

SOME PROBLEMS IN ASYMPTOTIC CONVEX GEOMETRY AND RANDOM MATRICES MOTIVATED BY NUMERICAL ALGORITHMS

ROMAN VERSHYNIN

ABSTRACT. The simplex method in Linear Programming motivates several problems of asymptotic convex geometry. We discuss some conjectures and known results in two related directions – computing the size of projections of high dimensional polytopes and estimating the norms of random matrices and their inverses.

1. ASYPTOTIC CONVEX GEOMETRY AND LINEAR PROGRAMMING

Linear Programming studies the problem of maximizing a linear functional subject to linear constraints. Given an *objective vector* $z \in \mathbb{R}^d$ and *constraint vectors* $a_1, \dots, a_n \in \mathbb{R}^d$, we consider the linear program

$$\begin{aligned} \text{(LP)} \quad & \text{maximize } \langle z, x \rangle \\ & \text{subject to } \langle a_i, x \rangle \leq 1, \quad i = 1, \dots, n. \end{aligned}$$

This linear program has d unknowns, represented by x , and n constraints. Every linear program can be reduced to this form by a simple interpolation argument [36]. The *feasible set* of the linear program is the polytope

$$P := \{x \in \mathbb{R}^d : \langle a_i, x \rangle \leq 1, \quad i = 1, \dots, n\}.$$

The solution of (LP) is then a vertex of P . We can thus look at (LP) from a geometric viewpoint:

for a polytope P in \mathbb{R}^d given by n faces, and for a vector z ,
find the vertex that maximizes the linear functional $\langle z, x \rangle$.

The oldest and still the most popular method to solve this problem is the *simplex method*. It starts at some vertex of P and generates a walk on the edges of P toward the solution vertex. At each step, a *pivot rule* determines a choice of the next vertex; so there are many variants of the simplex method with different pivot rules. (We are not concerned here with how to find the initial vertex, which is a nontrivial problem in itself).

Date: March 16, 2007.

Supported by the Alfred P. Sloan Foundation and by NSF DMS grant 0401032.

1.1. The shadow-vertex pivot rule and sections of polytopes. The most widely known pivot rule maximizes the objective function $\langle z, x \rangle$ over the neighboring vertices. The resulting walk on the vertices is defined iteratively, and thus is usually hard to analyze. An alternative *shadow-vertex* pivot rule [10] defines a walk on the polytope P as a preimage of a projection of P . The resulting walk can be described in a non-iterative way, so one hopes to analyze it with the methods of asymptotic convex geometry.

Suppose we know a solution x_0 of (LP) for some other objective vector z_0 . The shadow-vertex simplex method interpolates between z_0 and z by computing the solutions of (LP) for all z' in the plane

$$E = \text{span}(z_0, z).$$

From a geometric viewpoint, we consider the orthogonal projection $Q(P)$ of the feasible polytope P onto E . It is easily checked that the vertices x_0 and x of P will be preserved by the projection: $Q(x)$ and $Q(x_0)$ will be vertices of the polygon $Q(P)$.

The shadow-vertex simplex method thus computes the vertices of the polygon $Q(P)$ one by one, starting from $Q(x_0)$ and ending with $Q(x)$. So at the end it outputs x , which is the solution of (LP). One can express the computation of $Q(x)$ as a pivot rule, and check that each next vertex can be computed in polynomial time. The resulting walk on the polytope P is therefore the preimage of the vertices of the polygon $Q(P)$ under the projection Q .

It will be convenient to work in the dual setting. The polar of P is

$$K := P^\circ = \{x \in \mathbb{R}^d : \langle x, y \rangle \leq 1 \text{ for all } y \in P\} = \text{conv}(0, a_1, \dots, a_n)$$

and the polar of the projection $Q(P)$ is the section $K \cap E$. The length of the walk in the shadow-vertex simplex method is thus bounded by the size (the number of edges) of the polygon $K \cap E$.

1.2. Complexity of the simplex method and the size of sections. The running time of the simplex method is proportional to the length of the walk on the edges of P it generates. Hirsch's conjecture states that every polytope P in \mathbb{R}^d with n faces has *diameter* at most $n - d$. The diameter is the maximum of the shortest walk on the edges between any pair of the vertices. The best known bound on the diameter is $n^{\log_2 d + 2}$ due to Kalai and Kleitman [15].

For every known variant of the simplex method, an example of (LP) is known for which the length of the walk on P is not polynomial in n and

d .¹ For the the classical (maximizing) pivot rule, such an example was first constructed by Klee and Minty [18]: on a certain deformed cube, the walk visits each of the 2^d vertices [18].

Similar pessimistic examples are known for the the shadow-vertex simplex method: the size of the planar section $K \cap E$ that bounds the length of the walk is in general exponential in n, d . This follows for example from the seminal construction in semidefinite programming by Ben-Tal and Nemirovski [5], which yields a polytope K and a plane E such that the section $K \cap E$ is an approximation of the circle with error *exponentially small* in n, d .

Problem 1.1 (Sections of polytopes). *Let K be a polytope in \mathbb{R}^d with n vertices, and E be a two-dimensional subspace of \mathbb{R}^d . Estimate the size (the number of edges) of the polygon $K \cap E$. Under what conditions on K and E is this number polynomial in n, d ?*

This problem is somewhat opposite to the typical problems of the asymptotic convex geometry, whose ideal would be to produce the most round section (fine approximation to a circle). In Problem 1.1 our ideal is a section with fewest edges, thus farthest from the circle. From the viewpoint of the simplex method, “round” polytopes have high complexity, while polytopes with fewest faces have low complexity.

1.3. Smoothed analysis and randomly perturbed polytopes. Despite the known examples of exponentially long walks, on most problems that occur in practice the simplex algorithm runs in polynomial and even linear time. To explain this empirical evidence, the *average analysis* of the simplex method was developed in the eighties, where the (LP) was drawn at random from some natural distribution and the expected size of the walk was shown to be polynomial in n, d [4, 28, 30, 21, 12, 1, 33, 3, 2].

In particular, Haimovich showed ([12], see [26], Section 11.5) that if one chooses the directions of the inequalities in (LP) uniformly at random as \leq or \geq , then the expected length of the walk in the shadow-vertex simplex method is at most $d/2$. Note that the size does not depend on the number of inequalities n .

However, reversing inequalities is hard to justify in practice. Spielman and Teng [32] proposed to replace average analysis by a finer model, which they called *smoothed analysis*, and where the random inputs are replaced by slight *random perturbations* of arbitrary inputs. Smoothed analysis thus

¹Recently, a randomized polynomial time pivoting algorithm for (LP) was found by Kelner and Spielman [16]. However, their algorithm generates a walk on some other polytope related to (LP) and not on P .

interpolates between the worst case analysis (arbitrary inputs) and the average analysis of Smale (random inputs).

Spielman and Teng [31] first showed that the shadow-vertex simplex method has polynomial smoothed complexity. If the polytope K is randomly perturbed, then its section $K \cap E$ will have an expected polynomial size (which in turn bounds the length of the walk in the simplex method). Their result was improved in [7] and in [36], and the current best bound is as follows:

Theorem 1.2. [36] *Let a_1, \dots, a_n be independent Gaussian vectors in \mathbb{R}^d with centers of norm at most 1, and whose components have standard deviation $\sigma \leq 1/6\sqrt{d \log n}$. Let E be a plane in \mathbb{R}^d . Then the random polytope $K = \text{conv}(a_1, \dots, a_n)$ satisfies*

$$(1.1) \quad \mathbb{E} |\text{edges}(K \cap E)| \leq Cd^3\sigma^{-4},$$

where C is an absolute constant.

The prior weaker bound of Spielman and Teng [31] was $Cnd^3\sigma^{-6}$; the subsequent work of Deshpande and Spielman [7] improved upon the exponent of d but doubled the exponent of n .

Theorem 1.1 shows that the expected size of the section $K \cap E$ is *polylogarithmic* in n , while the previous bound were polynomial in n . Going back to the pre-dual polytope P , this indicates that random perturbations of the polytopes create short walks between any two *given* vertices. Note that for large n , this polylogarithmic bound becomes better than the bound $n - d$ in Hirsch's conjecture.

Theorem 1.1 provides a solution to Problem 1.1 for a randomly perturbed polytope and a fixed subspace. A seemingly harder problem, which is still open, is for an *arbitrary polytope and a randomly perturbed subspace*. This version would be significant for the analysis of the simplex method, because it allows one to leave the constraints intact and only perturb the objective function.

Another open problem is to estimate the diameter of randomly perturbed polytopes, rather than all polytopes as in Hirsch's conjecture.

Problem 1.3 (Spielman-Teng [31]). *Let K be a perturbed polytope as in Theorem 1.2. Estimate the expected diameter of $P = K^\circ$. Is it always polynomial in n , d and σ ? Perhaps even polylogarithmic in n ?*

Finally, no analog of Theorem 1.2 is known for *bounded perturbations*, i.e. for $a_i = \bar{a}_i + \sigma\theta_i$, where \bar{a}_i are arbitrary fixed vectors of norms at most 1 and θ are independent vectors chosen from $\{-1, 1\}^d$ or from $[-1, 1]$ uniformly at random. Such *bounded smoothed analysis* is a common model

for roundoff errors when real numbers are represented as binary numbers in computers [11].

1.4. Nondegeneracy of faces and invertibility of random matrices.

The approach to Theorem 1.2 developed by Spielman and Teng [31] is based on the intuition that most faces of K should be non-degenerate simplices (e.g. they have inscribed balls of polynomial radii). If the plane E intersects such a nondegenerate simplex F , the length of the interval $E \cap F$ is likely to be polynomially big (if the plane intersects a simplex, it is likely to pass through its “bulk” rather than touch the boundary only).

On the other hand, with high probability all vectors in the the perturbed polytope K have norms $O(\log n)$. (Its vertices are Gaussian perturbations of n vectors of norm at most 1). Therefore, the perimeter of the polygon $K \cap E$ can be at most $O(\log n)$. Since all edges $E \cap F$ of this polygon are polynomially big, we conclude that there are at most polynomially many edges, as desired.

There are several places where this approach breaks down or is not known to succeed. One such problem is the non-degeneracy of the faces. The nondegeneracy of a simplex S is usually quantified with the smallest singular value of the matrix A that realizes the change of the basis from the standard simplex to S . For the polytope $K = \text{conv}(a_1, \dots, a_n)$, each face is a simplex with vertices $(a_i)_{i \in I}$ for some d -element subset $I \subset \{1, \dots, n\}$. If the vertices a_i are Gaussian as in Theorem 1.2, the change of basis is a $d \times d$ matrix with Gaussian independent entries. Thus we need the random Gaussian matrices to be far from being singular. Quantitative theory of invertibility of random matrices is the subject of the next section.

2. INVERTIBILITY OF RANDOM MATRICES

For a one-to-one linear operator $A : X \rightarrow Y$ between two normed spaces X and Y , two quantities are central in functional analysis: the norm $\|A\|$ and the norm of the inverse $\|A^{-1}\|$. If the operator is not onto, then the inverse norm is computed for the restriction of A onto its image; so we identify $\|A^{-1}\|$ with $\|(A|_{A(X)})^{-1}\|$. Thus

$$\|A\| = \sup_{x \in X: \|x\|=1} \|Ax\|, \quad \frac{1}{\|A^{-1}\|} = \inf_{x \in X: \|x\|=1} \|Ax\|.$$

The operator A can be viewed as realizing an *embedding* of the space X into the space Y , and the product $\|A\|\|A^{-1}\|$ is the *distortion* of the embedding (see [13]).

The canonical example is when both X and Y are finite dimensional Euclidean spaces, say $X = \mathbb{R}^k$, $Y = \mathbb{R}^n$, where we identify the linear operator

A with its $k \times n$ matrix. The *singular values* of A are the eigenvalues of $|A| = \sqrt{A^*A}$, the largest and the smallest singular values being

$$\lambda_{\max}(A) = \|A\|, \quad \lambda_{\min}(A) = \frac{1}{\|A^{-1}\|}.$$

In the numerical linear algebra and scientific computing literature, the distortion

$$\kappa(A) = \lambda_{\max}(A)/\lambda_{\min}(A) = \|A\|\|A^{-1}\|$$

is commonly called the *condition number* of A .

We are interested estimating these quantities for random matrices A . For one reason, random matrices sometimes provide an intuition for what to expect in practice; we saw such reasoning about average analysis and smoothed analysis of the simplex method in the previous section. Random linear operators with controllable distortion (or their adjoints) also serve as handy tools in most randomized constructions in geometric functional analysis [7], geometric algorithms in theoretical computer science [34, 35], compressed sensing in information processing ([8], [6]), vector quantization [19] and some other fields.

2.1. Gaussian matrices. We start from the simplest case of a *Gaussian matrix*, those whose entries are i.i.d. standard normal random variables. The asymptotics of the largest and the smallest singular values is well understood in this case: for a $n \times d$ Gaussian matrix A with $n \geq d$, one has

$$\lambda_{\max}(A) \approx \sqrt{n} + \sqrt{d}, \quad \lambda_{\min}(A) \approx \sqrt{n} - \sqrt{d} \quad \text{with high probability.}$$

There is a long history of such asymptotic results. In particular, the largest and the smallest singular values converge almost surely to their corresponding values above as the dimension n grows to infinity and the aspect ratio d/n converges to a constant, see [7]. A sharp nonasymptotic result – for every fixed n and d – follows from Gordon’s inequality (see [7]):

$$\sqrt{n} - \sqrt{d} \leq \mathbb{E}\lambda_{\min}(A) \leq \mathbb{E}\lambda_{\max}(A) \leq \sqrt{n} + \sqrt{d}.$$

Combining with the concentration of measure inequality, one deduces a deviation bound [7]: for every $t > 0$, with probability at least $1 - 2e^{-t^2/2}$ one has

$$(2.1) \quad \sqrt{n} - \sqrt{d} - t \leq \lambda_{\min}(A) \leq \lambda_{\max}(A) \leq \sqrt{n} + \sqrt{d} + t.$$

Note that the lower bounds become meaningless for *square Gaussian matrices*, those with $n = d$. Yet this case is central in some applications: as we saw in Section 1.4, the square matrices determine the faces of the polytope K in linear programming, and nondegeneracy of such a face translates into

a lower bound for $\lambda_{\min}(A)$. To guess the order of the smallest singular value, note that for an $(n-1) \times n$ matrix, the lower bound is $\sqrt{n} - \sqrt{n-1} \sim n^{-1/2}$.

Von Neumann and his associates, who used random matrices as test matrices for their algorithms, indeed speculated that for random square matrices one should have

$$(2.2) \quad \lambda_{\min}(A) \sim n^{-1/2} \quad \text{with high probability}$$

(see [37], pp. 14, 477, 555.) In a more precise form, this estimate was conjectured by Smale [29] and proved by Edelman [9] for Gaussian matrices: for every $\varepsilon \geq 0$, one has

$$(2.3) \quad \mathbb{P}(\lambda_{\min}(A) \leq \varepsilon n^{-1/2}) \sim \varepsilon.$$

In particular, the smallest singular value is not concentrated: the mean and the standard deviation of $n^{1/2}\lambda_{\min}(A)$ are both of the order of a constant. This is very different from the behavior of the largest singular value, which by (2.1) is tightly concentrated around its mean.

An elegant argument by Sankar, Spielman and Teng [25] generalizes (2.3) for random Gaussian perturbations of an arbitrary matrix, i.e. for the smoothed analysis setting of Theorem 1.2:

Theorem 2.1 (Sankar, Spielman and Teng [25]). *Let A be an $n \times n$ matrix with independent Gaussian random entries (not necessarily centered), each of variance σ^2 . Then, for every $\varepsilon \geq 0$, one has*

$$\mathbb{P}(\lambda_{\min}(A) \leq \varepsilon n^{-1/2}) \leq C\varepsilon/\sigma,$$

where $C = 1.823$.

In applications for random polytopes such as in Section 1.4, we need *all faces* to be nondegenerate, thus all $d \times d$ submatrices of a random $n \times d$ Gaussian matrix (whose rows are the constraint vectors a_1, \dots, a_n from (LP)) be nicely invertible. This motivates the following problem:

Problem 2.2. *Let A be an $n \times d$ Gaussian matrix (with i.i.d. standard normal entries, or, more generally, as in Theorem 2.1). Estimate the expected minimum of the smallest singular values of all $d \times d$ submatrices of A .*

In particular, if $n = O(d)$, we want this minimum to be polynomially small rather than exponentially small in d .

2.2. General matrices with i.i.d. entries. Most problems we discussed become much harder once Gaussian matrices are replaced with other natural matrices with i.i.d. entries. Nevertheless, understanding of discrete matrices, whose entries can take finite set of values, is important in applications such as in numerical algorithms, which can only deal with discrete

values. A survey on random discrete matrices was recently written by Vu [38].

Asymptotic theory of random matrices has developed to a point where the behavior of the largest singular value is well understood. Suppose A is an $n \times d$ matrix with i.i.d. centered entries, which have variance 1. Then the finiteness of the *fourth moment* of the entries is necessary and sufficient that

$$(2.4) \quad \lambda_{\max}(A) \rightarrow \sqrt{n} + \sqrt{d} \quad \text{almost surely}$$

as the dimension n grows to infinity and the aspect ratio d/n converges to a constant [39]. A similar statement holds for the convergence in probability, and with a slightly weaker condition than the fourth moment [27].

Under a much stronger *subgaussian* moment assumption, which still holds discrete and gaussian random variables, a parallel non-asymptotic result is known (for all finite n and d). A random variable ξ is called *subgaussian* if its tail is dominated by that of the standard normal random variable: there exists a constant $B > 0$ such that

$$(2.5) \quad \mathbb{P}(|\xi| > t) \leq 2 \exp(-t^2/B^2) \quad \text{for all } t > 0.$$

The minimal B here is called the *subgaussian moment*.² Gaussian random variables and all bounded random variables, in particular the symmetric ± 1 random variable, are examples of subgaussian random variables. Inequality (2.5) is often equivalently stated as a moment condition

$$(2.6) \quad (\mathbb{E}|\xi|^p)^{1/p} \leq CB\sqrt{p} \quad \text{for all } p \geq 1,$$

where C is an absolute constant.

The following non-asymptotic result follows from a more general result proved by Klartag and Mendelson ([17], Theorem 1.4) with constant probability, which was later improved by Mendelson, Pajor and Tomczak-Jaegermann ([22], Theorem D) to an exponential probability:

Theorem 2.3 ([17, 22]). *Let A be an $n \times d$ matrix ($n \geq d$) with i.i.d. centered subgaussian entries with variance 1. Then, with probability at least $1 - Ce^{-m}$, one has*

$$\sqrt{n} - C\sqrt{d} \leq \lambda_{\min}(A) \leq \lambda_{\max}(A) \leq \sqrt{n} + C\sqrt{d},$$

where C depends only on the subgaussian moment of the entries.

This estimate approaches the sharp asymptotic bound (2.4) for very tall matrices (for small aspect ratios d/n). However, the lower bound becomes useless for the aspect ratios above some constant, and in particular says nothing about square matrices.

²In the literature in geometric functional analysis, the subgaussian moment is often called the ψ_2 -norm.

Problem 2.4. *Let A be an $n \times d$ matrix ($n \geq d$) with i.i.d. centered subgaussian entries with variance 1. Is it true that with high probability one has*

$$\lambda_{\min}(A) \geq c(\sqrt{n} - \sqrt{d}),$$

where $c > 0$ depends only on the subgaussian moment of the entries?

In a positive direction, lower bounds valid for all aspect ratios $y := d/n < 1$ were proved by Litvak, Rudelson, Pajor and Tomczak-Jaegermann [20] with an exponential dependence on $1 - y$, and improved to a linear dependence by Rudelson [23].

Nevertheless, even the positive solution of Problem 2.4 would not say anything for *square matrices*, those with $n = d$. This problem was recently solved in the work [24], which confirmed prediction (2.2) for general matrices with independent entries. Recall that the bounded fourth moment of the entries is necessary and sufficient to control the largest singular value as in (2.4). Then [24] proves that the fourth moment assumption (i.e. the fourth moments of the entries are uniformly bounded) is also sufficient to control the smallest singular value. For an $n \times n$ matrix A with random centered entries of variances at least 1,

Under the fourth moment assumption, prediction (2.2) holds.

The identical distribution of the entries is not needed in this result.

For a stronger *subgaussian* assumption on the entries, prediction (2.2) holds with exponentially high probability. This was conjectured by Spielman and Teng [31] for random ± 1 matrices:

$$\mathbb{P}(s_n(A) \leq \varepsilon n^{-1/2}) \leq \varepsilon + c^n,$$

and proved in [24] in more generality – for all matrices with subgaussian i.i.d. entries, and up to a constant factor which depends only on the subgaussian moment.

Theorem 2.5 ([24]). *Let A be an $n \times n$ matrix with i.i.d. centered subgaussian entries with variance 1. Then for every $\varepsilon \geq 0$ one has*

$$(2.7) \quad \mathbb{P}(\lambda_{\min}(A) \leq \varepsilon n^{-1/2}) \leq C\varepsilon + c^n,$$

where $C > 0$ and $c < 1$ are constants that depend (polynomially) only on the subgaussian moment of the entries.

In particular, for $\varepsilon = 0$ we deduce that any random square matrix with i.i.d. subgaussian entries with variance 1 is *singular with exponentially small probability*. For random matrices with ± 1 entries, this was proved by Kahn, Komlós and Szemerédi [14]. For more on prior work and related conjectures on the singularity probability, see [38, 24].

REFERENCES

- [1] I. Adler, *The expected number of pivots needed to solve parametric linear programs and the efficiency of the self-dual simplex method*. Technical Report, University of California at Berkeley, May 1983
- [2] I. Adler, R. M. Karp, R. Shamir, *A simplex variant solving an $m \times d$ linear program in $O(\min(m^2, n^2))$ expected number of pivot steps*, J. Complexity 3 (1987), 372–387
- [3] I. Adler, N. Megiddo, *A simplex algorithm whose average number of steps is bounded between two quadratic functions of the smaller dimension*, Journal of the ACM 32 (1985), 871–895
- [4] K.-H. Borgwardt, *The simplex method. A probabilistic analysis*. Algorithms and Combinatorics: Study and Research Texts, 1. Springer-Verlag, Berlin, 1987.
- [5] A. Ben-Tal, A. Nemirovski, *On polyhedral approximations of the second-order cone*, Math. Oper. Res. 26 (2001), 193–205
- [6] The Compressed Sensing website <http://www.dsp.ece.rice.edu/cs/>
- [7] A. Deshpande, D. A. Spielman, *Improved smoothed analysis of the shadow vertex simplex method*, FOCS 2005 (46th Annual Symposium on Foundations of Computer Science), 349–356
- [8] D. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory 52 (2006), 1289–1306
- [9] A. Edelman, *Eigenvalues and condition numbers of random matrices*, SIAM J. Matrix Anal. Appl. 9 (1988), 543–560
- [10] S. Gaas, T. Saaty, *The computational algorithm for the parametric objective function*, Naval Research Logistics Quarterly 2 (1955), 39–45
- [11] R.M.Gray, D.L.Neuhoff, *Quantization*, IEEE Trans. Inform. Theory 44 (1998), 2325–2383
- [12] M. Haimovich, *The simplex algorithm is very good!: On the expected number of pivot steps and related properties of random linear programs*. Technical report, Columbia University, April 1983
- [13] P. Indyk, J. Matoušek, *Low-distortion embeddings of finite metric spaces*, in: Handbook of discrete and computational geometry, 177–196. Second edition. Edited by Jacob E. Goodman and Joseph O’Rourke. Discrete Mathematics and its Applications (Boca Raton). Chapman & Hall/CRC, Boca Raton, FL, 2004.
- [14] J. Kahn, J. Komlós, E. Szemerédi, *On the probability that a random ± 1 -matrix is singular*, J. Amer. Math. Soc. 8 (1995), no. 1, 223–240
- [15] G. Kalai, D. J. Kleitman, *A quasi-polynomial bound for the diameter of graphs of polyhedra*, Bulletin Amer. Math. Soc. 26 (1992), 315–316
- [16] J. A. Kelner, D. A. Spielman, *A randomized polynomial-time simplex algorithm for Linear Programming*, submitted.
- [17] B. Klartag, S. Mendelson, *Empirical Processes and Random Projections*, Journal of Functional Analysis 225 (2005), 229–245
- [18] V. Klee, G. J. Minty, *How good is the simplex algorithm?* In Sisha, O., editor, *Inequalities – III*, pp. 159–175, Academic Press, 1972
- [19] Yu. Liubarskii, R. Vershynin, *Uncertainty principles and vector quantization*, submitted
- [20] A. Litvak, A. Pajor, M. Rudelson, N. Tomczak-Jaegermann, *Smallest singular value of random matrices and geometry of random polytopes*, Adv. Math. 195 (2005), 491–523

- [21] N. Megiddo, *Improved asymptotic analysis of the average number of steps performed by the self-dual simplex algorithm*, Math. Programming 35 (1986), 140–172.
- [22] S. Mendelson, A. Pajor, N. Tomczak-Jaegermann, *Reconstruction and subgaussian operators*, Geometric and Functional Analysis, to appear
- [23] M. Rudelson, *Lower estimates for the singular values of random matrices*, Compt. Rendus Math. de L’Academie des Sciences 342 (2006), 247–252.
- [24] M. Rudelson, R. Vershynin, *The Littlewood-Offord Problem and invertibility of random matrices*, submitted
- [25] S. Sankar, D. Spielman, S.-H. Teng, *Smoothed analysis of the condition numbers and growth factors of matrices*, SIAM J. Matrix Anal. Appl. 28 (2006), 446–476
- [26] A. Schrijver, *Theory of linear and integer programming*. Wiley-Interscience Series in Discrete Mathematics. A Wiley-Interscience Publication. John Wiley & Sons, Ltd., Chichester, 1986
- [27] J. W. Silverstein, *On the weak limit of the largest eigenvalue of a large-dimensional sample covariance matrix*, J. Multivariate Anal. 30 (1989), 307–311
- [28] S. Smale, *On the average number of steps of the simplex method of linear programming*, Math. Programming 27 (1983), 241–262.
- [29] S. Smale, *On the efficiency of algorithms of analysis*, Bull. Amer. Math. Soc. (N.S.) 13 (1985), 87–121
- [30] S. Smale, *The problem of the average speed of the simplex method*. Mathematical programming: the state of the art (Bonn, 1982), 530–539, Springer, Berlin, 1983.
- [31] D. A. Spielman, S.-H. Teng, *Smoothed analysis: why the simplex algorithm usually takes polynomial time*, Journal of the ACM 51 (2004), 385–463
- [32] D. A. Spielman, S.-H. Teng, *Smoothed analysis of algorithms*, Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002), 597–606, Higher Ed. Press, Beijing, 2002.
- [33] M. J. Todd, *Polynomial expected behavior of a pivoting algorithm for linear complementarity and linear programming problems*, Mathematical Programming 35 (1986), 173–192
- [34] S. Vempala, *The random projection method*, Handbook of randomized computing, Vol. I, II, 651–671, Comb. Optim., 9, Kluwer Acad. Publ., Dordrecht, 2001.
- [35] S. Vempala, *The random projection method*. With a foreword by Christos H. Papadimitriou. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 65. American Mathematical Society, Providence, RI, 2004.
- [36] R. Vershynin, *Beyond Hirsch Conjecture: walks on random polytopes and smoothed complexity of the simplex method*, FOCS 2006 (47th Annual Symposium on Foundations of Computer Science), 133–142
- [37] J. von Neumann, *Collected works. Vol. V: Design of computers, theory of automata and numerical analysis*. General editor: A. H. Taub. A Pergamon Press Book The Macmillan Co., New York 1963
- [38] V. Vu, *Random discrete matrices*, submitted
- [39] Y. Q. Yin, Z. D. Bai, P. R. Krishnaiah, *On the limit of the largest eigenvalue of the large-dimensional sample covariance matrix*, Probab. Theory Related Fields 78 (1988), 509–521

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, DAVIS, CA 95616,
USA

E-mail address: `vershynin@math.ucdavis.edu`