Electronic Transactions on Numerical Analysis.
Volume 44, pp. 177–188, 2015.

ISSN 1068–9613.

ETNA
Kent State University
http://etna.math.kent.edu

RANDOMIZED METHODS FOR RANK-DEFICIENT LINEAR SYSTEMS*

JOSEF SIFUENTES†, ZYDRUNAS GIMBUTAS‡, AND LESLIE GREENGARD§

Abstract. We present a simple, accurate method for solving consistent, rank-deficient linear systems, with or without additional rank-completing constraints. Such problems arise in a variety of applications such as the computation of the eigenvectors of a matrix corresponding to a known eigenvalue. The method is based on elementary linear algebra combined with the observation that if the matrix is rank-k deficient, then a random rank-k perturbation yields a nonsingular matrix with probability close to 1.

Key words. rank-deficient systems, null space, null vectors, eigenvectors, randomized algorithms, integral equations

AMS subject classifications. 15A03, 15A12, 15A18, 65F15, 65F99

1. Introduction.

A variety of problems in numerical linear algebra involve the solution of rank-deficient linear systems. The most straightforward example is that of finding the eigenspace of a matrix $A \in \mathbb{C}^{n\times n}$ corresponding to a known eigenvalue λ. One then wishes to solve

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

If A itself is rank-deficient, of course, then setting $\lambda = 0$ corresponds to seeking its null space.

A second category of problems involves the solution of an inhomogeneous linear system

$$Ax = b, \tag{1.1}$$

where A is rank-k deficient but b is in the range of A. A third category consists of problems like (1.1), but for which a set of k additional constraints are known of the form:

$$C^*x = f\,, \tag{1.2}$$

where the matrix

$$\begin{bmatrix} A \\ C^* \end{bmatrix}$$

is full-rank. Here, $C \in \mathbb{C}^{n\times k}$, C^* denotes its conjugate transpose, and $f \in \mathbb{C}^k$.

In this relatively brief note, we describe a very simple framework for solving such problems using *randomized* schemes. They are particularly useful when A is well-conditioned

*Received September 25, 2014. Accepted December 10, 2014. Published online on February 13, 2015. Recommended by L. Reichel. The work of the second author (Z. G.) was supported in part by the Office of the Assistant Secretary of Defense for Research and Engineering and AFOSR under NSSEFF Program Award FA9550-10-1-0180 and in part by the National Science Foundation under grant DMS-0934733. Contributions by staff of NIST, an agency of the U.S. Government, are not subject to copyright within the United States. The work of the third author (L. G.) was supported in part by the Office of the Assistant Secretary of Defense for Research and Engineering and AFOSR under NSSEFF Program Award FA9550-10-1-0180, by the National Science Foundation under grant DMS-0934733, and by the Applied Mathematical Sciences Program of the U.S. Department of Energy under Contract DEFGO288ER25053.

†Department of Mathematics, Texas A&M University, Mailstop 3368, College Station, TX 77843-3368 (josefs@math.tamu.edu).

‡Information Technology Laboratory, National Institute of Standards and Technology, 325 Broadway, Mail Stop 891.01, Boulder, CO 80305-3328 (zydrunas.gimbutas@nist.gov).

§Simons Center for Data Analysis, Simons Foundation, 160 Fifth Avenue, New York, NY 10010 and Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012-1110 (greengard@cims.nyu.edu).

ETNA
Kent State University
http://etna.math.kent.edu

in a suitable $(n-k)$-dimensional subspace. In terms of the singular value decomposition $A = U\Sigma V^*$, this corresponds to the case when $\sigma_1(A)/\sigma_{n-k}(A)$ is of modest size and $\sigma_{n-k+1}(A), \ldots, \sigma_n(A) = 0$, where the $\sigma_i(A)$ are the singular values of A. We do *not* address least squares problems, that is, we assume that the system (1.1), with or without (1.2), is consistent.

DEFINITION 1.1. *We will denote by $\mathcal{N}(A)$ the null space of A and by $\mathcal{R}(A)$ its range.*

There is a substantial literature on this subject, which we do not seek to review here. We refer the reader to the texts [15, 19] and the papers [2, 4, 5, 6, 7, 8, 9, 10, 14, 18, 20, 21, 30]. Of particular relevance are [24, 25, 26, 27, 28, 32], which demonstrate the power of randomized schemes using methods closely related to the ones described below. It is also worth noting that, in recent years, the use of randomization together with numerical rank-based ideas has proven to be a powerful combination for a variety of problems in linear algebra and theoretical computer science; see, for example, [17, 22, 29].

The basic idea in the present work is remarkably simple and summarized in the following theorem.

THEOREM 1.2. *Suppose A is a rank-1 deficient matrix and that $Ax = b$. Suppose further that $p \notin \mathcal{R}(A)$ and $q \notin \mathcal{R}(A^*)$. Then $(A + pq^*)y = b$ is a nonsingular system, and the solution satisfies $Ay = b$. Furthermore, the difference $x - y$ is in the null space of A.*

Proof. That $A + pq^*$ is nonsingular is implied by the fact that $p \notin \mathcal{R}(A)$ and $q \notin \mathcal{R}(A^*)$. It follows that $A(x-y) = b-(b-pq^*y) = p(q^*y)$. Since $A(x-y)$ must be in $\mathcal{R}(A)$ and p is not, both sides vanish, implying that $x - y$ is a null vector of A and q^*y must be zero. $Ay = b$ follows directly from $A(x - y) = 0$. ∎

Another perspective, which may be more natural to some readers, is to consider the affine space $\{x' + \mathcal{N}(A)\}$ consisting of solutions to $Az = b$, where, x' is the solution of minimal norm. The difference of any two vectors in the affine space clearly lies in the null space of A. If $A + pq^*$ is nonsingular, then y is the unique vector in the affine space orthogonal to q, implying that $x - y \in \mathcal{N}(A)$.

This suggests the following simple procedure for computing a null vector of a rank-1 deficient matrix A:

1. Choose a random vector $x \in \mathbb{C}^n$, and compute $b = Ax$.
2. Choose random vectors $p, q \in \mathbb{C}^n$, and solve

$$(A + pq^*)y = b. \tag{1.3}$$

Then, the difference $x - y$ is in the null space of A. Since p and q are random, the requirement $p \notin \mathcal{R}(A)$ and $q \notin \mathcal{R}(A^*)$ occurs with probability close to 1.

It is worth comparing the proposed method with a similar scheme in [27, 28] based on considering the system

$$(A + pq^*)y = p, \tag{1.4}$$

where p is a random vector in $\mathbb{C}^n$. By the same analysis, $Ay = p - pq^*y = p(1 - q^*y)$, and, since Ay is in the range of A and p is not, both $Ay = 0$ and $q^*y = 1$. This scheme can be viewed as dual to (1.3) since it enforces a non-homogeneous constraint on the solution y. By construction, equation (1.4) is unable to handle consistent right-hand sides since p can not be in the range of A in order for $A + pq^*$ to be invertible.

Our method extends the existing scheme (1.4) to handle an arbitrary consistent right-hand side in the range of A. In addition, the previous solutions can be reused more efficiently in iterative refinement settings. If the solution y must satisfy an additional non-homogeneous constraint, then equations (1.3) and (1.4) can be combined by solving $(A + pq^*)y = b + pw$,

ETNA
Kent State University
http://etna.math.kent.edu

where $b = Ax$ and w is an arbitrary constant, yielding $A(x - y) = 0$ and $Ay = b$ subject to $q^*y = w$.

The remainder of this note is intended to make the proposed procedure rigorous. While related algorithms have been described in the literature (particularly [24, 27, 28]), the scheme presented here provides a simple framework for solving a variety of problems such as (1.1), (1.2) in addition to the null space problem. It is easy to implement, permits iterative refinement in standard precision arithmetic, and is compatible with iterative solution techniques.

2. Mathematical preliminaries. Much of our analysis depends on estimating the condition number of a rank-k deficient complex $n \times n$ matrix A to which is added a rank-k random perturbation. For $P, Q \in \mathbb{C}^{n\times k}$, we let

$$\begin{aligned} P &= P_R + P_{N^*}, \qquad \mathcal{R}(P_R) \subset \mathcal{R}(A),\ \mathcal{R}(P_{N^*}) \subset \mathcal{N}(A^*), \\ Q &= Q_{R^*} + Q_N, \qquad \mathcal{R}(Q_{R^*}) \subset \mathcal{R}(A^*),\ \mathcal{R}(Q_N) \subset \mathcal{N}(A), \end{aligned} \tag{2.1}$$

and

$$\begin{aligned} \rho &:= \|P_R\| = \sigma_{\max}(P_R), \qquad \eta := \sigma_{\min}(P_{N^*}), \\ \xi &:= \|Q_{R^*}\| = \sigma_{\max}(Q_{R^*}), \qquad \nu := \sigma_{\min}(Q_N), \end{aligned} \tag{2.2}$$

where, for all norms, $\|\cdot\| = \|\cdot\|_2$.

THEOREM 2.1. *Let $b = Ax$ and let y be an approximate solution to*

$$(A + PQ^*)y = b$$

in that it satisfies

$$\|b - (A + PQ^*)y\| \le \delta. \tag{2.3}$$

Then

$$\|A(x - y)\| \le \delta \left(1 + \frac{\|P\|}{\sigma_{\min}(P_{N^*})}\right). \tag{2.4}$$

Proof. It follows from (2.3) and the triangle inequality that

$$\|A(x - y)\| \le \delta + \|P\|\|Q^*y\|. \tag{2.5}$$

Moreover,

$$b - Ay - P(Q^*y) = \delta f$$

for some vector $f \in \mathbb{C}^n$ with $\|f\| \le 1$. Now let U be a matrix whose columns form an orthonormal basis for $\mathcal{N}(A^*)$. Multiplying on the left by U^*, we have

$$-(U^*P)\,(Q^*y) = \delta(U^*f), \qquad \|Q^*y\| \le \frac{\delta}{\sigma_{\min}(P_{N^*})},$$

where the last inequality follows from the fact that

$$\delta \ge \inf_{\|z\|=1, z\in\mathbb{C}^k} \|U^*Pz\|\|Q^*y\| = \inf_{\|z\|=1, z\in\mathbb{C}^k} \|UU^*Pz\|\|Q^*y\| = \sigma_{\min}(P_{N^*})\|Q^*y\|,$$

which yields the desired result when combined with (2.5). ∎

ETNA
Kent State University
http://etna.math.kent.edu

The obtained bound (2.4) indicates that $x - y$ is an approximate null vector of the matrix A, therefore, y is also an approximate solution to $Ay = b$ for a given consistent right-hand side $b \in \mathcal{R}(A)$.

THEOREM 2.2. *Let $A \in \mathbb{C}^{n\times n}$ have a k-dimensional null space, and let P, $Q \in \mathbb{C}^{n\times k}$. Then*

$$\|(A + PQ^*)^{-1}\| \leq \frac{1}{\sigma_{n-k}(A)} \sqrt{1 + \left(\frac{\rho}{\eta}\right)^2 + \left(\frac{\xi}{\nu}\right)^2 + \left(\frac{\sigma_{n-k}(A) + \rho\xi}{\eta\nu}\right)^2},$$

where ρ, η, ξ, ν are defined in (2.2).

Proof. Let $A = U\Sigma V^*$ be the singular value decomposition of A. Let C and D be such that $P = UC$ and $Q = VD$. Let $C^T = [C_R^T \; C_{N^*}^T]$, where $C_R \in \mathbb{C}^{(n-k)\times k}$ and $C_{N^*} \in \mathbb{C}^{k\times k}$. The entries in the columns of C_R are coefficients of the corresponding columns of P in an orthonormal basis of the range of A. Thus $\|C_R\| = \rho$, and similarly, $\|C_{N^*}^{-1}\| = 1/\eta$. Let $D^T = [D_{R^*}^T \; D_N^T]$, where $D_{R^*} \in \mathbb{C}^{(n-k)\times k}$ and $D_N \in \mathbb{C}^{k\times k}$. By similar reasoning, we have that $\|D_{R^*}\| = \xi$ and $\|D_N^{-1}\| = 1/\nu$. Then

$$\|(A + PQ^*)^{-1}\| = \|(\Sigma + CD^*)^{-1}\|,$$

and

$$\begin{aligned}(\Sigma + CD^*)^{-1} &= \begin{bmatrix} \Sigma' + C_R D_{R^*}^* & C_R D_N^* \\ C_{N^*} D_{R^*}^* & C_{N^*} D_N^* \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \Sigma'^{-1} & -\Sigma'^{-1} C_R (C_{N^*})^{-1} \\ -(D_N^*)^{-1} D_{R^*}^* \, \Sigma'^{-1} & (D_N^*)^{-1} \left(I_k + D_{R^*}^* \, \Sigma'^{-1} C_R\right) (C_{N^*})^{-1} \end{bmatrix}, \end{aligned} \tag{2.6}$$

where $\Sigma' \in \mathbb{C}^{(n-k)\times(n-k)}$ is the upper $(n-k)\times(n-k)$ submatrix of Σ and $I_k \in \mathbb{C}^{k\times k}$ is the identity matrix. This gives

$$\begin{aligned}&\|(\Sigma + CD^*)^{-1}\| \\ &\leq \sqrt{\frac{1}{\sigma_{n-k}^2(A)} + \left(\frac{\rho}{\sigma_{n-k}(A)\,\eta}\right)^2 + \left(\frac{\xi}{\sigma_{n-k}(A)\,\nu}\right)^2 + \left(\frac{1 + \rho\xi/\sigma_{n-k}(A)}{\eta\nu}\right)^2} \\ &= \frac{1}{\sigma_{n-k}(A)} \sqrt{1 + \left(\frac{\rho}{\eta}\right)^2 + \left(\frac{\xi}{\nu}\right)^2 + \left(\frac{\sigma_{n-k}(A) + \rho\xi}{\eta\nu}\right)^2}. \qquad \square \end{aligned}$$

It follows from this result that one can bound the conditioning of the perturbed matrix.

THEOREM 2.3. *Let $A \in \mathbb{C}^{n\times n}$ have a k-dimensional null space, and let P, $Q \in \mathbb{C}^{n\times k}$. Then*

$$\kappa(A + PQ^*) \leq \frac{\sigma_1(A) + \|P\|\,\|Q\|}{\sigma_{n-k}(A)} \sqrt{1 + \left(\frac{\rho}{\eta}\right)^2 + \left(\frac{\xi}{\nu}\right)^2 + \left(\frac{\sigma_{n-k}(A) + \rho\xi}{\eta\nu}\right)^2},$$

where ρ, η, ξ, ν are defined in (2.2).

The estimates in Theorems 2.2 and 2.3 improve the upper bounds for the perturbed matrix given in [28]. The preceding theorems also indicate that, in the absence of additional information, it is reasonable to pick random vectors of approximately unit norm and multiply the perturbation term PQ^* by the norm of A.

REMARK 2.4. The above estimates are very pessimistic. For consistent right-hand sides, the inversion process involves only the first column of (2.6), therefore the solution accuracy mostly depends on the spectral properties of Q.

ETNA
Kent State University
http://etna.math.kent.edu

Since the condition number of the perturbed system largely depends on the projections of P and Q on generally unknown null spaces $\mathcal{N}(A^*)$ and $\mathcal{N}(A)$, respectively, the algorithm is relatively insensitive to the choice of random variables used to generate P and Q. In the context of sparse matrices, a fast algorithm is required to apply the perturbation term PQ^*; the random matrices can be constructed and applied using, for example, the fast Johnson-Lindenstrauss transform (FJLT) [1] or the subsampled randomized Fourier transform (SRFT) [29].

In this note, we use standard random Gaussian matrices whose elements are independent standard normal random variables. The behavior of the smallest singular values of such matrices is closely related to the spectral properties of Wishart-type matrices [11, 12, 17]. Since the distribution of a standard Gaussian matrix is invariant under projections and rotations, the parameter $\lambda_{\min} = \nu^2$ (or $\lambda_{\min} = \eta^2$) is distributed as the smallest eigenvalue of a $k \times k$ Wishart matrix. It is shown in [11] that, for the real-valued $k \times k$ Wishart matrices, the mathematical expectation of $\log(k\lambda_{\min})$ is finite, and, as $k \to \infty$,

$$\mathbf{E}[\log(k\lambda_{\min})] \to -1.68788\ldots$$

For complex-valued $k \times k$ Wishart matrices, a more precise statement can be made:

$$\mathbf{E}[\log(k\lambda_{\min})] = \log 2 - \gamma \approx 0.11593,$$

where $\gamma \approx 0.5772$ is Euler's constant. The above estimates show that, on average, the condition number of the perturbed matrix grows only moderately as the rank-deficiency increases. In order to estimate the probability that a perturbed matrix with a very large condition number may appear, we again refer the reader to [11, 12] for a more precise characterization of the tails of eigenvalue distributions for Wishart matrices.

3. Solving consistent, rank-deficient linear systems. Let us first consider the solution of the consistent, rank-k deficient linear system $Ax = b$ in the special case where $\mathcal{N}(A)$ and $\mathcal{N}(A^*)$ are spanned by the columns of known $n \times k$ matrices N and V, respectively. Suppose now that we solve the linear system

$$(A + VN^*)x = b\,. \tag{3.1}$$

It is then clear that $V^*Ax = V^*b = 0$, so that $(V^*V)(N^*x) = 0$, from which we get that $N^*x = 0$. Thus, x is the particular solution to $Ax = b$ that is orthogonal to the null space of A implying that x is the minimum-norm solution of $Ax = b$. From Theorem 2.3, the condition number of $A + VN^*$ is given by

$$\kappa(A + VN^*) \leq \frac{\sigma_1(A) + \|V\|\,\|N\|}{\sigma_{n-k}(A)} \sqrt{1 + \left(\frac{\sigma_{n-k}(A)}{\sigma_{\min}(V)\sigma_{\min}(N)}\right)^2}\,. \tag{3.2}$$

The estimate (3.2) shows that the condition number of the perturbed system is very nearly optimal, that is, approximately that of the original problem restricted to the range of A, namely σ_1/σ_{n-k}.

Suppose now that we have no prior information about the null spaces of A and/or A^*. We may then substitute random matrices P and Q for V and/or N and follow the same procedure. With probability close to 1, $(A + PQ^*)$ will be invertible, and we will obtain the particular solution to $Ax = b$ that is orthogonal to the columns of Q. This simply requires that the projections of P onto $\mathcal{N}(A^*)$ and of Q onto $\mathcal{N}(A)$, denoted by P_{N^*} and Q_N, respectively, must be full-rank; see (2.1). This implies that only a basis for $\mathcal{N}(A)$ is needed to compute the minimum-norm solution: with probability close to 1, it is given by the solution to

$$(A + PN^*)x = b.$$

ETNA
Kent State University
http://etna.math.kent.edu

REMARK 3.1. This procedure allows us to obtain the minimum-norm solution to the underdetermined linear system without recourse to the SVD or other dense matrix methods. Any method for solving (3.1) can be used. If the perturbed system is reasonably well-conditioned and A can be applied efficiently, Krylov space methods such as GMRES can be extremely effective.

REMARK 3.2. It is worth noting that under certain conditions, GMRES can be used directly on a singular or nearly singular system. This issue is carefully analyzed in [3].

3.1. Consistent, rectangular linear systems. We next consider the case where we wish to solve the system (1.1) together with (1.2). Note that the system

$$\begin{bmatrix} A \\ C^* \end{bmatrix} x = \begin{bmatrix} b \\ f \end{bmatrix} \tag{3.3}$$

is full-rank if and only if any vector in $\mathcal{N}(A)$ has a nontrivial projection onto the columns of C. There is no need, however, to solve a rectangular system of equations (3.3). One only needs to solve the $n \times n$ linear system

$$(A + VC^*)x = b + Vf\,.$$

If $\mathcal{R}(V) = \mathcal{N}(A^*)$, then from Theorem 2.3, the condition number of $A + VC^*$ is given by

$$\kappa(A + VC^*) \leq \frac{\sigma_1(A) + \|V\|\,\|C\|}{\sigma_{n-k}(A)} \sqrt{1 + \left(\frac{\xi}{\sigma_{\min}(C_N)}\right)^2 + \left(\frac{\sigma_{n-k}(A)}{\sigma_{\min}(V)\sigma_{\min}(C_N)}\right)^2},$$

where ξ is the norm of C_{R^*}.

In some applications, the data may be known to be consistent (b is in the range of A), but V may not be known. Then, one can proceed as above by solving

$$(A + PC^*)x = b + Pf\,,$$

where P is a random $n \times k$ matrix. From Theorem 2.3, the condition number of $A + PC^*$ is given by

$$\kappa(A + PC^*) \leq \frac{\sigma_1(A) + \|P\|\,\|C\|}{\sigma_{n-k}(A)} \times \sqrt{1 + \left(\frac{\rho}{\sigma_{\min}(P_{N^*})}\right)^2 + \left(\frac{\xi}{\sigma_{\min}(C_N)}\right)^2 + \left(\frac{\sigma_{n-k}(A) + \rho\xi}{\sigma_{\min}(P_{N^*})\sigma_{\min}(C_N)}\right)^2},$$

where ρ and ξ are the norms of P_R and C_{R^*}, respectively.

4. Computing the null space. Let us return now to the question of finding a basis for the null space of a rank-k deficient matrix $A \in \mathbb{C}^{n\times n}$. As in the introduction, we begin by describing the procedure:

1. Choose k random vectors $\{x_i, i = 1, \ldots, k\} \in \mathbb{C}^n$, and compute $b_i = Ax_i$.
2. Choose random matrices $P, Q \in \mathbb{C}^{n\times k}$, and solve

$$(A + PQ^*)y_i = b_i. \tag{4.1}$$

Then, $A(x_i - y_i) = b_i - (b_i - PQ^*y_i) = P(Q^*y_i)$. Since $A(x_i - y_i) \in \mathcal{R}(A)$ and assuming $P(Q^*y_i) \notin \mathcal{R}(A)$, it follows that both sides must equal zero and that each vector $z_i = x_i - y_i$ is a null vector. Since the construction is random, the probability that

ETNA
Kent State University
http://etna.math.kent.edu

the $\{z_i\}$ are linearly independent is 1. The result $P(Q^*y_i) \notin \mathcal{R}(A)$ follows from the fact that P is random and that the projection of each column of P onto $\mathcal{N}(A^*)$ will be linearly independent with probability close to 1. Theorem 2.3 tells us how to estimate the condition number of (4.1). Finally, the accuracy of the null vectors $\{z_i\}$ can be further improved by an iterative refinement $\tilde{z}_i = z_i - \tilde{y}_i$, where the correction vectors $\tilde{y}_i$ solve (4.1)

$$(A + PQ^*)\tilde{y}_i = \tilde{b}_i,$$

with the updated right-hand sides $\tilde{b}_i = Az_i$.

This version of iterative refinement works well in standard precision arithmetic. It is clear from (2.3) and (2.4) that the accuracy of computing the null space is controlled by the error parameter δ, which in turn scales proportionally to the norm of the right-hand side b. In practice, just one refinement step is necessary to fully tighten the null vectors.

4.1. Stabilization. Since the condition number of the randomly perturbed matrix is controlled only in a probabilistic sense, if high precision is required, then one can use a variant of iterative refinement to improve the solution. That is, one can first compute $q_1, \ldots, q_k$ as approximate null vectors of A and $p_1, \ldots, p_k$ as approximate null vectors of A^*.

With these at hand, one can repeat the calculation with P and Q whose columns are $\{p_1, \ldots, p_k\}$ and $\{q_1, \ldots, q_k\}$, respectively. The parameters ρ/η and ξ/ν in Theorem 2.3 will be much less than 1, and the condition number of a second iteration will be approximately

$$\kappa(A + PQ^*) \approx \frac{\sigma_1(A) + \|P\|\,\|Q\|}{\sigma_{n-k}(A)} \sqrt{1 + \left(\frac{\sigma_{n-k}(A)}{\sigma_{\min}(P_{N^*})\sigma_{\min}(Q_N)} \right)^2}.$$

4.2. Determining the dimension of the null space. When the dimension of the null space is unknown, the algorithm above can also be used as a *rank-revealing* scheme; see also [23]. For this, suppose that the actual rank-deficiency is k_A and that we carry out the above procedure with $k > k_A$. The argument that $P(Q^*y_i) \notin \mathcal{R}(A)$ will fail since the projection of each of the columns of P onto $\mathcal{N}(A^*)$ must be linearly dependent. As a result, $x_i - y_i$ will fail to be a null vector (which will be obvious from the explicit computation of $A(x_i - y_i)$). The estimated rank k can then be systematically reduced to determine k_A. If k_A is large, bisection can be used to accelerate this estimate.

5. Numerical experiments. In this section, we describe the results of several numerical tests of the algorithms discussed above. All computations were performed in IEEE double-precision arithmetic using MATLAB version R2012a [1].

We use a pseudorandom number generator (MATLAB's `randn`) to create $n \times 1$ vectors $\phi_1, \phi_2, \ldots, \phi_{n-k}$ and $\psi_1, \psi_2, \ldots, \psi_{n-k}$ with entries that are independent and identically distributed Gaussian random variables of zero mean and unit variance. We apply the Gram-Schmidt process with reorthogonalization to $\phi_1, \phi_2, \ldots, \phi_{n-k}$ and $\psi_1, \psi_2, \ldots, \psi_{n-k}$ to obtain orthonormal vectors $u_1, u_2, \ldots, u_{n-k}$ and $v_1, v_2, \ldots, v_{n-k}$, respectively. We define A to be the $n \times n$ matrix

$$A = \sum_{i=1}^{n-k} u_i \sigma_i v_i^*,$$

where $\sigma_i = 1/i$. The rank-deficiency of A is clearly equal to k.

[1] Any mention of commercial products or reference to commercial organizations is for information only; it does not imply recommendation or endorsement by NIST.

ETNA
Kent State University
http://etna.math.kent.edu

In Table 5.1, we compare the regular and stabilized versions of the new algorithm for finding the null space of a rank-deficient matrix A. The first and second columns contain the parameters n and k determining the size and the rank-deficiency of the problem, respectively. The third column contains the modified condition number σ_1/σ_{n-k} of the original matrix A ignoring the zero singular values for a more meaningful comparison between columns. The fourth column contains the true condition number σ_1/σ_n of a random rank-k perturbation $A + PQ^*$. Finally, the fifth and sixth columns contain the relative accuracy $\|AN\|/\|N\|$ in determining the null space N for the randomized rank-k correction scheme before and after iterative refinement, respectively.

In Table 5.2, we compare the accuracy of the regular and stabilized versions of the randomized rank-k correction scheme for solving a rank-deficient linear system $Ax = b$ with a consistent right-hand side b. The first and second columns contain the parameters n and k determining the size and the rank-deficiency of the problem, respectively. The third and fourth columns contain the modified condition number σ_1/σ_{n-k} of the original matrix A and the condition number σ_1/σ_n of a random rank-k perturbation $A + PQ^*$, respectively. The fifth column contains the condition number σ_1/σ_n of the rank-k perturbation $A + VN^*$, where V and N are the approximate null vectors spanning the left and right null spaces, respectively. Finally, the fifth and seventh columns contain the relative accuracy $\|Ax - b\|/\|b\|$ in determining the solution vector x for the regular and stabilized schemes, respectively.

It is clear from Table 5.2 that the condition number can be quite large for the non-stabilized version of the algorithm when the rank-deficiency is high. This is due to the difficulty of finding high-dimensional random matrices P and Q that have large projections onto the corresponding null spaces $\mathcal{N}(A^*)$ and $\mathcal{N}(A)$. In such cases, the algorithm will strongly benefit from the stabilization procedure.

6. Further examples. Our interest in the development of randomized methods was driven largely by issues in the regularization of integral equation methods in potential theory. For illustration, consider the Neumann problem for the Laplace equation in the interior of a simply-connected, smooth domain $\Omega \subset \mathbb{R}^2$ with boundary Γ.

$$\Delta u = 0 \text{ in } \Omega, \qquad \frac{\partial u}{\partial n} = f \text{ on } \Gamma\,.$$

Classical potential theory [16] suggests seeking the solution as a single layer potential

$$u(x) = \frac{1}{2\pi} \int_\Gamma \log \|x - y\| \sigma(y)\, ds_y\,.$$

Using standard jump relations, this results in the integral equation

$$\sigma(x) + \frac{1}{\pi} \int_\Gamma \frac{\partial}{\partial n_x} \log \|x - y\| \sigma(y)\, ds_y = 2f(x)\,, \tag{6.1}$$

which we write as

$$(I + K)\sigma = 2f\,.$$

It is well-known that (6.1) is solvable if and only if the right-hand side satisfies the compatibility condition $\int_\Gamma f(y) ds_y = 0$. Using the L_2 inner product (for real-valued functions)

$$\langle f, g \rangle = \int_\Gamma f(y) g(y) ds_y,$$

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 5.1
Relative errors in determining the null vectors for the randomized rank-k correction scheme before and after iterative refinement.

n	k	$\kappa(A)$	$\kappa(A+PQ^*)$	E_2	$E_2(ref)$
160	1	$1.6\,10^{+02}$	$2.0\,10^{+03}$	$1.4\,10^{-16}$	$8.1\,10^{-17}$
160	3	$1.6\,10^{+02}$	$4.3\,10^{+04}$	$2.2\,10^{-15}$	$2.7\,10^{-16}$
160	6	$1.5\,10^{+02}$	$1.1\,10^{+04}$	$2.7\,10^{-14}$	$6.4\,10^{-16}$
320	1	$3.2\,10^{+02}$	$5.3\,10^{+03}$	$9.1\,10^{-17}$	$3.6\,10^{-17}$
320	3	$3.2\,10^{+02}$	$9.3\,10^{+03}$	$1.9\,10^{-16}$	$6.0\,10^{-17}$
320	6	$3.1\,10^{+02}$	$3.4\,10^{+04}$	$7.5\,10^{-16}$	$2.5\,10^{-16}$
640	1	$6.4\,10^{+02}$	$3.9\,10^{+04}$	$1.9\,10^{-16}$	$2.1\,10^{-16}$
640	3	$6.4\,10^{+02}$	$1.3\,10^{+06}$	$3.9\,10^{-15}$	$5.8\,10^{-16}$
640	6	$6.3\,10^{+02}$	$3.9\,10^{+06}$	$5.9\,10^{-13}$	$5.8\,10^{-16}$
1280	1	$1.3\,10^{+03}$	$6.0\,10^{+06}$	$5.5\,10^{-16}$	$3.2\,10^{-16}$
1280	3	$1.3\,10^{+03}$	$4.0\,10^{+04}$	$1.0\,10^{-14}$	$6.9\,10^{-17}$
1280	6	$1.3\,10^{+03}$	$6.5\,10^{+05}$	$3.7\,10^{-15}$	$8.1\,10^{-16}$
160	75	$8.5\,10^{+01}$	$2.4\,10^{+05}$	$4.2\,10^{-13}$	$2.1\,10^{-14}$
160	80	$8.0\,10^{+01}$	$3.2\,10^{+04}$	$2.2\,10^{-13}$	$2.5\,10^{-15}$
320	155	$1.6\,10^{+02}$	$1.4\,10^{+06}$	$3.2\,10^{-12}$	$7.5\,10^{-15}$
320	160	$1.6\,10^{+02}$	$1.6\,10^{+06}$	$1.5\,10^{-11}$	$1.6\,10^{-14}$
640	315	$3.2\,10^{+02}$	$1.0\,10^{+07}$	$1.1\,10^{-11}$	$6.8\,10^{-15}$
640	320	$3.2\,10^{+02}$	$4.3\,10^{+06}$	$1.6\,10^{-11}$	$1.9\,10^{-14}$
1280	635	$6.4\,10^{+02}$	$3.5\,10^{+08}$	$2.7\,10^{-10}$	$4.3\,10^{-14}$
1280	640	$6.4\,10^{+02}$	$1.9\,10^{+08}$	$1.9\,10^{-11}$	$5.7\,10^{-14}$

we may write the compatibility condition as

$$\langle 1, f\rangle = 0\,,$$

where 1 denotes the function that is identically 1 on Γ. The function 1 is also in the null space of $I+K^*$, the adjoint of the integral operator in (6.1), which is clearly necessary for solvability. Following the procedure in Section 3, we regularize the integral equation by solving

$$\sigma(x) + \frac{1}{\pi}\int_\Gamma \frac{\partial}{\partial n_x}\log\|x-y\|\sigma(y)\,ds_y + \int_\Gamma [r(x)1(y)]\sigma(y)\,dy = 2f(x)\,, \tag{6.2}$$

or

$$(I+K)\sigma + r(x)\langle 1,\sigma\rangle = 2f\,,$$

where $r(x)$ is a random function defined on Γ. Taking the inner product of (6.2) with the function 1 yields

$$\langle 1, r\rangle\,\langle 1,\sigma\rangle = 0\,.$$

This is a well-known fact for the Neumann problem, and the obvious choice is simply $r(x)=1$, so that (6.2) becomes

$$\sigma(x) + \frac{1}{\pi}\int_\Gamma \left[\frac{\partial}{\partial n_x}\log\|x-y\| + 1\right]\sigma(y)\,ds_y = 2f(x)\,.$$

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 5.2
Relative errors for the regular and stabilized versions of the randomized rank-k correction scheme in determining the solution of the rank-k deficient linear system $Ax = b$ with the consistent right-hand side $b \in \mathcal{R}(A)$.

n	k	$\kappa(A)$	$\kappa(A+PQ^*)$	E_2	$\kappa(A+UV^*)$	$E_2(stab)$
160	1	$1.6\,10^{+02}$	$9.1\,10^{+02}$	$1.3\,10^{-15}$	$1.6\,10^{+02}$	$1.1\,10^{-15}$
160	3	$1.6\,10^{+02}$	$3.1\,10^{+03}$	$3.9\,10^{-15}$	$1.6\,10^{+02}$	$1.9\,10^{-15}$
160	6	$1.5\,10^{+02}$	$1.3\,10^{+06}$	$1.4\,10^{-13}$	$1.5\,10^{+02}$	$1.7\,10^{-15}$
320	1	$3.2\,10^{+02}$	$4.9\,10^{+05}$	$7.3\,10^{-15}$	$3.2\,10^{+02}$	$1.3\,10^{-15}$
320	3	$3.2\,10^{+02}$	$4.1\,10^{+05}$	$6.6\,10^{-14}$	$3.2\,10^{+02}$	$2.9\,10^{-15}$
320	6	$3.1\,10^{+02}$	$3.3\,10^{+04}$	$1.1\,10^{-14}$	$3.1\,10^{+02}$	$2.7\,10^{-15}$
640	1	$6.4\,10^{+02}$	$1.2\,10^{+05}$	$1.7\,10^{-14}$	$6.4\,10^{+02}$	$2.1\,10^{-15}$
640	3	$6.4\,10^{+02}$	$8.8\,10^{+04}$	$9.1\,10^{-15}$	$6.4\,10^{+02}$	$3.1\,10^{-15}$
640	6	$6.3\,10^{+02}$	$1.6\,10^{+05}$	$9.9\,10^{-15}$	$6.3\,10^{+02}$	$2.8\,10^{-15}$
1280	1	$1.3\,10^{+03}$	$8.3\,10^{+04}$	$4.5\,10^{-15}$	$1.3\,10^{+03}$	$3.5\,10^{-15}$
1280	3	$1.3\,10^{+03}$	$5.2\,10^{+05}$	$1.7\,10^{-14}$	$1.3\,10^{+03}$	$6.9\,10^{-15}$
1280	6	$1.3\,10^{+03}$	$7.7\,10^{+05}$	$3.9\,10^{-14}$	$1.2\,10^{+03}$	$4.7\,10^{-15}$
160	75	$8.5\,10^{+01}$	$7.1\,10^{+04}$	$3.8\,10^{-13}$	$8.5\,10^{+01}$	$4.2\,10^{-15}$
160	80	$8.0\,10^{+01}$	$2.4\,10^{+04}$	$9.3\,10^{-14}$	$8.0\,10^{+01}$	$3.9\,10^{-15}$
320	155	$1.6\,10^{+02}$	$1.7\,10^{+05}$	$1.9\,10^{-13}$	$1.6\,10^{+02}$	$1.2\,10^{-14}$
320	160	$1.6\,10^{+02}$	$9.4\,10^{+05}$	$6.1\,10^{-12}$	$1.6\,10^{+02}$	$8.9\,10^{-15}$
640	315	$3.2\,10^{+02}$	$5.5\,10^{+07}$	$8.5\,10^{-11}$	$3.2\,10^{+02}$	$2.6\,10^{-14}$
40	320	$3.2\,10^{+02}$	$2.6\,10^{+07}$	$1.6\,10^{-11}$	$3.2\,10^{+02}$	$1.9\,10^{-14}$
1280	635	$6.4\,10^{+02}$	$5.9\,10^{+06}$	$7.5\,10^{-12}$	$6.5\,10^{+02}$	$3.2\,10^{-14}$
1280	640	$6.4\,10^{+02}$	$1.1\,10^{+07}$	$1.2\,10^{-11}$	$6.4\,10^{+02}$	$7.5\,10^{-14}$

For an application of the preceding analysis in electromagnetic scattering, see [31]. In [13], a situation of the type discussed in Section 3.1 arises. Without entering into details, it was shown that the "magnetic field integral equation" is rank-k deficient in the static limit in exterior multiply-connected domains of genus k. A set of k nontrivial constraints was derived from electromagnetic considerations, which were added to the system matrix as described above. Since we have illustrated the basic principle in the context of the null space problem, we omit further numerical calculations.

7. Conclusions. We have presented a simple set of tools for solving rank-deficient, but consistent, linear systems and demonstrated their utility with some numerical examples. Since the perturbed/augmented linear systems are reasonably well-conditioned with high probability, one can rely on Krylov subspace based iterative methods (e.g., conjugate gradient for self-adjoint problems or GMRES for non-self-adjoint problems) avoiding the cost of dense linear algebraic methods such as Gaussian elimination or the SVD itself. This is a particularly powerful approach when A is sparse or when there is a fast algorithm for applying A to a vector. Finite rank-deficiency issues arise in the continuous setting as well, especially in integral equation methods, which we have touched on only briefly here.

We are currently working on the development of robust software for the null space problem that we expect will be competitive with standard approaches such as QR-based schemes [4], inverse iteration [9, 15], or Arnoldi methods [14].

Acknowledgment. We thank Mark Tygert for many helpful discussions.

ETNA
Kent State University
http://etna.math.kent.edu

REFERENCES

[1] N. AILON AND B. CHAZELLE, *The fast Johnson-Lindenstrauss transform and approximate nearest neighbors*, SIAM J. Comput., 39 (2009), pp. 302–322.
[2] J. BARLOW AND U. VEMULAPATI, *Rank detection methods for sparse matrices*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 1279–1297.
[3] P. N. BROWN AND H. F. WALKER, *GMRES on (nearly) singular systems*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 37–51.
[4] T. F. CHAN, *Rank revealing QR factorizations*, Linear Algebra Appl., 88/89 (1987), pp. 67–82.
[5] K. L. CLARKSON AND D. P. WOODRUFF, *Low rank approximation and regression in input sparsity time*, in Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing (STOC'13), ACM, New York, 2013, pp. 81–90.
[6] T. F. COLEMAN AND A. POTHEN, *The null space problem I: complexity*, SIAM J. Algebraic Discrete Methods, 7 (1986), pp. 527–537.
[7] ———, *The null space problem II: algorithms*, SIAM J. Algebraic Discrete Methods, 8 (1987), pp. 544–563.
[8] A. DASGUPTA, P. DRINEAS, B. HARB, R. KUMAR, AND M. W. MAHONEY, *Sampling algorithms and coresets for l_p regression*, SIAM J. Comput., 38 (2009), pp. 2060–2078.
[9] I. S. DHILLON, *Current inverse iteration software can fail*, BIT, 38 (1998), pp. 685–704.
[10] P. DRINEAS AND M. W. MAHONEY, *A randomized algorithm for a tensor-based generalization of the SVD*, Linear Algebra Appl., 420 (2007), pp. 553–571.
[11] A. EDELMAN, *Eigenvalues and condition numbers of random matrices*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 543–560.
[12] ———, *The distribution and moments of the smallest eigenvalue of a random matrix of Wishart type*, Linear Algebra Appl., 159 (1991), pp. 55–80.
[13] C. L. EPSTEIN, Z. GIMBUTAS, L. GREENGARD, A. KLÖCKNER, AND M. O'NEIL, *A consistency condition for the vector potential in multiply-connected domains*, IEEE Trans. Magn., 49 (2013), pp. 1072–1076.
[14] G. H. GOLUB AND C. GREIF, *An Arnoldi-type algorithm for computing PageRank*, BIT, 46 (2006), pp. 759–771.
[15] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.
[16] R. B. GUENTHER AND J. W. LEE, *Partial Differential Equations of Mathematical Physics and Integral Equations*, Prentice-Hall, Englewood Cliffs, 1988.
[17] N. HALKO, P. G. MARTINSSON, AND J. TROPP, *Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions*, SIAM Rev., 53 (2011), pp. 217–288.
[18] P. C. HANSEN, *Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank*, SIAM J. Sci. Statist. Comput., 11 (1990), pp. 503–518.
[19] P. C. HANSEN, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, 1998.
[20] M. E. HOCHSTENBACH AND L. REICHEL, *Subspace-restricted singular value decompositions for linear discrete ill-posed problems*, J. Comput. Appl. Math., 235 (2010), pp. 1053–1064.
[21] I. C. F. IPSEN, *Computing an eigenvector with inverse iteration*, SIAM Rev., 39 (1997), pp. 254–291.
[22] E. LIBERTY, F. WOOLFE, P. G. MARTINSSON, AND M. TYGERT, *Randomized algorithms for the low-rank approximation of matrices*, Proc. Natl. Acad. Sci. USA, 104 (2007), pp. 20167–20172.
[23] V. PAN, D. IVOLGIN, B. MURPHY, R. E. ROSHOLT, I. TAJ-EDDIN, Y. TANG, AND X. YAN, *Additive preconditioning and aggregation in matrix computations*, Comput. Math. Appl., 55 (2008), pp. 1870–1886.
[24] V. Y. PAN, D. IVOLGIN, B. MURPHY, R. E. ROSHOLT, Y. TANG, AND X. YAN, *Additive preconditioning for matrix computations*, Linear Algebra Appl., 432 (2010), pp. 1070–1089.
[25] V. Y. PAN AND G. QIAN, *Randomized preprocessing of homogeneous linear systems of equations*, Linear Algebra Appl., 432 (2010), pp. 3272–3318.
[26] ———, *Solving linear systems of equations with randomization, augmentation and aggregation*, Linear Algebra Appl., 437 (2012), pp. 2851–2876.
[27] V. Y. PAN AND X. YAN, *Null space and eigenspace computations with additive preprocessing*, in Proceedings of the 2007 International Workshop on Symbolic-Numeric Computation (SNC07), J. Verschelde and S. M. Watt, eds., ACM, New York, 2007, pp. 152–160.
[28] ———, *Additive preconditioning, eigenspaces, and the inverse iteration*, Linear Algebra Appl., 430 (2009), pp. 186–203.
[29] V. ROKHLIN AND M. TYGERT, *A fast randomized algorithm for overdetermined linear least-squares regression*, Proc. Natl. Acad. Sci. USA, 105 (2008), pp. 13212–13217.
[30] G. W. STEWART, *Rank degeneracy*, SIAM J. Sci. Statist. Comput., 5 (1984), pp. 403–413.
[31] F. VICO, Z. GIMBUTAS, L. GREENGARD, AND M. FERRANDO-BATALLER, *Overcoming low-frequency breakdown of the magnetic field integral equation*, IEEE Trans. Antennas and Propagation, 61 (2013), pp. 1285–1290.

[32] X. WANG, *Effect of small rank modification on the condition number of a matrix*, Comput. Math. Appl., 54 (2007), pp. 819–825.