$$\Omega_1(z)=\int_a^b z^2(x)\,dx, \qquad \Omega_2(z)=\int_a^b (dz/dx)^2\,dx.$$

We usually replace them with one stabilizer of the form $\Omega=\beta_1\Omega_1+\beta_2\Omega_2$, where the coefficient β_1 is chosen arbitrarily $(0<\beta_1<1,\ \beta_2=1-\beta_1)$. The method of solving multicriterial problems indicated in the previous paragraph enables us to consider, instead of (6), problem (1) with several criteria, for example $\Phi_1=\rho(A(z),u)$, $\Phi_2=\Omega_1(z)$, $\Phi_3=\Omega_2(z)$.

Secondly, in some problems the choice of the distance ρ in the space u contains great arbitrariness. For example, such will be the case when there is a set of equations connecting various physical quantities (different sizes, different scales, and different errors). Of course, we can always "lay responsibility" on the coefficients c_j in the expression

$$\rho(u,u')=\Big[\sum_{j=1}^{m} c_j(u_j-u_j')^2\Big]^{1/2},$$

but in reality nothing is known of these c_j. Therefore instead of the single-criterial problem (4) it is more reasonable to consider problems with several criteria, for example

$$\Phi_j=|A(z)_j-u_j|\to\min, \qquad 1\leqslant j\leqslant m.$$

It is easy to see that if a quasisolution $z=z^*$ exists, such that $\rho(A(z^*),u)=0$, then it is not quite so terrible to choose unsuccessful c_j: the problem will be badly considered, but the solution will finally be obtained. However, in real problems usually

$$\min_{z\in G}\rho(A(z),u)>0.$$

An unfortunate choice of c_j can greatly damage the calculated "best" solution.

In conclusion the author thanks the participants in the seminar on ill-posed problems and its leader V. Ya. Arsenin for useful discussions of the problems raised here.

REFERENCES

1. TIKHONOV A.N. and ARSENIN V.YA., Methods of solving ill-posed problems. Moscow: Nauka, 1979.
2. PODINOVSKII V.V. and NOGIN V.D., Paréto-optimal solutions of multicriterial problems. Moscow: Nauka, 1982.
3. IVANOV V.K., VASIN V.V. and TANANA V.P., The theory of linear ill-posed problems and its applications. Moscow: Nauka, 1978.
4. RAKHMATULINA A.KH., Ill-posed problems and methods of solving them. Preprint IPMatem. AN SSSR. Moscow,1, 1972.
5. MOROZOV V.A., Regularization of ill-posed problems and the choice of regularization parameter. Zh. vychisl. mat. mat. fiz., 6,1,170-175,1966.
6. SOBOL I.M. and STATNIKOV R.B., The choice of optimal parameters in problems with many criteria. Moscow: Nauka, 1981.
7. SOBOL I.M., Estimate of the accuracy of the simplest multidimensional search. Dokl. AN SSSR, 266,3,569-572,1982.

Translated by H.Z.

U.S.S.R. Comput.Maths.Math.Phys.,Vol.26,No.3,pp. 104-113,1986 0041-5553/86 $10.00+0.00
Printed in Great Britain

AN ASYMPTOTIC ESTIMATE OF THE AVERAGE NUMBER OF STEPS OF THE PARAMETRIC SIMPLEX METHOD*

A.M. VERSHIK and P.V. SPORYSHEV

The estimate, announced in /1/, of the average number of steps of the parametric version of the simplex method for solving linear programming problems with respect to some natural class of statistics in the space of the problems is solved. If the number of variables in the problem equals n, and the number of limitations of the equality type equals k, then for the avarage number of steps $s(n,k)$ the following estimate holds:

$$s(n,k)\leqslant\frac{(k+1)^{1/2}}{2}(\pi\ln n)^{k/2}+O((\ln n)^{(k-1)/2}), \qquad n\to\infty.$$

The Grassman approach to similar problems, which is important in iteself, is described.

*Zh.vychisl.Mat.mat.Fiz.,26,6,813-826,1986.

1. Introduction.

The problem of what is rate of convergence of the basic method of solving linear programming problems - the simplex method and its versions - emerged as far back as the first papers on this problem (see /2, 3/). At the same time it was understood that from one aspect the method is practical, but amongst all the problems of given dimensions there exist those for which the simplex method reduces to complete inspection of the possible basis plans. Appropriate examples were later constructed (see, in particular, /4/). Obviously, that is the case with any modification of the simplex method. However, computation specialists understood that these phenomena are not characteristic and in standard problems the simplex method works well. Time was necessary, and exact terms were obtained for the correct formulation of the problems and the justification of this thesis. In recent years several papers have appeared, simultaneously and independently, in which similar (but not matching) practical results are presented. For a more detailed analysis and historical remarks see Sect. 6.

The main result of this paper is based on the representation of linear programming problems like Grassman's manifold, which is interesting in itself, and on the concepts of integral geometry. This method ("the Grassman approach") was proposed by A.M. Vershik and consists of the following. Consider a set of all k-dimensional subspaces of an n-dimensional space, i.e. Grassman's manifold; then we can connect the linear programming problem to each of them, and the dual problem (see Sect. 2, Eq.(2.5)) to its orthogonal addition. The space of the problems was changed into a Grassman manifold; there is an isolated invariant measure in it, and the number of steps of the simplex method is a functional in it. Using the methods of integral geometry, we can estimate the asymptotic behaviour of the average number of steps Such is the general scheme. However, it is required to obtain in advance the exact geometrical meaning of the number of steps. This is a complex problem: there is still no simple description of a set of these permissible plans which occur in the usual simplex-method procedure for the specified initial plan. It is convenient to use the so-called parametric simplex method (Sect. 3), not the canonical (or dual) simplex method. It enables us to describe the permissible plans considered in the method as optimal for some values of the parameter in a single-parametric family of problems. It is interesting that if we use the Kantorovich-Rubinshtein geoemtrical interpretation of linear programming problems, the number of steps is the number - of special form - of the bounds of some subsidiary polyhedral set (Sect. 5). A similar problem is considered in convex integral geometry. But the functional remains cumbersome in this form also; a simplification is obtained if, instead of parametrically optimal plans, we enumerate the parametrically dual permissible plans. This sets an upper estimate to the number of steps, and the asymptotic behaviour of the new functional is expressed in terms of the asymptotic behaviour of a comparatively simple integral (see Sect.4), which gives the final result.

An outline of this proof is presented in /1/; it is somewhat improved below and a different parametric method is used. At almost the same time as /1/ Smale's estimate /5/ appeared. Although it makes wider assumptions regarding the statistics, the method itself, as shown in /6/, is coarser and cannot give the same powers of $\ln n$ as is presented in this paper (see Sect. 6). In /7/ the estimate is better, but for another parametric method (the parameter enters polynomially and not linearly). The problem of estimating the number of steps of the canonical simplex method remains unsolved. At the same time, the order-estimate is obviously nearly accurate for the given parametric method in this paper.

2. Universal formulations of dual problems. The Grassman space of the problems.

We shall begin from the usual formulation of dual linear programming problems in canonical form:

$$\max \{\langle c, x\rangle | x\in R^m,\ Ax\leqslant b,\ x\geqslant 0\}, \tag{2.1a}$$

$$\min \{\langle b, p\rangle | p\in R^k,\ A^t p\geqslant c,\ p\geqslant 0\}. \tag{2.1b}$$

Here A is a $k\times m$ matrix, $b\in R^k, c\in R^m$, t is the sign of transposition, $\langle,\rangle$ is the scalar product in R^m and R^k, and the inequalities are understood in the sense of coordinate ordering. Henceforth m and k are fixed; $n=m+k$.

We shall write problems (2.1) in another way which is more convenient for later. Consider the space R^{n+2} with the fixed basis $e^0,\ldots,e^{n+1}$, numbered $0,1,\ldots,n+1$; we shall insert R^k and R^m into R^{n+2} as coordinate subspaces withthe numbers $1,2,\ldots,k$ and $k+1,\ldots,n$ respectively. Using (A, b, c) from (2.1) we shall construct the $(k+1)\times(n+2)$ -matrix D:

$$D = D[\{0, 1, \ldots, k\}, \{0, 1, \ldots, n+1\}] = \left\|\begin{array}{c|c|c} 1\ 0\ldots 0 & -c^T & 0 \\ \hline 0\ 1\ldots 0 & & \\ \vdots\ \ddots & A & b \\ 0\ 0\ldots 1 & & \end{array}\right\|, \tag{2.2}$$

and shall consider the linear programming problems in R^{n+2}:

$$\max \{x_0 | x\in R^{n+2};\ Dx=0;\ x_i\geqslant 0,\ i=1,2,\ldots,n;\ x_{n+1}=-1\}, \tag{2.3a}$$

$$\min \{y_{n+1} | y=D^t z;\ z\in R^{k+1};\ y_i\geqslant 0,\ i=1,2,\ldots,n;\ y_0=1\}. \tag{2.3b}$$

It is obvious that (2.3) and (2.1) are identical. However, the formulation of (2.3) is entirely symmetric: both the direct and dual problems are formulated in the same space, and the limitations in both problems are of the equality type.

We shall consider the matrix D as the matrix of the operator from R^{n+2} into R^{k+1}, denoted by the same letter D, and D^t as the matrix of the conjugate operator D^* from R^{k+1} into R^{n+2} (since the basis if fixed, R^p is identified with the conjugate space).

We shall reformulate (2.3), denoting the kernel of the operator F by $\operatorname{Ker} F$, and the following form by $\operatorname{Im} F$:

$$\max\{x_0 | x \in \operatorname{Ker} D;\ x_i \geqslant 0,\ i=1, 2, \ldots, n;\ x_{n+1}=-1\}, \tag{2.4a}$$

$$\min\{y_{n+1} | y \in \operatorname{Im} D^*;\ y_i \geqslant 0,\ i=1, 2, \ldots, n;\ y_0=1\}. \tag{2.4b}$$

We can now dispense with the special form (2.2) of the matrix D and consider it to be entirely arbitrary. Henceforth, it will only be necessary that the determinant of the submatrix $\bar{D}=D[\{0, 1, \ldots, k\}, \{0, 1, \ldots, k\}]$ should be non-vanishing. In this case, multiplying D from the left by $\bar{D}^{-1}$ will give the form (2.2). We shall call the set $\mathcal{D}$ of all the $(k+1)\times(n+2)$ -matrices D with a non-zero minor of the form shown the matrix space of linear programming problems. This space of the problems is only formally distinct from the space of the triple (A, b, c), which is usually considered. The advantages of form (2.4) will be seen later, but now we take the following step in the direction of an invariant formulation. We shall denote it by $E_D=E=\operatorname{Im} D^*$, $E^D=\operatorname{Ker} D$; and the other is the subspace in R^{n+2}, whilst $E^D=E_D{}^{\perp}=E^{\perp}$, where the sign $\perp$ is an orthogonal addition with respect to the natural scalar product in R^{n+2}. We can now rewrite (2.3) thus, denoting the cone $\{x \in R^{n+2} | x_i \geqslant 0,\ i=1, 2, \ldots, n\}=Q$:

$$\max\{x_0 | x \in E^{\perp} \cap Q,\ x_{n+1}=-1\}, \tag{2.5a}$$

$$\min\{y_{n+1} | y \in E \cap Q,\ y_0=1\}. \tag{2.5b}$$

Here we can assume E is an entirely arbitrary $(k+1)$-dimensional subspace in R^{n+2}. Both problems now have an identical form and are symmetric with respect to the conversion from E to $E^{\perp}$ with the replacement $x_0 \to -y_{n+1}$, $x_{n+1} \to -y_0$. We can say that (2.5) is a limitation on the subspaces E, $E^{\perp}$, respectively, of the trivial universal problems

$$\max\{z_0 | z \in Q,\ z_{n+1}=-1\}\ \ (=\infty),\ \min\{z_{n+1} | z \in Q,\ z_0=1\}\ \ (=-\infty).$$

We recall that the set of all q-dimensional subspaces of the p-dimensional space is called a Grassman manifold and is denoted by $G(p, q)$. We shall call problems (2.5) a Grassman formulation of dual linear programming problems. Now, already, the parameter of the problem is the subspace E, i.e. the element $G(n+2, k+1)$ (or $E^{\perp}$, i.e. the element $G(n+2, m+1)$). We shall call Grassman's manifold the Grassman space of linear programming problems, bearing in mind the correspondence of the subspaces and linear programming problems in the form (2.5). When changing from the matrix to the Grassman space of the problems we identify all the matrices D with one and the same kernel (or form D^*). This identification is not essential, since any two matrices D with a common kernel are distinguished by a non-singular transformation in the form. However, when examining these or other methods and the properties of linear programming problems, it is necessary to verify the correctness with respect to this identification. This is done in Sect. 3 for the parametric method. We emphasize the importance of the formulation (2.5) for the general theory of linear programming problems. A similar approach, in which the problem (e.g., linear programming) or geometric object (e.g. a polyhedron) are considered as the limitation (or projection, see Sect. 5) of the universal problem or geometric object, is called the Grassman approach. It is useful in problems of linear algebra, convex analysis, etc. (comp. /8/).

We shall turn to the statistics in the space of the problems. Usually some or other statistics in the space of the set (A, b, c) (see (2.1) are considered. We can transpose it to the matrix D (see (2.2)) or directly introduce it in $\mathcal{D}$, bearing in mind the formulations (2.3) and (2.4). When changing to a Grassman space of the problems we can specify the statistics directly in $G(n+2, k+1)$. At the same time there is some isolated measure in $G(p, q)$ which will be used later. We recall its description.

In Grassman's manifold $G(p, q)$ the orthogonal group $O(p)$ operates transitively, and at the same time we can represent $G(p, q)$ as a homogeneous space of the group $O(p)$. It is well known that a unique $O(p)$-invariant normalized measure exists in $G(p, q)$, which we will call a Grassman measure and denote it by $\mu_{p,q}$. The uniqueness follows from the uniqueness of Haar's measure in $O(p)$. The expression for the measure $\mu_{p,q}$ (true, with another normalization) can be obtained, for example, in /9/. For us the following simple proposition is important:

Proposition 1. Suppose φ is the mapping of the space of the $q\times p$ -matrices of rank q in Grassman's manifold $G(p, q)$:

$$\varphi(D)=\operatorname{Im} D^*.$$

Then every probability measure in the matrix space, which is invariant with respect to right-multiplication by any orthogonal p-order matrix, converts into a Grassman measure for the

mapping φ.

The proof of the proposition clearly follows from the uniqueness of the orthogonally-invariant measure.

Corollary 1. Every probability measure in the space of matrices $\mathcal{D}$, for which the rows are statistically independent and the distribution of each of them is orthogonally invariant, becomes - under the influence of φ - a Grassman measure.

In this connection the main result of this paper (see Sect. 4) holds for any such measure in matrix space.

3. Geometry of the parametric method.

To estimate the number of steps of any iterative process we first need to describe this number in an obvious way as a function of the data of the problem and of the initial approximation. Unfortunately, there is still no simple description of this function for the canonical simplex method. All the existing theoretical estimates are described in the so-called parametric simplex methods. Their advantage is that the basic plans which can be found during the iterations permit of a simple description: they are optimal for some subsidiary problem with a parameter. Here Lemke's parametric method - also called "Danzig's self-dual algorithm" /3/ (for details see /10/) - is used. Another, less convenient, parametric method was used in /1/.

We shall introduce the parameter $t\in\mathbf{R}_+$ into problem (2.1a):

$$\max\{\langle c-tq_m, x\rangle \mid x\in R^m,\ Ax\leqslant b+tq_k,\ x\geqslant 0\}, \tag{3.1}$$

where $q_m=(1,\ldots,1)^T\in R^m$, $q_k=(1,\ldots,1)^T\in R^k$. Lemke's method consists of a sequential review of the optimal plans in problem (3.1) for t, changing from $+\infty$ to 0. When $t=0$ we obtain problem (2.1a): if t is fairly large, then the optimal plan is $x=0$ and we can take it as the initial plan. The process of converting, using the optimal plans in (3.1), from one t to a similar one is analogous to the simplex structure (see /10/). Finally we either obtain the optimal plan in (2.1a) when $t=0$, or we establish its unimportance if for some $t>0$ there is no optimal plan. Naturally, we can introduce the parameter in a more complex way (see /7/) and Sect. 6 below). The most common method in some sense must be universal deformation for problems (2.1).

The introduction of the parameter t makes the matrix D from (2.2)-(2.4) linearly dependent on t:

$$D(t)=D+\bar{D}R(t),\qquad R(t)=\left\|\begin{array}{c|c} 0 & \begin{array}{c} tq_m^T \\ 0 \end{array} \\ \hline \begin{array}{c} 0 \\ tq_k \end{array} & \end{array}\right\|$$

(if D has the form (2.2), then $\bar{D}=\mathrm{Id}$). A one-parametric family of problems is obtained, which is written in a form similar to (2.4):

$$\max\{x_0 \mid x\in\operatorname{Ker}D(t);\ x_i\geqslant 0,\ i=1,2,\ldots,n;\ x_{n+1}=-1\}, \tag{3.2a}$$

$$\min\{y_{n+1} \mid y\in\operatorname{Im}D(t)^*;\ y_i\geqslant 0,\ i=1,2,\ldots,n;\ y_0=1\}. \tag{3.2b}$$

We shall introduce the notation $N=\{1,2,\ldots,n\}$, $K=\{1,2,\ldots,k\}$, $M=N\setminus K$, $\mathcal{J}=\{J\subset N \mid |J|=k,\ J\neq K\}$.

Below we present the definitions necessary to work with the parametric method. The definitions are given for problem (2.3a). Their reformulations for problems (2.1), (2.4), (2.5) are obvious, there fore if necessary the concepts introduced will also be used as it applies to them.

We shall call the set $J\in\mathcal{J}$ a basis set for problem (2.3) if the limitation $\operatorname{Ker}D(=E^\perp)\cap\{x\in R^{n+1} \mid x_i=0,\ i\in\{1,2,\ldots,n+1\}\setminus J\}=\{0\}$. This definition is equivalent to the usual one.

The basis set is called permissible (or doubly permissible) if the vector $x\in R^{n+1}$ exists, (correspondingly, $y\in R^{n+1}$), satisfying the limitations of problem (2.3a), whilst $x_i=0$ for $i\in N\setminus J$ (correspondingly, satisfying the limitations of (2.3b), whilst $y_j=0$ for $j\in J$). The basis set, which is simultaneously permissible and doubly permissible, is called optimal).

The basis set for problem (2.3a) is called parametrically permissible (correspondingly, parametrically doubly permissible and parametrically optimal (in /1/ - fully permissible)), if $t\geqslant 0$ exists, for which it is permissible for problem (3.2a) (correspondingly, double permissible and optimal). We can also give similar definitions for other parametric methods.

Lemke's method for the initial problem (2.3a) consists of a sequential inspection with respect to t of the parametrically optimal sets. These sets, being basis sets, are generally neither permissible nor doubly permissible (a drawback of the method). The set K is parametrically optimal (for large t), and we shall take it as the initial set; the optimal set is parametrically optimal when $t=0$.

Not every parametrically optimal set is necessarily encountered in some step of Lemke's method; this will happen in the case of degeneracies. However we have the following

Proposition 2. For an open and everywhere dense set in the space of the given problems (2.3a) the parametric optimality of the set is equivalent to its appearance in some step of Lemke's method.

Proof. It is sufficient to note that for an open and everywhere dense set in the space of the given problems (2.3a) the parametric family of problems (3.2a) is such that the optimal plan is unique in all the range of variation of the parameter t, with the exception of a finite number of points, and at these points there are precisely two optimal basis plans: one for small values of t and one for large values of t.

Corollary 2. For almost all problems (2.3a) with respect to Lebesgue's measure in the matrix space $\mathcal{D}$ the number of steps in Lemke's method is the same as the number of parametrically optimal sets.

We shall now explicitly describe parametrically permissible and parametrically doubly permissible sets.

Proposition 3. Suppose $J\in\mathcal{J}$ is the basis set. The set J is permissible in (3.2a) when and only when

$$\operatorname{Ker} D\cap\{x\in R^{n+2}\mid x_i=0,\ i\in M\setminus J;\ x_i=-t,\ i\in K\setminus J;\quad x_i\geqslant 0,\ i\in M\cap J;\ x_i\geqslant -t,\ i\in K\cap J;\ x_{n+1}=-1\}\neq\varnothing. \tag{3.3}$$

The set J is doubly permissible in (3.2a) when and only when

$$\operatorname{Im} D^*\cap\{y\in R^{n+2}\mid y_i=0,\ i\in K\cap J;\ y_i=-t,\ i\in M\cap J;\ y_i\geqslant 0,\quad i\in K\setminus J;\ y_i\geqslant -t,\ i\in M\setminus J;\ y_0=1\}\neq\varnothing. \tag{3.4}$$

The simultaneous satisfaction of conditions (3.3) and (3.4) for one and the same $t\geqslant 0$ is equivalent to parametric optimality.

The proof amounts to a verification of the definition applied to problem (3.2a).

Corollary 3. Whether the property of the set is parametrically permissible or parametrically doubly permissible depends only on the kernel D. Therefore, these properties are correct with respect to the conversion to the Grassman formulation of problem (2.5).

We shall give a direct geometrical interpretation of parametrically doubly permissible sets. Note that we can give a similar condition for parametrically permissible sets. However, the interpretation of the simultaneous completion of these conditions (i.e. of the parametric optimality) is rather cumbersome due to the connecting parameter t. It is convenient to introduce the following notation. We shall compare the following subspace in R^{n+1} to each set of indices $I\subset\{0, 1, \ldots, n\}$ such that $M\cap I\neq\varnothing$:

$$L(I)=\{z\in R^{n+1}\mid z_i=0,\quad i\in\{0, 1, \ldots, k\}\cap I;\quad z_i=z_j,\ i, j\in M\cap I\}.$$

In the subspace $L(I)$ we shall fix an orthonormalized basis, consisting of the unit vectors of the standard basis e^i, $i\in\{0, 1, \ldots, n+1\}\setminus I$, and of the vector

$$\bar{e}(I\cap M)=r^{-1/2}\sum_{i\in I\cap M}e^i \text{ where } r=|I\cap M|.$$

We shall denote the vector coordinates $z\in L(I)$ in this basis by z_i, $i\in\{0, 1, \ldots, n+1\}\setminus I$ and $\bar{z}$ respectively. We shall introduce a cone into $L(I)$:

$$C(I)=\{z\in L(I)\mid z_i>0,\ \ i\in\{0, 1, \ldots, k\}\setminus I;\ \ z_i>r^{-1/2}\bar{z},\qquad i\in M\setminus I;\ \bar{z}<0\}.$$

From Proposition 3 we obtain

Proposition 4. Suppose $J\in\mathcal{J}$. Then for almost all problems (2.3a) with respect to Lebesgue's measure in the matrix space $\mathcal{D}$ the following statements are equivalent:

a) the set J is a parametrically doubly permissible set;

b) $\operatorname{Im} D^*\cap C(J)\neq\varnothing$.

Thus, the parametric double permissibility of set J is equivalent to the non-emptiness of the intersection of the subspace $\operatorname{Im} D^*$ and the cone $C(J)$. From Proposition 4 we directly obtain

Corollary 4. Suppose $J\in\mathcal{J}$. Then for almost all $E\in G(n+2,\ k+1)$ (using the Grassman measure) the following statements are equivalent:

a) the set J is a parametrically doubly permissible set in problem (2.5a);

b) $E\cap C(J)\neq\varnothing$.

We shall sum up the results of this paragraph in the following theorem. We shall denote the number of steps in Lemke's method for problem (2.3a), or, which amounts to the same in problems (2.4) or (2.5) (in the first two cases $E=\operatorname{Im} D^*$), by $s(E)$.

Theorem 1. For almost all $E\in G(n+2,\ k+1)$ (using the Grassman measure)

$$s(E)\leqslant|\{J\in\mathcal{J}\mid E\cap C(J)\neq\varnothing\}|.$$

Proof. The number of steps for almost all E is the number of steps of the parametrically optimal sets (Corollary 2). But every such set is a parametrically doubly permissible set by definition. Applying Corollary 4, we obtain the necessary result for problem (2.5). But, using Corollary 3, the number of steps in problems (2.3) and (2.4) is the same as in (2.5).

4. A fundamental theorem.

We shall obtain the asymptotic behaviour of an average number - using an orthoinvariant (Grassman) measure in Grassman's manifold - of parametrically doubly permissible sets which, using Theorem 1, will provide an estimate of the number of steps of Lemke's method. We shall denote the number of parametrically permissible sets, distinct from K, for E in problem (2.5a) by $s_1(E)$. The integration with respect to the Grassman measure is denoted by

$$\int_{G(p,\ q)}\ldots d\theta,\qquad s_1(n, k)\equiv\int_{G(n+2,\ k+1)}s_1(E)\,dE.$$

Theorem 2. For fixed k and as $n\to\infty$ we have

$$s_1(n, k) = \frac{(k+1)^{1/2}}{2}(\pi \ln n)^{k/2} + O((\ln n)^{(k-1)/2}).$$

Proof. Suppose $J \in \mathcal{J}$. We shall introduce into R^{n+2} the subspace

$$L^0(J) = \{z \in R^{n+2} \mid z_i = 0, \quad i \in J\}$$

and the cone

$$C^0(J) = \{z \in L^0(J) \mid z_i > 0, \quad i \in \{0, 1, \ldots, n\} \setminus J\}.$$

Suppose χ is a function of the set: $\chi(\alpha) = 1$ if $\alpha \neq \varnothing$ and $\chi(\varnothing) = 0$. We shall present a convenient formula for the functions s_1.

Lemma 1. For almost all subspaces $E \in G(n+2, k+1)$ with respect to the measure $\mu_{n+2, k+1}$ the following relation holds:

$$s_1(E) = \sum_{J \in \mathcal{B}} \chi(E \cap C(J)) = \frac{1}{2} \sum_{J \in \mathcal{B}} \Big(\sum_{j \in \{0, 1, \ldots, n\} \setminus J} \chi(E \cap C(J \cup \{j\})) + \chi(E \cap C^0(J)) \Big).$$

Proof. The first equation in Lemma 1 is directly obtained from Corollary 4. We shall fix the set $J \in \mathcal{J}$ and show that for almost all E

$$\chi(E \cap C(J)) = \frac{1}{2} \Big[\sum_{j \in \{0, 1, \ldots, n\} \setminus J} \chi(E \cap C(J \cup \{j\})) + \chi(E \cap C^0(J)) \Big], \tag{4.1}$$

whence the validity of the second equation in Lemma 1 will also follow. For almost all E (using the Grassman measure)

$$\dim(E \cap L(J)) = \dim E + \dim L(J) - (n+2) = 2.$$

Consequently, the cone $E \cap C(J)$ is two-dimensional. It is obvious that it has almost exactly two one-dimensional bounds, which are intersections of E with the bounds of the leading dimension of the cone $C(J)$. The cones $C(J \cup \{j\})$, $j \in \{0, 1, \ldots, n\} \setminus J$ and the cone $C^0(J)$ are the bounds of the leading dimension of the cone $C(J)$. Thus, Eq. (4.1) is proved.

Using Lemma 1,

$$s_1(n, k) = \frac{1}{2} \sum_{J \in \mathcal{B}} \Big[\sum_{j \in \{0, 1, \ldots, n\} \setminus J} P(J \cup \{j\}) + P^0(J) \Big], \tag{4.2}$$

where

$$P(I) = \int_{G(n+2, k+1)} \chi(E \cap C(I))\, dE, \quad I \subset \{0, \ldots, n\}, \quad M \cap I \neq \varnothing,$$

$$P^0(J) = \int_{G(n+2, k+1)} \chi(E \cap C^0(J))\, dE, \quad J \in \mathcal{B}.$$

We shall fix the set of indices $J \in \mathcal{J}$ and the index $j \in \{0, 1, \ldots, n\} \setminus J$. Suppose $I = J \cup \{j\}$. We shall evaluate the integrals $P(I)$ and $P^0(J)$. Note that the function $\chi(E \cap C(I))$ depends only on the straight line $E \cap L(I)$ (and not on the subspace E itself). Using Chzhen's formula (see /9, p.207/), we have

$$P(I) = \int_{\mathrm{RP}(L(I))} \chi(\theta \cap C(I))\, d\theta,$$

where $\mathrm{RP}(L)$ is a projective space associated with the subspace L, i.e. the space of the straight lines in L (for $L = R^p$ the standard notation is RP_p), and the integration is carried out over the corresponding orthoinvariant measure. In a similar way,

$$P^0(J) = \int_{\mathrm{RP}(L^0(J))} \chi(\theta \cap C^0(J))\, d\theta.$$

Later we will use the following lemma.

Lemma 2. Suppose C is an open convex cone in R^p, $C \neq R^p$, μ is any probability measure in R^p, which is invariant with respect to the effect of the orthogonal group, whilst $\mu(\{0\}) = 0$. Then

$$\int_{\mathrm{RP}_p} \chi(\theta \cap C)\, d\theta = 2\mu(C),$$

where the integration is carried out over the orthoinvariant normalized measure μ_p.

Proof. This lemma is an obvious corollary of the orthoinvariance of the measures μ and μ_p.

We shall use Lemma 2 to calculate $P(I)$. We shall take as the measure μ the Gaussian measure in $L(I)$, i.e. the measure with the density

$$g(z)=[(2\pi)^{-1/2}]^{m+2}\exp\left[-\frac{1}{2}\Big(\sum_{i\in\{0,\,1,\,\ldots,\,n+1\}\setminus I} z_i^2+\bar{z}^2\Big)\right],$$

where $z_i (i\in\{0,\,1,\ldots,\,n+1\}\setminus I)$, $\bar{z}$ are the coordinates of the vector z in the orthonormalized base fixed in $L(I)$ (see the end of Sect.3). Then

$$P(I)=2\int_{C(I)} g(z)\,dz, \tag{4.3}$$

where the integration is carried out over Lebesgue's measure in $L(I)$.

Suppose

$$P_r=2[(2\pi)^{-1/2}]^{m+2}\int_{\Omega}\exp\left(-\frac{1}{2}\sum_{i=1}^{m+2}x_i^2\right)dx,$$

where $1\leqslant r\leqslant m+1$, $\Omega=\{x\in R^{m+2}\mid x_i<0;\ x_i>0,\ i=2,\ldots,r+1;\ x_i>r^{-1/2}x_1,\ i=r+2,\ldots,\ m+1\}$. From the definition of the cone $C(I)$ and from Eq. (4.3) we have the following lemma.

Lemma 3. The relation $P(I)=P_r$, where $r=|I\cap M|$, is valid.

For $P^0(J)$ the calculations are carried out in a similar way. From Lemma 2 with a Gaussian measure as the measure μ and from the equation

$$(2\pi)^{-1/2}\int_0^{\infty} e^{-t^2/2}\,dt=1/2 \tag{4.4}$$

we have

Lemma 4. We have the formula $P^0(J)=2^{-m}$.

We shall fix r, such that $1\leqslant r\leqslant k+1$, and shall calculate the asymptotic behaviour of P_r. Integrating with respect to all the variables, besides x_1, and using Eq. (4.4), we will obtain

$$P_r=2^{-r+1}(2\pi)^{-1/2}\int_{-\infty}^{0} e^{-t^2/2}[\Phi(r^{-1/2}t)]^{m-r}\,dt, \tag{4.5}$$

where

$$\Phi(v)=(2\pi)^{-1/2}\int_v^{\infty} e^{-t^2/2}dt.$$

Lemma 5. For fixed $r\geqslant 1$ and as $n\to\infty$ for the integral

$$I(r,n)=(2\pi)^{-1/2}\int_{-\infty}^{0} e^{-t^2/2}[\Phi(r^{-1/2}t)]^n\,dt$$

the following asymptotic equation holds:

$$I(r,n)=r^{1/2}(r-1)!\,n^{-r}(4\pi\ln n)^{(r-1)/2}[1+O(\ln n)^{-1}].$$

Remark 1. The asymptotic behaviour of the integral $I(r,n)$ has already been calculated for a similar reason in /11, Sect.5/, but the calculations were carried out in a different way.

Proof. Making the change of variable $v=r^{-1/2}t$, we obtain

$$I(r,n)=\left(\frac{r}{2\pi}\right)^{1/2}\int_{-\infty}^{0} e^{-rv^2/2}[\Phi(v)]^n\,dv.$$

This integral is a Laplace integral. The function $\Phi(v)$ is less than 1 and $\Phi(v)\to 1$ as $v\to-\infty$. We shall make a change of variable, auch that we imbed in 0 the maximum point of the function raised to the n-th power. Assuming $u=1-\Phi(v)$, we have

$$I(r,n)=r^{1/2}\int_0^{1/2} e^{-(r-1)v^2/2}(1-u)^n\,du, \tag{4.6}$$

where $v=v(u)$ is a function determined from the equation $u=1-\Phi(v)$. Note that $v(u)\to-\infty$ as $u\to 0$. The asymptotic expansion for the function of the errors as $v\to-\infty$ gives (see /12, p.119/)

$$u=1-\Phi(v)=(2\pi)^{-1/2}\frac{e^{-v^2/2}}{-v}[1+O(v^{-2})].$$

Putting this equation into algorithmic form, we obtain

$$\ln u = -\frac{v^2}{2} - \ln(-v) - \ln[(2\pi)^{1/2}] + \ln[1 + O(v^{-2})],$$

whence follow the expressions for v^2 and v and we shall substitute them into (4.6). We have

$$I(r, n) = r^{1/2}(4\pi)^{(r-1)/2}\int_0^{1/2} u^{r-1}(-\ln u)^{(r-1)/2}(1-u)^n[1 + O(v^{-1})]\,du.$$

Making the change of variable $t=-\ln(1-u)$ and noting that $u=1-e^{-t}=t+O(t^2), t\to 0,$ we obtain

$$I(r, n) = r^{1/2}(4\pi)^{(r-1)/2}\int_0^{-\ln 1/2} t^{r-1}(-\ln t)^{(r-1)/2}e^{-t(n+1)} \times [1 + O(t^2)][1 + O(v^{-1})]\,dt.$$

To clarify the asymptotic behaviour of the last integral we will use Watson's lemma for the logarithmic singularity.

Lemma 6 (see /13, p.46/). Suppose $\gamma, \lambda\in\mathbf{R}$, $\beta>0$, the function $f(x)$ is continuously differentiable for small $x\geqslant 0$ and is continuous for $0\leqslant x\leqslant a<\infty$. Then as $\lambda\to\infty$ the following asymptotic equation holds:

$$\int_0^a x^{\beta-1}|\ln x|^\gamma e^{-\lambda x}f(x)\,dx = \lambda^{-\beta}(\ln\lambda)^\gamma\Gamma(\beta)f(0)[1 + O((\ln\lambda)^{-1})].$$

Using this lemma, we will obtain a statement of Lemma 5. Using Lemmas 3, 4, 5 and Eqs.(4.2), (4.5), we will obtain a statement of Theorem 2. From Theorem 2 we obtain directly

Theorem 3. The average number of $s(n, k)$ steps in Lemke's method for problem (2.5a) using the orthoinvariant measure $\mu_{n+2,k+1}$ in the manifold $G(n+2, k+1)$ for fixed k and as $n\to\infty$ has the asymptotic estimate

$$s(n, k)\leqslant {}^1/_2(k+1)^{1/2}(\pi\ln n)^{k/2} + O((\ln n)^{(k-1)/2}). \tag{4.7}$$

From Theorem 2 and Proposition 1 we can obtain

Theorem 4. Suppose in the space of the $(k+1)\times(n+2)$-matrices of rank $k+1$ the probability measure μ is specified, which is invariant with respect to right-multiplication to any orthogonal $n+2$-order matrix. Then the average number of $s(n, k)$ steps in Lemke's method for problem (2.3a) using the measure μ in the above matrix space for fixed k and as $n\to\infty$ has the asymptotic estimate (4.7)

Remark 2. The problem solved in this paragraph has several interesting reformulations, of which we will present two:

1) the probability reformulation: if $\xi_1, \ldots, \xi_n$ are independent random $(0, 1)$-Gaussian quantities, what kind of sequence β_n need there be, such that

$$\mathbf{P}\left\{n^{-1/2}\sum_{i=1}^n \xi_i \geqslant \frac{1}{\beta_n}\max_{1\leqslant i\leqslant n}|\xi_i|\right\} \to o(1);$$

answer: $\beta_n\sim c(\ln n)^{1/2}$.

2) from spherical geometry: for what kind of sequence γ_n does the spherical size of the cone, drawn to the vectors $x_i=e_i-\gamma_n 1$, $1=(1, \ldots, 1)$ - e_i are basis vectors in R^n, $i=1, 2, \ldots, n$, - have a non-vanishing limit as $n\to\infty$; answer: $\gamma_n\sim c(n^2\ln n)^{-1/2}$.

Both answers follows from the theorems obtained above.

5. A dual geometrical interpretation.

The most fruitful geometrical interpretation of the linear programming problem was introduced by L.V. Kantorovich and considered in detail by G. Sh. Rubinshtein in /14/; it is implemented in the space of the "right-hand sides" and reduces the linear programming problem to the problem of the least point of intersection of a half-line and convex cone. Its advantage lies in the possibility of simultaneously interpreting direct and dual problems. In the usual form and in the notation of problem (2.1a) and (2.2) it looks as follows:

$$\max\left\{\lambda \,\middle|\, \sum_{i=1}^n x_i d^i = d^{n+1} - \lambda d^0,\ x_i \geqslant 0,\ i = 1, 2, \ldots, n\right\}. \tag{5.1}$$

where d^i is the form of the basis vector e^i for the mapping $D: R^{n+2}\to R^{k+1}$. If we introduce the cone $C=\mathrm{Con}\{d^i, i=1, 2, \ldots, n\}\subset R^{k+1}$ and the straight line $\Pi=\{-\lambda d^0+d^{n+1}|\lambda\in\mathbf{R}\}\subset R^{k+1}$, then problem (5.1) consists of obtaining the least point (with a minimum zero coordinate) of the intersection $\Pi\cap C$.

We shall write in the spirit of the Grassman interpretation. For this we shall consider the universal cone $Q=\mathrm{Con}\{e^i, i=1, 2, \ldots, n\}\subset R^{n+2}$ and the straight line $L=\{-\lambda e^0+e^{n+1}|\lambda\in\mathbf{R}\}\subset R^{n+2}$. Then the trivial universal problem consists of obtaining the least point (with a minimum zero coordinate) of the intersection $L\cap Q$ $(=\varnothing)$. Consider $D: R^{n+2}\to R^{k+1}\approx R^{n+2}/\mathrm{Ker}\,D$. Suppose, as in Sect. 2, $E=\mathrm{Im}\,D^*$, $E^\perp=\mathrm{Ker}\,D$. Identifying $R^{n+2}/E^\perp$ with E using D^*, we can

assume that the cone $C \approx D^*DQ$ and the straight line $\Pi \approx D^*DL$ lie in E; then we obtain the problem

$$\max \{\lambda | D^*DQ \ni D^*De^{n+1} - \lambda D^*De^0\}. \tag{5.1a}$$

Since the conditions of problem (5.1a) depend only on the kernel of the operator D^*D, the following equivalent problem arises:

$$\max \{\lambda | PQ \ni Pe^{n+1} - \lambda Pe^0\}, \tag{5.2a}$$

where P is an orthoprojector to E. It is now natural to supplement this same problem for the subspace $E^\perp$ with the replacement (as in Sect. 2) of e^0, e^{n+1} by $-e^{n+1}, -e^0$:

$$\min \{\mu | P_\perp Q \ni -P_\perp e^0 - \mu P_\perp e^{n+1}\}, \tag{5.2b}$$

where $P_\perp$ is an orthoprojector on $E^\perp$.

When transfering from (5.1) to (5.2a) we again identify the problems in which the matrices have identical kernels, and obtain a new method of transforming Grassman's manifold into the space of the linear programming problems, for which (5.2a) and (5.2b) are completely identical within the replacement of E by $E^\perp$ and e^0, e^{n+1} by $-e^{n+1}, -e^0$. Unlike (2.5), problems (5.2) are projections (and not sections) in E or $E^\perp$ of the universal trivial problem.

We shall present, without details, a new interpretation of the parametric dual permissibility of the set $J \in \mathcal{B}$ (see Sect. 3). At the same time one more universal set is needed, due to the fact that the set K, which is initial, is fixed.

Suppose $T = \operatorname{Con}\{e^i, i=0, 1, \ldots, k\} + \operatorname{co}\{0, e^i, i \in M\}$ (the sum of the cone and the simplex), $S(J) = \operatorname{Con}\{e^i, i \in J\}$, $T(J) = \operatorname{Con}\{e^i, i \in K \cap J\} + \operatorname{co}\{e^i, i \in M \cap J\}$,

$$\Pi(t) = \left\{-\lambda e^0 + e^{n+1} + t\sum_{i=1}^{k} e^i \,|\, \lambda \in \mathbf{R}\right\},$$

$H(t) = \operatorname{Aff}\{T(J), -te^0\}$ ($\operatorname{Con}\{\cdot\}$ is a conical, $\operatorname{co}\{\cdot\}$ is a convex, and $\operatorname{Aff}\{\cdot\}$ is an affine envelope of the sets of vectors in brackets).

Proposition 5. The base set $J \in \mathcal{B}$ is parametrically optimal in problem (2.3a) and others when and only when $t \geqslant 0$ exists, such that the following conditions hold:

a) $PS(J) \cap P\Pi(t) \neq \varnothing$ (parametric direct permissibility);

b) the hyperplane $PH(t)$ in E is supporting to PT (parametric dual permissiblity).

Corollary 5. The basis set $J \in \mathcal{B}$ is a parametrically doubly permissible set when and only when $PT(J)$ is a $(k-1)$-dimensional bound of the polyhedral set PT.

Corollary 6. The number of steps in Lemke's method does not exceed the number of bounds of codimension 2 of the polyhedral set PT for almost all $E \in G(n+2, k+1)$.

Thus, a new interpretation of the functional $s_1(E)$ from Sects. 3,4 is the number of bounds of codimension 2 of some polyhedral set in E. Changing E, we obtain radom polyhedral sets and arrive at the following problem which is well-known in random convex geometry /8/:

to obtain the average number of bounds of a random polyhedral set.

Thereby our fundamental result (Theorem 2) solves this problem for some statistics, and also the methods of constructing the set presented above (the projection of T into the random subspace E). This result agrees with the facts obtained in /9/.

6. Comments.

The Grassman approach to solving the problems of linear programming and convex geometry was suggested by A.M. Vershik in /8, 15, 16/. It should be expected that in the convex geometry problem we can use it to solve the old problem of the standard properties of convex polyhedra of large dimensions. Note the similarity of the Grassman approach to the ideas of Gayley's diagrams /17/. There are four models: the section or projection of a standard object with a subspace and its orthogonal addition. These are problems (2.5) and (5.2). The idea of including the parametric method inthe problem of estimating the number of steps of the simplex method was independently suggested in /6/, and also in /18, 19 and 5/. As a result it turned out that the methods and results of our paper and of /5/ differ: in /5/ a rougher method of estimation was used, which was, however, suitable for a wider class of statistics. In addition, Lemke's parametric method was considered in /5/, while a simpler parametric method was considered in /1/. In this paper, unlike /1/, an accurate constant is given in the estimate, and the order of entry of $\ln n$ is one and the same $(k/2)$. As shown in /6/, using the method in /5/ it is impossible to obtain our estimate, and we can very slightly improve the estimate from /5/ using the method of the author ($(\ln n)^{k/2}$ instead of $(\ln n)^{[(k+1)/2]}$). The Grassman approach in /5/ is not used (see also the review in /20/).

It should be noted that /11/ and other papers examined the problem of the average number of bounds of the leading dimension in a random polyhedron regardless of linear programming. The results of /11/ also reduce to near estimates.

After /5, 18, 19/ several interesting papers were published. One of them /7/ contains an estimate of the number of steps which is quadratic with respect to $\min(m, k)$. At the same time a more powerful parametric method is used - polynomial perturbation, similar to versal deformation in singularity theory and to Charns' method of controlling degeneration in linear programming problems.

In conclusion we mention that in /21 and 22/ the authors approach the linear programming problem from a quite different angle: they generally give an infinite polynomial algorithm for solving the linear programming problem. But the problem of the relation between these methods and simplex methods has not yet been solved in anyone's favour. To clarify this a purposeful calculational experiment is needed.

REFERENCES

1. VERSHIK A.M. and SPORYSHEV P.V., Estimates of the average number of steps of the simplex method and of the problems of asymptotic integral geometry. Dokl. AN SSR, 271,5,1044-1048,1983.
2. KANTOROVICH L.V., An economic calculation of the best use of resources. Moscow, Nauka, 1959.
3. DANTZIG J.B., Linear programming, its applications and generalizations. Moscow, Progress, 1966.
4. KLEE V. and MINTY G., How good is the simplex algorithm?-In: Inequalities III. N.Y.: Acad. Press, 159-175,1972.
5. SMALE S., On the average number of steps of the simplex method of linear programming-Math. Program. 27,3,241-262,1983.
6. SPORYSHEV P.V., The use of integral geometry in linear programming: Candidate Dissertation. Leningrad, LGU, 1984.
7. ADLER I., Megqido N. and Todd M.J., New results on the average behaviour of symplex algorithms. Bull. Amer. Math. Soc., 11, 2, 378-382,1984
8. VERSHIK A.M. and CHERNYAKOV A.G., Critical points of fields of convex polyhedra and the Paréto-Smale optimum with respect to a convex cone. Dokl. AN SSSR, 263,3,529-532,1982.
9. SANTALO L., Integral geometry and geometric probabilities. Moscow: Nauka, 1983.
10. GOL'SHTEIN E.G. and YUDIN D.B., New trends in linear programming. Moscow, Sov. radio, 1966.
11. REYNAUD H., Sur l'envelope des nuages des points aleatoires dans R_n. I.-J. Appl. Probability, 7,35-48,1970.
12. DVAIG G.B., Tables of integral and other mathematical formulas. Moscow, Nauka, 1977.
13. FEDORYUK M.V., The method of steepest descents. Moscow, Nauka, 1977.
14. RUBINSHTEIN G.SH., The problem of the least point of intersection of an axis with a polyhedron and some of its applications. Uspekhi matem. nauk, 10,4,206-207,1955.
15. VERSHIK A.M., Some comments on the infinite-dimensional problem of linear programming. Uspekhi matem. Nauk, 25, 5, 117-124, 1970.
16. VERSHIK A.M., Description of invariant measures for the effect of some infinite-dimensional groups. Dokl. AN SSSR, 213,4,749-752,1974.
17. GRUNBAUM B., Convex polytopes. N.Y. Intersci., 1967.
18. BORGWARDT K.H., Some distribution-independent results about the asymptotic order of the average number of pivot steps of the simplex method. Math. Operat. Res., 7,441-462, 1982.
19. BORGWARDT K.H., The avarage number of steps required by the simplex method is polynomial. J. Operat. Res. A-B 26,157-177,1982.
20. SMALE S., On the efficiency of algorithms of analysis. Bull. Amer. Math. Soc., 13,2,87-121 1985.
21. KHACHIYAN L.G., Polynomial algorithms in linear programming. Zh. vychisl. matem. matem. fiz., 20, 1, 51-68,1980.
22. KARMARKAR N., A new polynomial-time algorithm for linear programming. Combinatorica, 4,4,373-395,1984.

Translated by H.Z.

U.S.S.R. Comput.Maths.Math.Phys.,Vol.26,No.3,pp. 113-120,1986 0041-5553/86 $10.00+0.00
Printed in Great Britain.

ESTIMATES OF THE RATE OF CONVERGENCE OF DIFFERENCE SCHEMES FOR VARIATIONAL ELLIPTIC SECOND-ORDER INEQUALITIES IN AN ARBITRARY DOMAIN*

S.A. VOITSEKHOVSKII, I.N. GAVRILYUK and V.S. SAZHENYUK

Second-order variational elliptic inequalities in an arbitrary domain with a constraint inside the domain or at the boundary with respect to the penalty and fictitious domains method are approximated using non-linear boundary value problems in a rectangle. Difference schemes are constructed for these problems and the rates of convergence of difference solutions to solutions of the variational inequalities are established.

*Zh.vychisl.Mat.mat.Fiz.,26,6,827-836,1986.