Journal of Computational and Applied Mathematics 303 (2016) 206–217

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Adaptive cross approximation for ill-posed problems☆

T. Mach [a], L. Reichel [b], M. Van Barel [a,*], R. Vandebril [a]

[a] Department of Computer Science, KU Leuven, Celestijnenlaan 200A, 3001 Leuven (Heverlee), Belgium
[b] Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA

ARTICLE INFO

Article history:
Received 11 August 2015
Received in revised form 21 January 2016

Keywords:
Ill-posed problem
Inverse problem
Sparse discretization
Regularization
Adaptive cross approximation

ABSTRACT

Integral equations of the first kind with a smooth kernel and perturbed right-hand side, which represents available contaminated data, arise in many applications. Discretization gives rise to linear systems of equations with a matrix whose singular values cluster at the origin. The solution of these systems of equations requires regularization, which has the effect that components in the computed solution connected to singular vectors associated with small singular values are damped or ignored. In order to compute a useful approximate solution typically approximations of only a fairly small number of the largest singular values and associated singular vectors of the matrix are required. The present paper explores the possibility of determining these approximate singular values and vectors by adaptive cross approximation. This approach is particularly useful when a fine discretization of the integral equation is required and the resulting linear system of equations is of large dimensions, because adaptive cross approximation makes it possible to compute only fairly few of the matrix entries.

1. Introduction

This paper considers the approximate solution of Fredholm integral equations of the first kind,

$$\int_{\Omega_1} \kappa(s,t)x(t)dt = g(s), \quad s \in \Omega_2, \tag{1.1}$$

with a smooth kernel κ. The Ω_i are subsets of $\mathbb{R}^{d_i}$ for some positive integers d_i, $i = 1, 2$. Integral equations of this form arise in many applications, including remote sensing, computerized tomography, and image restoration. The solution of (1.1) is a so-called ill-posed problem. A reason for this is that the singular values of the integral operator cluster at the origin; see, e.g., [1,2].

Discretization of (1.1) by a Galerkin, Petrov–Galerkin, or Nyström method yields a linear system of equations

$$A\boldsymbol{x} = \boldsymbol{g}, \quad A \in \mathbb{R}^{n\times n}, \ \boldsymbol{g} \in \mathbb{R}^n, \tag{1.2}$$

☆ The research was partially supported by the Research Council KU Leuven, project OT/10/038 (Multi-parameter model order reduction and its applications), PF/10/002 Optimization in Engineering Centre (OPTEC), CREA-13-012 (Can Unconventional Eigenvalue Algorithms Supersede the State of the Art), OT/11/055 (Spectral Properties of Perturbed Normal Matrices and their Applications), by the Fund for Scientific Research—Flanders (Belgium), G.0828.14N (Multivariate polynomial and rational interpolation and approximation), G.0342.12N (Reestablishing Smoothness for Matrix Manifold Optimization via Resolution of Singularities), and by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office, Belgian Network DYSCO (Dynamical Systems, Control, and Optimization). This research also is supported in part by NSF grant DMS-1115385.

* Corresponding author.
E-mail addresses: thomas.mach@cs.kuleuven.be (T. Mach), reichel@math.kent.edu (L. Reichel), marc.vanbarel@cs.kuleuven.be (M. Van Barel), raf.vandebril@cs.kuleuven.be (R. Vandebril).

http://dx.doi.org/10.1016/j.cam.2016.02.020
0377-0427/

with a matrix with many singular values close to the origin. The matrix may be very ill-conditioned already for small to moderate values of n, where we measure the conditioning as the ratio between the largest and smallest singular values of A. In fact, A may be singular. Linear systems of equations with a matrix of this kind are commonly referred to as discrete ill-posed problems. We will for notational simplicity in this paper assume the matrix A to be square, however, the methods described can also be applied after minor modifications when A is rectangular, in which case the linear system of equations (1.2) may be replaced by a least-squares problem.

In many applications, the right-hand side vector $\boldsymbol{g}$ represents measured data and is contaminated by measurement and discretization errors. Due to these errors and the ill-conditioning of A, straightforward solution of (1.2) typically yields a computed solution that is severely contaminated by propagated error and is therefore not useful. To circumvent this difficulty, the system (1.2) is commonly replaced by a nearby problem that is less sensitive to the error in $\boldsymbol{g}$. This replacement is referred to as regularization. The possibly most popular regularization methods include truncated singular value decomposition (TSVD) and Tikhonov regularization.

Define the singular value decomposition (SVD)

$$A = U\Sigma V^T, \tag{1.3}$$

where $U = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n] \in \mathbb{R}^{n\times n}$ and $V = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n] \in \mathbb{R}^{n\times n}$ are orthogonal matrices and

$$\Sigma = \operatorname{diag}[\sigma_1, \sigma_2, \ldots, \sigma_n] \in \mathbb{R}^{n\times n}$$

is a diagonal matrix. Its nontrivial entries are the singular values of A; they are ordered according to $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_n \geq 0$. The columns of U and V are commonly referred to as the left and right singular vectors of A, respectively. The superscript ^{T} denotes transposition. The matrix

$$A_k = \sum_{j=1}^{k} \sigma_j \boldsymbol{u}_j \boldsymbol{v}_j^T \tag{1.4}$$

is a closest matrix of rank at most k to A in the spectral norm; see, e.g., [3]. The TSVD method determines, for some suitable $k \geq 0$, the solution of minimal Euclidean norm, denoted by $\boldsymbol{x}_k$, of the least-squares problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|A_k\boldsymbol{x} - \boldsymbol{g}\|. \tag{1.5}$$

Here and throughout this paper $\|\cdot\|$ stands for the Euclidean vector norm or the spectral matrix norm. The parameter k is a regularization parameter that determines how many singular values and vectors of A are used to compute the approximate solution $\boldsymbol{x}_k$ of (1.2).

Tikhonov regularization replaces the system (1.2) by the penalized least-squares problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n}\{\|A\boldsymbol{x} - \boldsymbol{g}\|^2 + \mu\|\boldsymbol{x}\|^2\}, \tag{1.6}$$

which has a unique solution $\boldsymbol{x}_\mu$ for any positive value of the regularization parameter μ. Substituting the SVD (1.3) into (1.6) shows that Tikhonov regularization dampens the contributions to $\boldsymbol{x}_\mu$ of singular values and vectors with large index k the most; increasing $\mu > 0$ results in more damping. We refer to [4,1,5,2,6–8] for details and computed examples with these regularization methods.

The determination of suitable values of the regularization parameters, k in (1.5) and μ in (1.6), is important for the quality of the computed approximate solution. Several methods have been described in the literature including the discrepancy principle, the L-curve criterion, and generalized cross validation; see [9–11] for recent discussions of their properties and illustrations of their performance. Regularization methods typically require that regularized solutions for several parameter values be computed and compared in order to determine a suitable value.

The present paper is concerned with the situation when the data vector $\boldsymbol{g}$ in (1.2) is of high dimension. Then the matrix A is large. The repeated solution of (1.6) can be carried out by iterative methods; see, e.g., [4,1,5–7]. These methods require the evaluation of matrix–vector products with A, and possibly with A^T as well, and this can be expensive when A is large. Moreover, all entries of the matrix have to be computed. There are iterative methods for computing the first k singular values and associated singular vectors of the matrix A; see, e.g., [12–14]. These methods also require matrix–vector product evaluations with A and A^T, as well as the evaluation of all matrix elements.

Cross approximation, sometimes also referred to as skeleton approximation, of matrices has been proposed as an approach to approximate a large dense matrix by a matrix of low rank; see, e.g., [15–19] and references therein. This method seeks to select a subset of k rows and columns of the matrix A to obtain a matrix $M_k \in \mathbb{R}^{n\times n}$ of rank at most k so that $\|A - M_k\|$ is small. Due to the optimality of the SVD of A, we have

$$\|A - A_k\| \leq \|A - M_k\|.$$

However, M_k is much cheaper to compute than A_k. In particular, the determination of M_k does not require that all entries of A be evaluated. The good performance of cross approximation for the approximation of a large matrix A by a matrix M_k of low rank k is well documented in the literature; see, e.g., [15–19]. However, we are not aware of applications of cross