \in D$ such that $\rho(P; d^*) = \rho^*(P)$ is called a *Bayes Decision* (BD).

In general, BD may not exist (inf is not attained). Below we consider the DMPs where it exists (in practice, it is possible to make the decision whose risk is as close to BR as we want). In particular, it is always the case when D is finite.

EX. 1. $\Omega = \{\omega_1; \omega_2; \omega_3; \omega_4\}$, $P = \begin{pmatrix} 1 & 2 & 3 & 4 \\ \frac{5}{10} & \frac{2}{10} & \frac{1}{10} & \frac{2}{10} \end{pmatrix}$, $D = \{d_1; d_2; d_3\}$, L is defined by the table

	d_1	d_2	d_3
ω_1	1	0	3
ω_2	3	2	0
ω_3	0	4	3
ω_4	2	1	1

Let us find BR and BD. We calculate the risk of each decision:
$\rho(P;d) = \sum\limits_{\omega_i} L(\omega_i;d)P(\omega_i) = \frac{5}{10}L(\omega_1;d) + \frac{2}{10}L(\omega_2;d) + \frac{1}{10}L(\omega_3;d) + \frac{2}{10}L(\omega_4;d)$
so $\rho(P;d_1) = \frac{5}{10}\cdot 1+\frac{2}{10}\cdot 3+0+\frac{2}{10}\cdot 2 = 1.5$, $\rho(P;d_2) = 0+\frac{2}{10}\cdot 2+\frac{1}{10}\cdot 4+\frac{2}{10}\cdot 1 = 1$, $\rho(P;d_3) = \frac{5}{10}\cdot 3+0+\frac{1}{10}\cdot 3+\frac{2}{10}\cdot 1 = 2$.

BR $\rho^*(P) = \min\{1.5;1;2\} = 1$ and BD $d^* = d_2$.

. . . .

EX. 2. $\Omega = \{0;1\}$, $P = \begin{pmatrix} 0 & 1 \\ \frac{3}{5} & \frac{2}{5} \end{pmatrix}$, $D = [0;1]$, $L(\omega;d) = |\omega - d|^\alpha$; $\alpha \geqslant 1$.

a) $\alpha = 1$: $\rho(P;d) = \frac{3}{5}d + \frac{2}{5}(1-d) = \frac{2}{5} + \frac{1}{5}d$, BR $\rho^*(P) = \frac{2}{5}$ is attained at BD $d^* = 0$. Note that for $D = (0;1]$ BR is the same, but there is no BD.

b) $\alpha > 1$: $\rho(P;d) = \frac{3}{5}d^\alpha + \frac{2}{5}(1-d)^\alpha$. To find BR and BD, consider

$$\frac{\partial}{\partial d}\rho(P;d) = \alpha\left[\tfrac{3}{5}d^{\alpha-1} - \tfrac{2}{5}(1-d)^{\alpha-1}\right]$$

which is equal to 0 when $d\sqrt[\alpha-1]{\frac{3}{2}} = (1-d) \Leftrightarrow d = \left[1+(\frac{3}{2})^{\frac{1}{\alpha-1}}\right]^{-1} =: d_0$. It is easy to see that $d_0 \in (0;1)$ is a p. of min for $\rho(P;d)$ ($\frac{\partial}{\partial d}\rho(P;d) < 0$ if $d < d_0$, > 0 if $d > d_0$), thus BD $d^* = d_0$, and the BR is

$$\rho^*(P) = \tfrac{3}{5}\left[1+(\tfrac{3}{2})^{\frac{1}{\alpha-1}}\right]^{-\alpha} + \tfrac{2}{5}\left\{1-\left[1+(\tfrac{3}{2})^{\frac{1}{\alpha-1}}\right]^{-1}\right\}^\alpha$$

REM. In this Ex., for $\alpha \in (0;1)$ d_0 is a p. of max; $\rho(P;0) = \frac{2}{5}$, $\rho(P;1) = \frac{3}{5}$, so $d^* = 0$ and $\rho^*(P) = \frac{2}{5}$.

. . . .

3.2. Mixed decisions

We can interpret the choice of decision by Statistician as a "game" with Nature (or Chance), where she/he aims to $\to$ min her/his mean loss. Recall that, in MG and BG, a player does not always have the single optimal action, but always has an OMS — a random choice of action accordingly to some prob. distrib. Analogously, we can choose the decision $d \in D$ randomly, w.r.t. a prob. distrib. $Q(\cdot)$ on D.

Then the risk $\rho(P;Q) = \int\limits_D \rho(P;d)Q(\mathbf{d}d) = \int\limits_D\int\limits_\Omega L(\omega;d)P(\mathbf{d}\omega)Q(\mathbf{d}d)$.

DEF. $d \in D$ is a *pure decision* (PD), a prob. distrib. $Q(\cdot)$ on D is a *mixed decision* (MD).

Let the set of all MDs be S. PD d is a "degenerate" MD — a discrete prob. distrib., concentrated in d.

BR on the set of MDs is $\rho^*_S(P) = \inf\limits_{Q\in S} \rho(P;Q)$, and a mixed BD Q^* is a MD at which inf is attained.

But, the randomization does not improve/decrease BR. Indeed, let it be $\rho^* = \inf\limits_{d\in D} \rho(P;d)$. Then $\forall d \in D$: $\rho(P;d) \geqslant \rho^*$, hence

$$\forall Q \in S\colon \rho(P;Q) = \int\limits_D \rho(P;d)Q(\mathbf{d}d) \geqslant \rho^* \int\limits_D Q(\mathbf{d}d) = \rho^*$$

thus $\rho^*_S = \inf\limits_{Q\in S} \rho(P;Q) \geqslant \rho^*$.

Moreover, $\forall \varepsilon > 0$ $\exists d_\varepsilon \in D$: $\rho(P;d_\varepsilon) < \rho^* + \varepsilon$. Let Q_ε be a discrete prob. distrib., concentrated in d_ε (actually, as a decision, it is d_ε), and we have $\rho(P;Q_\varepsilon) = \rho(P;d_\varepsilon) < \rho^* + \varepsilon$, so $\rho^*_S = \inf\limits_{Q\in S} \rho(P;Q) < \rho^* + \varepsilon$ as well; due to arbitrariness of ε we get $\rho^*_S \leqslant \rho^*$, which, together with previous inequality, gives $\rho^*_S = \rho^*$.

Thus, if a mixed BD exists, then a pure BD exists too, which provides the same BR.

What is the difference, why it was more optimal to randomize an action and obtain, on the average, a non-zero gain in MG and BG (without saddle points or NE), while in such DMPs it is not the case?

It was one of OMS properties in a MG that $\max\limits_{1\leqslant i\leqslant m} W(i, Y^*) = W(X^*, Y^*)$; let the PS i_0 of P1 be the strategy at which max is attained. In a MG this property does not mean that P1 can, instead of her/his OMS X^*, always apply the PS i_0 and get the mean payoff $W(X^*, Y^*)$ as before, because, knowing about it, P2 can, instead of her/his OMS Y^*, switch to a PS j' that provides her/him a larger payoff, so that P1 has larger loss, — gets less than $W(X^*, Y^*)$.

And the "Statistician vs. Nature" game is not antagonistic; we also say that Nature is not a "sentient adversary", it is "indifferent". Its "strategy", a prob. distrib. $P(\cdot)$, does not depend on Statistician's one. Nature here does not have a "payoff": neither its own, nor associated with Statistician's payoff, and does not have a "goal", so it will not exploit "weaknesses" of Statistician to make it "better for itself" or "worse for an adversary".

On the one hand, we interpret this equality $\rho^*_S = \rho^*$ as follows: at making decisions under "non-ill-intentioned" indetermined conditions, a randomization does not improve an outcome; that is, to attain the best, on the average, outcome, it suffices to act determinably, always choosing pure BD.

On the other hand, a randomization allows to surmount some subjective deviations and persuade others that these deviations have not affected the results obtained.

3.3. Criteria of Decision Making under Indeterminacy

Here we consider DMPs where a prob. distrib. $P(\cdot)$ on Ω — "Nature's strategy" — is unknown, thus Statistician cannot calculate risks, determine BR and BD that corresponds to BR. What she/he knows are the states $\Omega = \{\omega_1; \ldots; \omega_m\}$, the decisions $D = \{d_1; \ldots; d_n\}$, and the *loss matrix* $L = \|L(\omega_i; d_j)\|_{\substack{1 \leqslant i \leqslant m, \\ 1 \leqslant j \leqslant n.}}$

<u>EX. 3.</u> $m = n = 4$, $L = \begin{pmatrix} 8 & 7 & 6 & 10 \\ 7 & 11 & 1 & 4 \\ 1 & 3 & 9 & 2 \\ 5 & 2 & 4 & 6 \end{pmatrix}$. What decision is the best?

1. Laplace Criterion. This criterion corresponds with the so-called *insufficient basis principle*, accordingly to which, if we have no information about the distribution of the states ω, then we can, due to symmetry, assume the states to be equiprobable: $P(\omega_i) \equiv \frac{1}{m}$ (there are no reasons to prefer some specific deviation from equiprobability). Then we apply Bayesian approach: the BR $\rho^* = \min\limits_{1 \leqslant j \leqslant n} \frac{1}{m} \sum\limits_{i=1}^{m} L(\omega_i; d_j)$, BD d^* is any d_j at which $\min\limits_j$ is attained. Obviously, the set of such j is the same for $m\rho^* = \min\limits_j \sum\limits_i L(\omega_i; d_j)$, — we choose the column where the sum of the coefficients is minimal.

We apply this criterion to Ex. 3: $4\rho(d_1) = 8 + 7 + 1 + 5 = 21$, $4\rho(d_2) = 23$, $4\rho(d_3) = 20$, $4\rho(d_4) = 22$, hence $\rho^* = \frac{1}{4} \min\{21; 23; 20; 22\} = 5$, OD/BD is d_3.

2. Minimax Criterion. Sometimes this criterion is called "cautious" or "pessimistic", — we choose the best of the worst possibilities. When we apply this criterion, we may conceive Nature as an "adversary", who aims to maximize the loss of Statistician (and knows her/his strategy).

If we choose the decision d_j, then the largest possible loss is $\max\limits_{1 \leqslant i \leqslant m} L(\omega_i; d_j)$. As the risk, we take the smallest of these: $\rho^* = \min\limits_{1 \leqslant j \leqslant n} \max\limits_{1 \leqslant i \leqslant m} L(\omega_i; d_j)$, and OD d^* is any d_j at which $\min\limits_j$ is attained. That is, Statistician acts so that to have the minimal loss in the antagonistic game with Nature.

In Ex. 3, by this criterion, $\rho^* = \min\{8; 11; 9; 10\} = 8$, OD d^* is d_1.

3. Savage Criterion. "Excessive pessimism" of min max-criterion and non-optimality of corresponding decision are illustrated by situations such as:

<u>EX. 4.</u> $m = n = 2$ and $L = \begin{pmatrix} 1001 & 1000 \\ 1 & 1000 \end{pmatrix}$.

min max-criterion recommends the decision d_2, which guarantees the loss 1000.

However, d_1 looks better: even if "Nature tends to harm", there is a "chance" (which may be significant) that the random state will be ω_2 and the loss will be only 1.

To get over this shortcoming, we introduce a *regret function/matrix* $R = \|R(\omega_i; d_j)\|_{\substack{1\leqslant i\leqslant m,\\ 1\leqslant j\leqslant n}}$, where

$$R(\omega_i; d_j) = L(\omega_i; d_j) - \min_{d\in D} L(\omega_i; d)$$

expresses the "regret" of Statistician as an additional loss that she/he will have if she/he makes this decision in this state, comparing to the best decision in this state. After that, we apply min max-criterion to R instead of L: the risk $\rho^* = \min_j \max_i R(\omega_i; d_j)$, and OD d^* is d_j, at which $\min_j$ is attained. Now Statistician aims to minimize regret, not loss, in the antagonistic game with Nature.

In Ex. 4 the regret matrix $R = \begin{pmatrix} 1 & 0 \\ 0 & 999 \end{pmatrix}$, and $\rho^* = \min\{1; 999\} = 1$ is attained for OD $d^* = d_1$, which corresponds to an "adequate" comparison of decisions. Note that Laplace criterion provides the same result.

In Ex. 3 we have $R = \begin{pmatrix} 2 & 1 & 0 & 4 \\ 6 & 10 & 0 & 3 \\ 0 & 2 & 8 & 1 \\ 3 & 0 & 2 & 4 \end{pmatrix}$, $\rho^* = \min\{6; 10; 8; 4\} = 4$, OD $d_4 \neq d_1$.

4. Hurwitz Criterion. In contrast to preceding criteria, this one has additional "regulator", or, more precisely, the parameter that determines the ratio of "pessimism" and "optimism".

Under the "most pessimistic" assumption, — when Nature tries to maximize the (Statistician's) loss, — for any Statistician's decision d_j such loss is $L_-(d_j) = \max_i L(\omega_i; d_j)$; under the "most optimistic" assumption, when Nature, on the contrary, tries to minimize the loss, its minimal value is $L_+(d_j) = \min_i L(\omega_i; d_j)$; either way, OD is d_j that provides $\min_j L_\pm(d_j)$.

Let an *optimism index* be $\alpha \in [0; 1]$. We assume the risk $\rho^* = \min_j \big[\alpha L_+(d_j) + (1-\alpha)L_-(d_j)\big]$ and let OD d^* be any d_j, at which $\min_j$ is attained.

We apply this criterion in Ex. 3 for $\alpha = 0.5$. Intermediate results, as a table:

	L_+	L_-	$\alpha L_+ + (1-\alpha)L_-$
d_1	1	8	4.5
d_2	2	11	6.5
d_3	1	9	5
d_4	2	10	6

This implies $\rho^* = \min\{4.5; 6.5; 5; 6\} = 4.5$ and OD d^* is d_1.

3.4. Statistical Decision Making Problems

DMP with observations. Let Statistician have the additional information: before making a decision $d \in D$ she/he observes the RE whose outcome, $\xi \in X$, is known to her/him and has a distribution that depends on parameter $W = \omega \in \Omega$ of the initial RE. $W = \omega$ remains unknown though. For example,

– ω is RandVar with uniform distribution on $\Omega = [0; 1]$,

– observation ξ is the outcome, known to Statistician, of a coin toss, where the probability of heads is $p = \omega$ ($\xi \in X = \{0; 1\}$, 0 for tails, 1 for heads),

– d is an estimate of heads probability ω, at that $d \in D = [\frac{1}{8}; \frac{7}{8}]$,

– if Statistian makes the decision d, she/he will lose $L(\omega; d) = |\omega - d|^4$ "units" (of money, blood, ...). Naturally, Statistician aims at $d = \omega$.

We can expect that this information helps to make a better — more accurate — decision and decrease a loss. All the more we should expect this when Statistician observes several realizations of ξ for the same parameter $W = \omega$; on the other hand, we can consider $X = Y^n$ right away, where Y is made from possible values of a single observation, n is the number of observations. We assume that

• For any $\omega \in \Omega$ we know the conditional distrib. $F(\cdot|\omega)$ of ξ at the value of the parameter $W = \omega$.

Such DMPs are the ones we call *statistical* (SDMP).

DEF. $\delta\colon X \to D$ that maps any PV x of the observation ξ to a decision d is called a *decision function* (DF). We denote the set of all DFs by Δ.

Although, in practice, it suffices to be able to find $\delta(x)$ for the concrete PV x that RandVar ξ has taken in a given RE, we assume that the goal of Statistician is to determine $\delta(x)$ for all $x \in X$. Put differently, we require that Statistician is able to explain, *before* the observation, how she/he will make a decision according to any possible outcome of the observation.

Let $W = \omega$ have the distribution $P(\cdot)$. Statistician chooses DF $\delta \in \Delta$, aiming to minimize the mean loss (the risk). In SDMPs there are several "risk" notions:

1) The risk introduced before, for a fixed decision $d \in D$:

$$\rho(P; d) = \mathrm{E}\, L(W; d) = \int_\Omega L(\omega; d) P(d\omega)$$

2) For any $\omega \in \Omega$ the corresponding risk at applying DF δ when $W = \omega$:

$$\rho(\omega; \delta) = \mathrm{E}\, L(\omega; \delta(\xi)) = \int_X L(\omega; \delta(x)) F(dx|\omega)$$

For this definition of risk to be proper, we require the following additional condition to hold:

• $\forall \omega \in \Omega$: $L(\omega; \delta(x))$ is measurable & integrable w.r.t. measure $F(\cdot|\omega)$ on X.

DEF. For a fixed $\delta \in \Delta$ the function $\rho(\cdot;\delta)\colon \Omega \to \mathbb{R}$ is called a *risk function* (RF) of a given DF δ.

3) The risk at applying DF δ:
$\rho(P;\delta) = \mathrm{E}\, L(W;\delta(\xi)) = \int\limits_{\Omega} \rho(\omega;\delta)P(d\omega) = \int\limits_{\Omega}\int\limits_{X} L(\omega;\delta(x))F(dx|\omega)P(d\omega)$
Presumably, this is the "main" risk, — the one Statistician wants to minimize.

DEF. $\rho^*(P) = \inf\limits_{\delta\in\Delta} \rho(P;\delta)$ is called a *Bayes risk (in SDMP).*

DEF. DF $\delta^* \in \Delta$ such that $\rho(P;\delta^*) = \rho^*(P)$ is called a *Bayes DF* (BDF).
Obviously, Statistician should choose and apply BDF as her/his DF.

Construction of BDF. We assume that 1) $F(dx|\omega) = f(x|\omega)G(dx)$ and 2) in the expression for $\rho(P;\delta)$ we can change the order of integration; in particular, it is possible when $\exists C > 0$: $\forall\omega \in \Omega, \forall d \in D$: $0 \leqslant L(\omega;d) \leqslant C$, that is, the loss function $L(\cdot)$ is non-negative and bounded.

Then $\rho(P;\delta) = \int\limits_{X}\big[\int\limits_{\Omega} L(\omega;\delta(x))f(x|\omega)P(d\omega)\big]G(dx)$ and, as a solution of the problem $\rho(P;\delta) \to \min$, we can take DF $\delta\colon X \to D$ such that $\int\limits_{\Omega}(...) \to \min$ for any $x \in X$. Since this $\int\limits_{\Omega}(...)$ looks like "usual" risk relative to the decision $d = \delta(x)$, we assign $\delta^*(x) := d^*$, where

$$\int\limits_{\Omega} L(\omega;d^*)f(x|\omega)P(d\omega) = \inf_{d\in D}\int\limits_{\Omega} L(\omega;d)f(x|\omega)P(d\omega)$$

To give another interpretation of this approach, we define the function $f_1(x) = \int\limits_{\Omega} f(x|\omega)P(d\omega)$.

Since $\int\limits_{X} f_1(x)G(dx) = \int\limits_{\Omega}\int\limits_{X} f(x|\omega)G(dx)P(d\omega) = \int\limits_{\Omega} P(d\omega) = 1$ and $f_1(x) \geqslant 0$, it follows that this function is a density (w.r.t. the measure $G(\cdot)$) of some prob. distrib. on X. We divide by it the expression that we need to $\to$ min and obtain the problem of determining $d = \delta(x)$ such that $\int\limits_{\Omega} L(\omega;d)\frac{f(x|\omega)}{f_1(x)}P(d\omega) \to \min$.

For the sake of convenience, we also assume that $P(d\omega) = p(\omega)H(d\omega)$, that is, prob. distrib. $P(\cdot)$ is absolutely continuous (abs. cont.) w.r.t. the measure $H(\cdot)$ with the density $p(\omega)$. Then $\int\limits_{\Omega}(...) = \int\limits_{\Omega} L(\omega;d)\big[\frac{f(x|\omega)p(\omega)}{f_1(x)}\big]H(d\omega)$, where the expression in [...] is the conditional density $p(\omega|x)$ of W distrib. for $\xi = x$, thus the $\int\limits_{\Omega}(...)$ itself is a conditional expectation, $\mathrm{E}\{L(W;d)|\xi = x\}$. Accordingly, d^* is a (Bayes) decision that minimizes the mean loss for the conditional distribution of W under the condition $\xi = x$.

DEF. In a SDMP, the unconditional distrib. $P(d\omega)$ of the parameter W is called *prior* (before the observation/experiment), and the conditional distrib. $P(d\omega|x) = \frac{f(x|\omega)}{f_1(x)}P(d\omega)$ of this parameter is called *posterior* (after observation).

So, either without observations, or with them, it is more advantageous for Statistician to choose BD at which BR is attained; the difference is as follows: in case when there are no observations, BR is determined w.r.t. the prior distrib. of W, and in case when they are present, w.r.t. the posterior one.

EX. 5. $\Omega = \{\omega_1; \omega_2\}$, $D = \{d_1; d_2\}$, $L = \begin{pmatrix} 0 & 4 \\ 8 & 0 \end{pmatrix}$. The parameter $W \in \Omega$ has the prior distrib. $P(\omega_1) = p$, $P(\omega_2) = 1 - p$, where $p \in [0; 1]$ ("parameter of parameter's distrib."). RV $\xi \in X = \{0; 1\}$ has the following distrib.: when $W = \omega_1$, $\xi \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{5} & \frac{4}{5} \end{pmatrix}$, and when $W = \omega_2$, $\xi \sim \begin{pmatrix} 0 & 1 \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}$.

Statistician knows $P(\cdot)$ (in fact, p) and observes ξ, but does not know ω. Let us find BDF $\delta^*(x)$ — what decision she/he should make depending on the realization of ξ — and BR. In order to do so, we determine the posterior distrib. of W for each PV of ξ, applying Bayes formulae

$$P\{W = \omega_i|\xi = x\} = \frac{P\{\xi = x|W = \omega_i\}P\{W = \omega_i\}}{\sum\limits_j P\{\xi = x|W = \omega_j\}P\{W = \omega_j\}}$$

Since there are 2 values of W, it suffices to know $P\{W = \omega_1|\xi = x\} =: p(x)$, then $P\{W = \omega_2|\xi = x\} = 1 - p(x)$. We have

$p(0) = P\{W = \omega_1|\xi = 0\} = \frac{\frac{1}{5}p}{\frac{1}{5}p + \frac{3}{4}(1-p)}$, $\quad p(1) = P\{W = \omega_1|\xi = 1\} = \frac{\frac{4}{5}p}{\frac{4}{5}p + \frac{1}{4}(1-p)}$

The risk of decision d_1 is $\rho(P(\cdot|x); d_1) = 0 \cdot p(x) + 8(1 - p(x)) = 8(1 - p(x))$, the risk of d_2, analogously, is $4p(x)$. BR is the smallest of them; $8(1 - p(x)) \leqslant 4p(x)$ $\Leftrightarrow 8 \leqslant 12p(x) \Leftrightarrow p(x) \geqslant \frac{2}{3}$. Thus, for $p(x) > \frac{2}{3}$ BD is d_1, for $p(x) < \frac{2}{3}$ BD is d_2, and for $p(x) = \frac{2}{3}$ they both are BD.

a) $x = 0$: $p(0) < \frac{2}{3} \Leftrightarrow \frac{3}{5}p < \frac{2}{5}p + \frac{3}{2}(1 - p) \Leftrightarrow 6p < 4p + 15(1 - p) \Leftrightarrow p < \frac{15}{17}$. That is, for $p < \frac{15}{17}$ BD is d_2, for $p > \frac{15}{17}$ BD is d_1, for $p = \frac{15}{17}$ both are BD.

b) $x = 1$: $p(1) < \frac{2}{3} \Leftrightarrow \frac{12}{5}p < \frac{8}{5}p + \frac{1}{2}(1 - p) \Leftrightarrow 24p < 16p + 5(1 - p) \Leftrightarrow p < \frac{5}{13}$; if $p < \frac{5}{13}$, then BD is d_2; if $p > \frac{5}{13}$, then BD is d_1; if $p = \frac{5}{13}$, then both are BD.

$$\text{BDF } \delta^*(x) = \delta^*(x; p) = \begin{cases} d_1, & \big((x = 0) \wedge (p \geqslant \frac{15}{17})\big) \vee \big((x = 1) \wedge (p \geqslant \frac{5}{13})\big), \\ d_2, & \big((x = 0) \wedge (p \leqslant \frac{15}{17})\big) \vee \big((x = 1) \wedge (p \leqslant \frac{5}{13})\big). \end{cases}$$

Let us represent $\rho^* = \rho^*(p)$. Since BR is attained at BDF δ^* described above, we have $\rho^* = \sum\limits_\omega \sum\limits_x L(\omega; \delta^*(x))P\{\xi = x|W = \omega\}P\{W = \omega\} =$

$= \frac{1}{5}pL(\omega_1; \delta^*(0)) + \frac{4}{5}pL(\omega_1; \delta^*(1)) + \frac{3}{4}(1 - p)L(\omega_2; \delta^*(0)) + \frac{1}{4}(1 - p)L(\omega_2; \delta^*(1))$

1) $p \leqslant \frac{5}{13} < \frac{15}{17}$: $\delta^*(0) = \delta^*(1) = d_2$, so $\rho^* = \frac{4}{5}p + \frac{16}{5}p + 0 + 0 = 4p$;
2) $\frac{5}{13} \leqslant p \leqslant \frac{15}{17}$: $\delta^*(0) = d_2$, and $\delta^*(1) = d_1$, thus

$$\rho^* = \frac{4}{5}p + 0 + 0 + 2(1-p) = 2 - \frac{6}{5}p$$

3) $p \geqslant \frac{15}{17} > \frac{5}{13}$: $\delta^*(0) = \delta^*(1) = d_1$, hence

$$\rho^* = 0 + 0 + \frac{24}{4}(1-p) + 2(1-p) = 8(1-p)$$

We remark that $\rho^*(p)$ is continuous on $[0;1]$.

If Statistician did make a decision from D without the observation ξ, — a priori, — then BR would be $\min\{4p; 8(1-p)\}$; accordingly, BD d^* for $p \leqslant \frac{2}{3}$ would be d_2, and for $p \geqslant \frac{2}{3}$ it would be d_1.

Observation Cost. In a SDMP, the observation, by Statistician, of a RV whose distribution depends on unknown parameter may lead to additional losses (e.g. to carry out the experiment where realization of ξ is obtained, Statistician buys equipment, employs staff, makes agreements with suppliers of resources etc.)

Let $c(\omega; x)$ be the *cost of observation* (OC) of the PV x of RV ξ for $W = \omega$, that is, what Statistician loses additionally at that. Then the mean OC

$$\mathrm{E}\,C = \mathrm{E}\,C(W;\xi) = \int\limits_{\Omega}\int\limits_{X} c(\omega;x)F(dx|\omega)P(d\omega)$$

Accordingly, the total risk at applying DF δ is $\mathrm{E}\,L(W;\delta(\xi)) + \mathrm{E}\,C(W;\xi)$. Applying it, we assume that

• Statistician's losses are additive, that is, under these conditions she/he loses $L(\omega;\delta(x)) + c(\omega;x)$. In particular, it means that the loss L and the cost C are measured in the same "units".

REM. Often OC depends on the number of observations — sample size — and does not depend on parameter or PV: if the cost of 1 observation is c, and their number is n, then Statistician loses nc.

EX. 6. Let in SDMP from Ex. 5 the cost of ξ observation be $c > 0$. Before making a decision d_1 or d_2 Statistician can either do not observe ξ (then she/he does not need to "pay" c), or observe (and pay c). Thus, the "full" decision is 2-stage: at first, decide, to observe or not to observe, and then decide, d_1 or d_2.

How should she/he act, knowing p and c?

Without observation BR $\rho_0^* = \min\{4p; 8(1-p)\} = \begin{cases} 4p, & p \leqslant \frac{2}{3}, \\ 8(1-p), & p \geqslant \frac{2}{3}, \end{cases}$ and

with observation total BR $\rho_1^* = c + \begin{cases} 4p, & p \leqslant \frac{5}{13}, \\ 2 - \frac{6}{5}p, & \frac{5}{13} \leqslant p \leqslant \frac{15}{17}, \\ 8(1-p), & p \geqslant \frac{15}{17}. \end{cases}$ We remark that $\frac{5}{13} < \frac{2}{3} < \frac{15}{17}$. Obviously, when $p \in \left[0; \frac{5}{13}\right] \cup \left[\frac{15}{17}; 1\right]$, we have $\rho_1^* > \rho_0^*$, so it is more advantageous for Statistician to make a decision without observation.

When $p \in \left[\frac{5}{13}; \frac{15}{17}\right]$, since for $p \in \left[\frac{5}{13}; \frac{2}{3}\right]$ inequality $\rho_0^* = 4p \geqslant 2 - \frac{6}{5}p = \rho_1^*$ holds, and for $p \in \left[\frac{2}{3}; \frac{15}{17}\right]$ inequality $\rho_0^* = 8(1-p) \geqslant 2 - \frac{6}{5}p = \rho_1^*$ holds as well, it is more advantageous for Statistician to make a decision <u>with</u> observation if $c \leqslant \rho_0^* - \rho_1^* = \min\{4p; 8(1-p)\} - \left(2 - \frac{6}{5}p\right)$, and <u>without</u> it otherwise. The largest difference $\rho_0^* - \rho_1^*$ between risks is attained when $p = \frac{2}{3}$, it is equal to $\frac{22}{15}$.

It is possible that observations and decisions occur for several stages, and the decision made at the current moment affects the information that Statistician will obtain in the future (e.g. at stage I she/he observes the realization of RV ξ and, depending on it, decides, how many realizations of RV η she/he will observe at stage II). Such DMPs are called *consecutive decision making problems.*

3.5. Conjugate families of distributions

<u>DEF.</u> Let ξ_1, ..., ξ_n be a (repeated) sample from the prob. distrib. $F(\cdot; W)$. The family $\mathcal{P}$ of prob. distributions is called *closed w.r.t. sampling (from $F(\cdot; W)$)* or *conjugate (for $F(\cdot; W)$) family of distributions* when it has the following property: if the prior distrib. of parameter W belongs to $\mathcal{P}$, then, for any sample size n and for any values of ξ_i, the posterior distrib. of W belongs to $\mathcal{P}$ too.

<u>REM.</u> In the following, the symbol "$\sim$" in expressions such as $A(\omega) \sim B(\omega)$ means that $\frac{A(\omega)}{B(\omega)} = C$, where C does not depend on ω. $C = 1$ is possible.

<u>EX.</u> Let $\{\xi_i\}_{i=1}^n$ be a (repeated) sample from Bernoulli distribution with unknown parameter W, which is a probability of success: $P\{\xi_i = 1\} = W$, $P\{\xi_i = 0\} = 1 - W$, or, more generally, $P\{\xi_i = t\} = W^t(1-W)^{1-t}$ for $t = 0, 1$. Also, let W have β distrib. with given parameters $(\alpha; \beta)$, that is, prior density of W distrib. is $p(\omega) \sim \omega^{\alpha-1}(1-\omega)^{\beta-1}$ for $\omega \in (0;1)$; to be more precise, $p(\omega) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\omega^{\alpha-1}(1-\omega)^{\beta-1}I_{(0;1)}(\omega)$.

We denote the conditional distrib. of the sample $\{\xi_i\}_{i=1}^n$, — i.e. the conditional joint distrib. of its elements ξ_1, ..., ξ_n, — under condition $W = \omega$ by $f(x|\omega) = \prod_{i=1}^n P\{\xi_i = x_i|\omega\}$, where $x = (x_1; ...; x_n)$. Then

$$f(x|\omega) = \prod_{i=1}^n \omega^{x_i}(1-\omega)^{1-x_i} = \omega^{\sum\limits_i x_i}(1-\omega)^{n-\sum\limits_i x_i}$$

Due to aforementioned reasonings (see "Construction of BDF"), the posterior density of W distrib. $p(\omega|x) \sim p(\omega)f(x|\omega)$, thus $p(\omega|x) \sim \omega^{\alpha+\sum x_i - 1}(1-\omega)^{\beta+n-\sum x_i-1}$. We denote $\sum x_i =: y$ and rewrite $p(\omega|x) \sim \omega^{\alpha+y-1}(1-\omega)^{\beta+n-y-1}$, which is the β distrib. with parameters $\alpha' = \alpha + y$, $\beta' = \beta + n - y$.

This Ex. proves the following

THEOR. The family of β distributions is conjugate for the class of Bernoulli distributions by the parameter W — probability of success at a single trial. At that the posterior parameters are related to the prior ones as follows:

$$\alpha' = \alpha + y,\ \beta' = \beta + n - y$$

REM. Were in this Ex. the probability of success $\frac{W}{2}$, then W with β distrib. would be, as well, a parameter of the (Bernoulli) distrib. from which the sample was taken; however, the conditional density would have another form ($f_1(x|\omega) \sim \omega^{\sum x_i}(2-\omega)^{n-\sum x_i}$) and the posterior distrib. of W would not be a β distrib. anymore. So the Theor. specifies how, exactly, the prob. distrib. from which the sample is taken depends on the parameter. On the other hand, if this dependency is clearly determined from the context, we may omit such specification.

....

THEOR. The family of β distrib-s with param-s $(\alpha;\beta)$ is conjugate for $B_{m;W}$ ($F(k) = C_m^k W^k (1-W)^{m-k}$ for $k = \overline{0,m}$), where m is fixed, by W; at that posterior param-s $\alpha' = \alpha + \sum\limits_i x_i$, $\beta' = \beta + mn - \sum\limits_i x_i$.

◀ $p(\omega) \sim \omega^{\alpha-1}(1-\omega)^{\beta-1}$, for $y = \sum\limits_i x_i$: $f(x|\omega) = \omega^y (1-\omega)^{mn-y}$; $p(\omega|x) \sim \omega^{\alpha+y-1}(1-\omega)^{\beta+mn-y-1}$. ▶

....

THEOR. The family of Γ distrib-s with param-s $(\theta;\nu)$ (density $f(t) = \frac{\theta^\nu}{\Gamma(\nu)} t^{\nu-1} e^{-\theta t} I_{(0;+\infty)}(t)$) is conjugate for Poisson distrib. ($F(k) = \frac{W^k}{k!} e^{-W}$, $k \in \mathbb{Z}_+$) by W; posterior param-s $\theta' = \theta + n$, $\nu' = \nu + \sum\limits_i x_i$.

◀ The density of W's prior distrib. is $p(\omega) \sim \omega^{\nu-1} e^{-\theta\omega}$. Let $y = \sum\limits_{i=1}^n x_i$, where $x = (x_1; ...; x_n)$ are PVs of Poisson sample ξ. The cond. distrib. of ξ under $W = \omega$ is $f(x|\omega) = \prod\limits_{i=1}^n \frac{\omega^{x_i}}{x_i!} e^{-\omega} \sim \omega^y e^{-n\omega}$. Then the posterior distrib. density $p(\omega|x) \sim p(\omega) f(x|\omega) \sim \omega^{\nu+y-1} e^{-(\theta+n)\omega}$, which, to within a multiplier independent of ω, is the density of Γ distrib. with param-s $\theta' = \theta + n$, $\nu' = \nu + y$. ▶

THEOR. The family of β distrib-s with param-s $(\alpha;\beta)$ is conjugate for negative binomial distrib. with param-s $(r;W)$, where r is fixed ($F(k) = C_{r+k-1}^{r-1} W^r (1-W)^k$, $k \in \mathbb{Z}_+$), by param. W (probability of success at one trial), at that posterior param-s $\alpha' = \alpha + rn$, $\beta' = \beta + \sum\limits_i x_i$.

◀ Prior distrib. of W has density $p(\omega) \sim \omega^{\alpha-1}(1-\omega)^{\beta-1}$, conditional distrib. $f(x|\omega) \sim \omega^{rn}(1-\omega)^{\sum x_i}$, hence for $y = \sum x_i$ we have $p(\omega|x) \sim \omega^{\alpha+rn-1}(1-\omega)^{\beta+y-1}$ — β distrib. with given param-s. ▶

COR. Under the conditions of this Theor. for $r = 1$ we have the geometric distrib., for which, therefore, the family of β distrib-s is conjugate by param. W — success probability.

THEOR. The family of Γ distrib-s with param-s $(\theta;\nu)$ is conjugate for exponential distrib. (whose density is $f(t){=}We^{-Wt}I_{[0;+\infty)}(t)$) by param. W, at that posterior param-s $\theta' = \theta + \sum\limits_i x_i$, $\nu' = \nu + n$.

◄ $p(\omega) \sim \omega^{\nu-1}e^{-\theta\omega}$, $f(x|\omega) \sim \omega^n e^{-\omega y}$, where $y = \sum\limits_i x_i$, so $p(\omega|x) \sim \omega^{\nu+n-1}e^{-(\theta+y)\omega}$. ►

THEOR. The family of Pareto distrib-s with param-s $(a;b)$ (density $f(t) = \frac{ab^a}{t^{a+1}}I_{[b;+\infty)}(t)$) is conjugate for uniform distrib. on $[0;W]$ by param. W, at that posterior param-s $a' = a+n$, $b' = \max\{b;x_1;...;x_n\}$.

◄ For $\omega \geqslant b$: $p(\omega) \sim \frac{1}{\omega^{a+1}}$, for $\omega < b$: $p(\omega) = 0$. The conditional density of the sample $f(x|\omega) = \prod\limits_{i=1}^{n}\begin{cases}0, & x_i \notin [0;\omega],\\ \frac{1}{\omega}, & x_i \in [0;\omega]\end{cases} = \begin{cases}0, & \exists x_i \notin [0;\omega],\\ \omega^{-n}, & \forall x_i \in [0;\omega].\end{cases}$

Therefore the posterior density $p(\omega|x) > 0$ **iff** $\omega \geqslant b$ and $\omega \geqslant x_i$, $i = \overline{1,n}$, which is equivalent to $\omega \geqslant b'$. For such ω: $p(\omega|x) \sim \omega^{-a-1}\cdot\omega^{-n} \sim \omega^{-(a+n+1)}$, and for $\omega < b'$ we have $p(\omega|x) = 0$. Thus $p(\omega|x)$ is the density of Pareto distrib. with given posterior parameters. ►

DEF. The *precision measure* (PM) of $N_{a;\sigma^2}$ is $\tau = \sigma^{-2}$.

Accordingly, the density $p(u) = \sqrt{\frac{\tau}{2\pi}}e^{-\frac{1}{2}\tau(u-a)^2}$. To distinguish from the variance σ^2, we denote $N_{a;(\tau)}$.

THEOR. The family of N distrib-s with param-s $(a;(\tau))$ (density is given above) is conjugate for N distrib. with fixed PM $r > 0$ and unknown mean W (density $f(u) = \sqrt{\frac{r}{2\pi}}e^{-\frac{1}{2}r(u-W)^2}$) by param. W, at that posterior param-s $a' = \frac{\tau a + nr\overline{x}}{\tau + nr}$ and $\tau' = \tau + nr$, where $\overline{x} = \frac{1}{n}\sum\limits_{i=1}^{n} x_i$.

◄ $p(\omega) = \sqrt{\frac{\tau}{2\pi}}e^{-\frac{1}{2}\tau(\omega-a)^2} \sim e^{-\frac{1}{2}\tau(\omega-a)^2}$, $f(x|\omega) \sim e^{-\frac{1}{2}r\sum\limits_i(x_i-\omega)^2}$. Therefore $p(\omega|x) \sim e^{-\frac{1}{2}\left[\tau(\omega-a)^2 + r\sum\limits_i(x_i-\omega)^2\right]}$.

$$\tau(\omega-a)^2 + r\sum_i(x_i-\omega)^2 = \tau\omega^2 - 2\tau\omega a + \tau a^2 + r\sum_i(\omega^2 - 2\omega x_i + x_i^2) =$$
$$= \tau\omega^2 - 2\tau\omega a + \tau a^2 + nr\omega^2 - 2r\omega\sum_i x_i + r\sum_i x_i^2 =$$
$$= (\tau+nr)\omega^2 - 2(\tau a + nr\overline{x})\omega + \left[\tau a^2 + r\sum_i x_i^2\right] =$$

$$= (\tau + nr)\left[\omega^2 - 2\tfrac{\tau a + nr\overline{x}}{\tau+nr}\omega + \left(\tfrac{\tau a+nr\overline{x}}{\tau+nr}\right)^2\right] + \underbrace{\left[\tau a^2 + r\sum_i x_i^2 - \frac{(\tau a + nr\overline{x})^2}{\tau+nr}\right]}_{A,\text{ does not depend on }\omega} =$$

$$= (\tau+nr)\left[\omega - \tfrac{\tau a+nr\overline{x}}{\tau+nr}\right]^2 + A$$

Thus $p(\omega|x) \sim e^{-\frac{1}{2}\tau'(\omega-a')^2-\frac{1}{2}A} \sim e^{-\frac{1}{2}\tau'(\omega-a')^2}$ — the density of $N_{a';(\tau')}$ ▶

REM. $\tau' = \tau + nr \xrightarrow[n\to\infty]{} +\infty$: as the sample size increases, the posterior distrib. "concentrates" at its mean a'.

THEOR. The family of Γ distrib-s with param-s $(\theta;\nu)$ is conjugate for N distrib. with fixed mean μ and unknown PM W ($f(u) = \sqrt{\frac{W}{2\pi}}e^{-\frac{1}{2}W(u-\mu)^2}$) by W; posterior param-s $\theta' = \theta + \frac{1}{2}\sum\limits_i(x_i-\mu)^2$, $\nu' = \nu + \frac{n}{2}$.

◀ $p(\omega) \sim \omega^{\nu-1}e^{-\theta\omega}$, $f(x|\omega) \sim \omega^{\frac{n}{2}}e^{-\frac{1}{2}\omega\sum\limits_i(x_i-\mu)^2}$, hence

$$p(\omega|x) \sim \omega^{\nu+\frac{n}{2}-1}e^{-\left[\theta+\frac{1}{2}\sum\limits_i(x_i-\mu)^2\right]\omega} \quad ▶$$

DEF. *2-side 2-dimensional Pareto distrib.* with param-s $a > 0$, $b_1 < b_2$ is the distrib. on $\mathbb{R}^2$ that has the density $p(u;v) = \begin{cases} 0, & u > b_1 \vee v < b_2, \\ \frac{a(a+1)(b_2-b_1)^a}{(v-u)^{a+2}}, & u \leqslant b_1 \wedge v \geqslant b_2. \end{cases}$

PROP. If $\xi = (\xi_1;\xi_2)$ has such distrib., then $\mathrm{E}\,\xi_1 = \frac{ab_1-b_2}{a-1}$, $\mathrm{E}\,\xi_2 = \frac{ab_2-b_1}{a-1}$, $\mathrm{Var}\,\xi_1 = \mathrm{Var}\,\xi_2 = \frac{a(b_2-b_1)^2}{(a-1)^2(a-2)}$.

THEOR. The family of 2-side 2-dim. Pareto distrib-s with param-s a, b_1, b_2 is conjugate for the uniform distrib. on $[W_1;W_2]$ by param. $W = (W_1;W_2)$, at that posterior param-s $a' = a+n$, $b_1' = \min\{b_1;x_1;...;x_n\}$, $b_2' = \max\{b_2;x_1;...;x_n\}$.

◀ $p(\omega) = p(\omega_1;\omega_2) \sim \begin{cases} 0, & \omega_1 > b_1 \vee \omega_2 < b_2, \\ (\omega_2-\omega_1)^{-(a+2)}, & \omega_1 \leqslant b_1 \wedge \omega_2 \geqslant b_2, \end{cases}$

$$f(x|\omega) = \begin{cases} 0, & \exists x_i \notin [\omega_1;\omega_2], \\ (\omega_2-\omega_1)^{-n}, & \forall x_i \in [\omega_1;\omega_2]. \end{cases}$$

Thus $p(\omega|x) \sim \begin{cases} (\omega_2-\omega_1)^{-(a+n+2)}, & \omega_1 \leqslant b_1, x_1, ..., x_n \wedge \omega_2 \geqslant b_2, x_1, ..., x_n, \\ 0, & \text{otherwise.} \end{cases}$

The inequalities are equivalent to $\omega_1 \leqslant b_1' \wedge \omega_2 \geqslant b_2'$, so $p(\omega|x)$ is the density of 2-side 2-dim. Pareto distrib. with param-s a', b_1', b_2'. ▶

3.6. Bayes Estimators

Let us consider a problem of estimating parameters by a sample as a (S)DMP, where the decision is the estimate $d = (d_1; ...; d_k) \in \mathbb{R}^k$ of the value $w = (w_1; ...; w_k) \in \mathbb{R}^k$ of the parameter $W = (W_1; ...; W_k) \in \mathbb{R}^k$. In this case the set D of decisions coincides with the set Θ of parameter W values.

Let $D = \Theta = \mathbb{R}^k$. Probability that W belongs to some parts of $\mathbb{R}^k$ may be 0.

A *loss (function)* (LF) $L(w; d)$ is a real number that reflects the difference between the exact (unknown) w and the estimate-decision d. We assume that

$$L(w; d) = \gamma(w)\Lambda(w - d)$$

where $\Lambda(z) \geqslant 0$ is a function of the vector of errors $z = w - d$, $\Lambda(0) = 0$, and $\gamma(w) \geqslant 0$ is a weight function that determines the relative significance of errors vector for different w. Particular case: $\gamma(w) \equiv \text{const}$.

Let $p_x = p(w|x)$ ($\sim p(w)f(x|w)$) be the posterior density of the W distrib. w.r.t. the measure $\mu(\cdot)$, under condition that the observation ξ, whose distrib. depends on W, takes its possible value $x \in X$.

DEF. The *risk (w.r.t. PV x)* of a decision d is

$$\rho(p_x; d) = \int_{\mathbb{R}^k} L(w; d)p_x(w)\mu(dw) = \int_{\mathbb{R}^k} \gamma(w)\Lambda(w - d)p(w|x)\mu(dw)$$

DEF. The *Bayes estimator (w.r.t. PV x)* (BE) is a decision $d^* \in \mathbb{R}^k$ at which the minimal risk is attained: $\rho(p_x; d^*) = \rho^*(p_x) := \inf\limits_{d \in \mathbb{R}^k} \rho(p_x; d)$.

For $k = 1 \Leftrightarrow D = \Theta = \mathbb{R}$ (W is a scalar) we often choose LF of the form $L(w; d) = a|w - d|^b$, where $a, b > 0$. Here we consider $b = 2$ and $b = 1$ cases.

Quadratic LF $L(w; d) = a(w - d)^2$.

At first, we provide (vague) reasonings in favour of this very LF. Let $L(w; d)$ depend on $w - d$ only, i.e. $L(w; d) = \Lambda(w - d)$. If $\Lambda(z) \geqslant 0$ is doubly differentiable and $\Lambda(0) = 0$, then its expansion in Taylor series to within the 2nd degree terms $\Lambda(w - d) \approx a_0 + a_1(w - d) + a_2(w - d)^2$ (when d is close enough to w, higher degree terms are negligible). $\Lambda(0) = 0 \Rightarrow a_0 = 0$, and it follows from $\Lambda(z) \geqslant 0$ that $a_1 = 0$ (otherwise $\Lambda(w - d) = (w - d)\big[a_1 + a_2(w - d)\big]$ would be < 0 either for $d < w$, or for $d > w$, close enough to w). Therefore we obtain $L(w; d) \approx a_2(w - d)^2$.

Choosing (scaling) the units system, we can assume $a = 1$.

Then the risk $\rho(p_x; d) = \int_{\mathbb{R}} (w - d)^2 p_x(w)\mu(dw) = \mathrm{E}(W - d)^2 = \mathrm{E}\,W^2 - 2d\,\mathrm{E}\,W + d^2$ ($\mathrm{E}\,W^2 < \infty$ is assumed). Note that the mean here is determined, in general, by posterior distrib. of W (density $p_x(w)$). Since $\rho(p_x; d) = \mathrm{Var}\,W + (\mathrm{E}\,W - d)^2$, we have $\rho(p_x; d) \to \min = \mathrm{Var}\,W$ when $d = \mathrm{E}\,W$; thus BE $d^* = \mathrm{E}\,W$.

1. No observations: $p_x(w) = p(w)$, — we consider the distrib. of W to be the prior one. By this distrib. we calculate BE $d^* = \mathrm{E}\,W$ and BR $\rho^* = \mathrm{Var}\,W$.

2. Now, let there be observations — RV $\xi = (\xi_1; ...; \xi_n)$ that takes values $x = (x_1; ...; x_n)$. Accordingly, conditional density $f(x|w)$ of ξ's distrib. under the condition $W = w$ and posterior density $p(w|x)$ of W's distrib. under condition $\xi = x$ appear.

At $\xi = x$ BE $d^* = \mathrm{E}\{W|x\} = \mathrm{E}\{W|\xi = x\} = \int\limits_{\mathbb{R}} wp(w|x)\mu(dw)$.

For any $x \in X$ the risk $\rho^*(p_x) = \mathrm{Var}\{W|x\}$, and *Bayes risk* (BR) is the mean $\rho^* = \rho^*(p) = \mathrm{E}\,\mathrm{Var}\{W|\xi\} = \mathrm{E}\,\rho^*(p_\xi)$, where the expectation is determined by ξ ($\rho^*(p_\xi)$ is the function in RV ξ).

But the distrib. of ξ depends on W, isn't there a "vicious circle"?

REM. 1. In the expression for BR, $\mathrm{E}\,\mathrm{Var}\{W|\xi\}$, the density of distrib. of ξ is the *martingale* (generalized) density $g(x) = \int\limits_{\mathbb{R}^k} f(x|w)p(w)\mu(dw)$.

REM. 2. To obtain BR $\mathrm{E}\,\mathrm{Var}\{W|\xi\}$ we can apply the equality

$$\mathrm{Var}\,W = \mathrm{E}\,\mathrm{Var}\{W|\xi\} + \mathrm{Var}\,\mathrm{E}\{W|\xi\} \Rightarrow \mathrm{E}\,\mathrm{Var}\{W|\xi\} = \mathrm{Var}\,W - \mathrm{Var}\,\mathrm{E}\{W|\xi\}$$

EX. Let ξ_1, ..., ξ_n be a sample from Poisson distrib., whose parameter W is unknown and has prior Γ distrib. with parameters $(\theta; \nu)$. We have the realization of the sample: x_1, ... x_n. Let us find BE d^* of parameter W and BR $\rho^*(p)$, assuming $L(w; d) = (w - d)^2$.

It follows from the Theor. established earlier that the posterior distrib. of W is also the Γ distrib. with param-s $\theta' = \theta + n$, $\nu' = \nu + y$, де $y = \sum\limits_i x_i$. We also know that for $\eta \sim \gamma(\theta; \nu)$: $\mathrm{E}\,\eta = \frac{\nu}{\theta}$, $\mathrm{Var}\,\eta = \frac{\nu}{\theta^2}$.

Due to results obtained above, BE $d^* = d^*(x) = \mathrm{E}\{W|x\} = \mathrm{E}\,\eta$, where $\eta \sim \gamma(\theta'; \nu')$; so $d^* = \frac{\nu'}{\theta'} = \frac{\nu+y}{\theta+n}$.

For $\xi = x$ the risk $\rho^*(p_x) = \mathrm{Var}\{W|x\} = \frac{\nu'}{(\theta')^2}$. Taking into account $\mathrm{E}\{\xi_i|W\} = W$ and $\mathrm{E}\,W = \frac{\nu}{\theta}$, we have BR $\rho^*(p) = \mathrm{E}\,\mathrm{Var}\{W|\xi\} =$

$$= \mathrm{E}\Big\{\frac{\nu+\sum\limits_i \xi_i}{(\theta+n)^2}\Big|\xi\Big\} = \frac{1}{(\theta+n)^2}\big[\nu + n\,\mathrm{E}\{\xi_i|\xi\}\big] = \frac{\nu}{(\theta+n)^2}\big(1 + \frac{n}{\theta}\big) = \frac{\nu}{\theta(\theta+n)}$$

Suppose now that each observation has the cost $c > 0$ and we choose the sample size n. Then the total (Bayes) risk is $\rho^*(p) + nc$. To minimize it, consider $R(n) = \rho^*(p) + nc = \frac{\nu}{\theta(\theta+n)} + nc$; its derivative $R'(n) = -\frac{\nu}{\theta} \cdot \frac{1}{(n+\theta)^2} + c = \frac{1}{\theta(n+\theta)^2}\big[c\theta(n+\theta)^2 - \nu\big]$, and $R'(n) = 0 \Leftrightarrow c\theta(n+\theta)^2 = \nu \Leftrightarrow n = \sqrt{\frac{\nu}{c\theta}} - \theta$. This is a p. of min; at the same time, the condition $n \in \mathbb{Z}_+$ must hold.

Hence, if $n_0 := \sqrt{\frac{\nu}{c\theta}} - \theta \leqslant 0$, we take $n = 0$, that is, estimate W by prior distrib. without observations; if $n_0 > 0$, we take, as n, the integer closest to n_0.

REM. If we let LF $L(w; d) = \frac{1}{w}(w - d)^2$, then for $\forall x$ such that $\nu + y > 1$: BE $d^* = \frac{\nu+y-1}{\theta+n}$, BR $\rho^*(p) = \frac{1}{\theta+n}$, and the optimal sample size is $n \approx n_0 = \frac{1}{\sqrt{c}} - \theta$.

LF proportional to absolute error $L(w;d) = a|w-d|$.

PROP. For such LF the BE d^*, whatever distrib. W has, is d that minimizes $\mathrm{E}\,|W-d|$.

THEOR. Let $\mathrm{E}\,|W| < \infty$. $d^* \in \mathbb{R}$ is BE of W, $\mathrm{E}\,|W-d^*| = \inf\limits_d \mathrm{E}\,|W-d|$, **iff** d^* is a median of distrib. of W: $P\{W \leqslant d^*\} = P\{W \geqslant d^*\} = \frac{1}{2}$.

This is implied by the fact that for $d \neq d^*$, where d^* is a median of distrib. of W, and $W \sim F_W(\cdot)$:

$$\mathrm{E}\,|W-d| = \mathrm{E}\,|W-d^*| + 2\begin{cases} \int\limits_d^{d^*}(x-d)F_W(dx), & d < d^*, \\ \int\limits_{d^*}^{d}(d-x)F_W(dx), & d > d^*. \end{cases}$$

(where, obviously, $\int(\ldots) \geqslant 0$). Indeed, for e.g. $d < d^*$: $\mathrm{E}\,|W-d| - \mathrm{E}\,|W-d^*| =$

$$= \int\limits_{-\infty}^{d}(d-d^*)F_W(dx) + \int\limits_d^{d^*}(2x-d-d^*)F_W(dx) + \int\limits_{d^*}^{+\infty}(d^*-d)F_W(dx) =$$

$$= (d-d^*)P\{W \leqslant d\} + \int\limits_d^{d^*}\big[(2x-2d)+(d-d^*)\big]F_W(dx) - (d-d^*)P\{W \geqslant d^*\} =$$

$$= 2\int\limits_d^{d^*}(x-d)F_W(dx) + (d-d^*)\big[P\{W \leqslant d^*\} - P\{W \geqslant d^*\}\big] = 2\int\limits_d^{d^*}(x-d)F_W(dx)$$

REC. Every abs. cont. prob. distrib. has a median, which may be non-unique.

EX. Let ξ_1, ..., ξ_n be a sample from $N_{W;(r)}$. We need to estimate W when LF $L(w;d) = a|w-d|$, having the sample realization x_1, ..., x_n, if the prior distrib. of W is $N_{a;(\tau)}$.

Adjusting the units system, we assume $a = 1$. By earlier Theor., the posterior distrib. of W is $N_{a';(\tau')}$, where $a' = \frac{\tau a + nr\overline{x}}{\tau + nr}$, $\tau' = \tau + nr$. The single median of N distrib. coincides with its mean (and mode), so $d^* = d^*(x) = a'$.

The risk $\rho^*(p_x) = \mathrm{E}\,|W - d^*(x)|$. Posterior distrib. of $W - d^*(x)$ is $N_{0;(\tau')}$, which does not depend on x; we also suppose it to be known that for $\eta \sim N_{0;(\tau)}$: $\mathrm{E}\,|\eta| = \sqrt{\frac{2}{\pi\tau}}$. Thus, regardless of x, $\mathrm{E}\,|W - d^*(x)| = \sqrt{\frac{2}{\pi\tau'}}$, hence BR $\rho^*(p) = \sqrt{\frac{2}{\pi(\tau+nr)}}$ as well.

If the cost of one observation is $c > 0$, then the total risk $\rho^*(p) + nc \to \min$ for $n \approx n_0 = \frac{1}{\sqrt[3]{2\pi rc^2}} - \frac{\tau}{r}$. In particular, as $\tau \to 0$, the limit value, $\frac{1}{\sqrt[3]{2\pi rc^2}}$, is the optimal number of observations when there is "not much" prior information about W (prior distrib. of W is "spread" over $\mathbb{R}$).

Estimation of vectorial parameter.

Now let $W = (W_1; ...; W_k)^T$, $k \geqslant 2$. For such problems we usually apply the "quadratic" LF

$$L\colon \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}_+\colon L(w; d) = (w - d)^T A(w - d)$$

where the matrix $A = \|a_{ij}\|_{1\leqslant i,j\leqslant k}$ is symmetric and non-negative definite ($\forall x = (x_1; ...; x_k)^T \in \mathbb{R}^k\colon \langle x^T; Ax\rangle \geqslant 0$).

If A is positive definite ($x \neq \theta_k \Rightarrow \langle x^T; Ax\rangle > 0$), then for $d \neq w$ (for any deviation of the estimate from true value of the parameter) the loss is positive.

If A is not positive definite, then for some $d \neq w$ the loss can be 0 (e.g. if we estimate only a part of coordinates of W, and the rest of them do not matter).

Let there exist $\mathrm{E}\, W = \mu$ and the *covariation matrix*

$$\mathrm{Cov}\, W = C = \|\, \mathrm{E}\big[(W_i - \mathrm{E}\, W_i)(W_j - \mathrm{E}\, W_j)\big]\,\|_{1\leqslant i,j\leqslant k}$$

For a given distrib. of W, its BE is $d \in \mathbb{R}^k$ that minimizes

$$\mathrm{E}\, L(w; d) = \mathrm{E}\big[(W - d)^T A(W - d)\big] =$$
$$= \mathrm{E}\big[\big((W - \mu)^T + (\mu - d)^T\big)\big] A\big((W - \mu) + (\mu - d)\big) =$$
$$= \mathrm{E}\big[(W - \mu)^T A(W - \mu)\big] + (\mu - d)^T A(\mu - d)$$

(other 2 summands are equal to 0 by properties of expectation of random matrices and because of $\mathrm{E}(W - \mu) = \theta_k$). In this expression, there is no d under E sign, and, since A is non-negative definite, we have $(\mu - d)^T A(\mu - d) \geqslant 0$. Thus d is BE **iff** $(\mu - d)^T A(\mu - d) = 0$.

If A is positive definite, then $d = \mu$ is the unique BE. Otherwise, there may be other BEs d.

PROP. $\mathrm{E}\big[(W - \mu)^T A(W - \mu)\big] = \mathrm{tr}(AC)$ (sum of A's diagonal elements).
That is, for any BE, the mean loss is $\mathrm{tr}(AC)$.

PROP. Let the sample $\xi = (\xi_1; ...; \xi_n)$ with PVs $x = (x_1; ...; x_n)$ have the conditional density of distrib. $f(x|w)$ for $W = w$. Then BE $d^* = \mathrm{E}\{W|x\}$ and BR $\rho^*(p) = \mathrm{tr}(A\, \mathrm{E}\, \mathrm{Cov}\{W|\xi\})$ ($\mathrm{Cov}\{W|x\}$ is the covariation matrix of posterior distrib. of W under the condition $\xi = x$).

REM. For the LFs that are proportional to absolute error, there are (by now) no analogous results due to absence of a "standard" definition of a multidimensional distribution's median.

“And the future has passed”

We have given a glance at some basics of linear programming, game theory, and optimal statistical decisions.

Almost all notions and results presented were known as early as 1950–1960, some even earlier, — look through, for example, [2], [5], and [3].

If we compare these themes with tree branches, then, firstly, we have looked at them only from some points of view, and, secondly, they have since extended, ramified, other branches grew from them.

• For the transport LPP, there is a special method of solving it, somewhat simpler than the general simplex-method. This method uses the analogous *transport table*. The criterion of solution existence is the balance condition (per time unit, the production is equal to the consumption). See [10, 1.9].

• *Network models* are directed graphs, which describe various “flows” (of resources, data, money) and require algorithms that, in some sense, optimize such flow. Among these, there are algorithms to find minimal skeleton, shortest or critical path, maximal flow, to minimize the cost of the flow. Similar to the transport LPP, methods exist here that are more effective than SM, which take into account the specifics of a problem. See [7, 6].

• LPPs where all or some variables must take integer values lead to *integer programming*. The main methods to solve such problems are *branch and bound* and *cutting planes*. We switch to them after transforming the problem into a “usual” LPP with continuous variables. See [7, 9].

• To make optimal decisions in multistep processes, where these decisions are able to change in time, we apply *dynamic programming*, in particular, when variables take integer values. The foundation of this method is the *R. Bellman's Optimality Principle*: to make optimal decision at a given step, we need optimality of decisions at all preceding steps. See [10, 6].

• *Markov decision processes* apply the methods of dynamic programming to the problems where Markov chain defines the probabilities of transitions between states, and payoffs are defined by *payoff matrix* at transition from one state to another. Matrices of transition probabilities and payoffs depend on decisions made. The goal is to maximize the mean payoff. See [8, 3].

• Searching for a point-set of parameters/variables, at which a function attains its “optimal”, the largest or the smallest, value, is the *optimization* problem. “Classical” iterative methods to solve such problems are Jacobi, Lagrange, Kuhn-Tucker, Newton-Raphson etc. Add indeterminacy of parameters' change between iterations, and we come to *stochastic gradient descent* and its variations. See [1].

• *Queueing theory* deals with determining the optimal parameters of the so-called *queueing systems*, such as: number of service channels, service rate, queue discipline, — so that certain statistical characteristics of the system's functioning (mean length of queue, time spent in queue, or probability of waiting for service) are as good as possible. See [4, 1].

• If, in a game, some players can act in concord during several moves, we have a *cooperative* game. This notion is applied to the 2-player games as well. For example, before the game begins, they can agree on something, and then each player decides, independently, if she/he keeps to the agreement. See [5, 6].

• If players decide to act in such a way as to get the result that is the best for them all, they nevertheless can have disagreements on what (alternative) such result consists in. To come to agreement, they begin negotiations; if we assume that during these negotiations the players, in turn, make proposals, then we have a *bargaining*. In the simplest case 2 players, in turn, give proposals, until one player agrees with a proposal of the other player. See [6, 7.2].

• Decisions can be made not by a single person or "side", they can be the outcome of independent decisions of several participants, for instance, of a voting. One of the principal problems in such situations is the optimal joining of individual preferences so that the common decision is also, in a sense, "good". *Arrow's Impossibility Theorem* states that some "natural" requirements for such (desirable) decision system are inconsistent. See [5, 14].

• The games with infinite set of pure strategies of a given player may not have the value in the sense that was defined for the finite case; one typical example is the "choose the larger number" game. For some games of this type it is, however, possible to generalize the notion of a value. One important case is the so-called "games on unit square" (P1 chooses $x \in [0; 1]$ and P2 chooses $y \in [0; 1]$). The games *with time choice* belong here too. See [5, A7].

"There are more things …"

Abbreviations

DEF. — Definition.
REC. — Recalling.
PROP. — Proposition.
REM. — Remark.
EX. — Example.
COR. — Corollary.
LEMMA. — Lemma.
THEOR. — Theorem.
◀ — proof begins
▶ — proof ends
iff — if and only if
w.r.t. — with respect to
$\otimes$ — inconsistency

References

[1] C.L. Byrne: Applied Iterative Methods. — A K Peters, Ltd., 2008.

[2] G. Dantzig: Linear Programming and Extensions. — Princeton University Press, 1963.

[3] M.H. DeGroot: Optimal Statistical Decisions. — McGraw-Hill Company, 1970.

[4] D. Gross, J.F. Shortle, J.M. Thompson, C.M. Harris: Queueing Theory. 4th ed. — Wiley, 2008.

[5] R.D. Luce, H. Raiffa: Games and Decisions. Introduction and critical survey. — Wiley, 1957.

[6] M.J. Osborne, A. Rubinstein: A Course in Game Theory. — MIT, 1994.

[7] H.A. Taha: Operations Research: An Introduction. 7th ed. — Pearson Education, Inc., 2003.

[8] J.N. Webb: Game Theory. Decisions, Interaction and Evolution. — Springer, 2007.

[9] О.Ю. Зайченко, М.П. Зайченко: Дослідження операцій. Збірник задач. — Слово, 2007.

[10] М.П. Моклячук, Р.Є. Ямненко: Теорія вибору та прийняття рішень. — ВПЦ «Київський університет», 2013.

`http://probability.univ.kiev.ua/userfiles/mmp/decion-lectures_6_10_2014.pdf`