Computational Techniques in Linear Algebra : from the S. Kaniel's Results to Today's Solution of Large Problems in Linear Algebra

Michel Bercovier

In homage to S. Kaniel to celebrate his (not so long ago) 75th Birthday

Born 8/6/1934

1955-1959 studies at the Hebrew University of Jerusalem Mathematics, Physics and statistics.

1959-1961 Ph.D. under the guidance of Prof. S. Agmon. Thesis title: *Unbounded normal operators*.

1961-1963 Visiting Assistant Professor– The University of Chicago.

1963-1964 Assistant Professor – The University of Chicago.

1964-1965 Visiting Assistant Professor– Stanford University

1965-1967 Research Associate – The Hebrew University

1967-1969 Assistant Professor – The Hebrew University.

1970-1981 Associate Professor – The Hebrew University

1977-1980 Chairman, Computer Science – The Hebrew University

1981- 2002 Professor – The Hebrew University

2002 – Professor Emeritus, The Hebrew University

1970-1971 Visiting Associate Professor– Northwestern University

1974-1975 Visiting Associate Professor– University of California, Los Angeles

1975-1976 Visiting Scientist – Victoria University, Canada

1980 – 1981 Einstein Fellow – The Institute of Advanced Study, Princeton.

Math of Comp 20 (1966)

Estimates for Some Computational Techniques in Linear Algebra

By Shmuel Kaniel*

1. In this paper we shall be concerned with two methods. The first one is that of conjugate gradients or least squares used for approximate computation of $Ax = f$ where A is a positive definite (symmetric) $n \times n$ matrix [2, 3, 6, 11]. It amounts to minimizing (with respect to the α_i) the expression:

$$\left\| A \sum_{i=0}^{k-1} \alpha_i A^i f - f \right\| \tag{1.1}$$

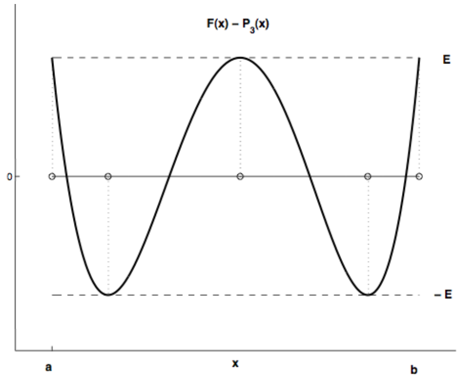
where k is a predetermined integer (usually much smaller than n) and where $\| \ \|$ denotes the L^2 norm. Then $g = \sum_{i=0}^{k-1} \alpha_i A^i f$ is taken as an approximate solution.

The second method is the generalized gradient or minimal iteration method used for an approximate computation of eigenvalues and eigenfunctions [3, 5].

It can be described as follows: Denote by H_k the subspace spanned by the vectors $f, Af, \cdots, A^k f$; denote by P the orthogonal projection on H_k and by B the restriction of PA to H_k (it is a well-defined $k + 1 \times k + 1$ matrix). Then $\mu_{\max}$, the largest eigenvalue of B is an approximation to $\lambda_{\max}$, the largest eigenvalue of A, likewise $\mu_{\min}$ approximates $\lambda_{\min}$ (we exclude here certain singular cases). There are, though, quite a few computational techniques for a numerical solution of this problem.

Central Result

Consider the minimax polynomial approximation I.e. the polynomial of degree n that approximates a given function in the given interval such that the absolute maximum error is minimized. The error is defined here as the difference between the function and the polynomial.



Let g be the approximate solution of Ax = f which is constructed by the conjugate gradients method. Then

$$\| Ag - f \| \leq \left(\sum_{j=1}^{k+1} \left| \frac{\prod_{i \neq j} \mu_i}{\prod_{i \neq j} (\mu_j - \mu_i)} \right| \right)^{-1} \| f \|,$$

where the μ_i are the eigenvalues of B

Central Result

Let g be the approximate solution of Ax = f which is constructed by the conjugate gradients method. Then

$$\| Ag - f \| \leq \left(\sum_{j=1}^{k+1} \left| \frac{\prod_{i \neq j} \mu_i}{\prod_{i \neq j} (\mu_j - \mu_i)} \right| \right)^{-1} \| f \|,$$

where the μ_i are the eigenvalues of B

B is k+1 X k+1 matrix, restriction of PA to the iteration subspace spanned by (f,Af, …., A^k)

Example. If $\lambda_{\min}/\lambda_{\max}$ *is small,* $\mu_{\min} = \mu_1 < \mu_2 < \cdots < \mu_{2k+1} = \mu_{\max}$ *and* $\mu_{i+1} - \mu_i = d$, *then estimate* (4.2) *yields:*

$$\| Ag - f \| < \left[\frac{\mu_{\min}}{\mu_{\max}} \binom{2k}{k} \right]^{-1} \| f \| \sim \frac{\mu_{\max}}{\mu_{\min}} \cdot k^{1/2} \cdot 2^{-2k} \| f \|.$$

Central Result II

μ_i *eigenvalues of B is k+1 X k+1 matrix, restriction of PA to the iteration subspace spanned by (f,Af, …., A^k)* **,** ***Ψ_i*** the corresponding eigenvector**,** ***λ_j*** eigenvalues of A,and ***Φ_j*** the eigenvectors

THEOREM 5.1. *Let d denote the distance between $\lambda_{\max}$ and the rest of the eigenvalues. Let f be normalized: $\| f \| = 1$ and let b denote $(f, \phi_{\max})$. Let k be chosen and B be constructed. Then:*

$$\lambda_{\max} \geqq \mu_{\max} \geqq \lambda_{\max} - \frac{\lambda_{\max} - \lambda_{\min}}{b^2}\left[T_k\left(\frac{\lambda_{\max} - \lambda_{\min} + d}{\lambda_{\max} - \lambda_{\min} - d}\right)\right]^{-2}. \tag{5.1}$$

Follow some refinement exploiting the structure of the eigenvalues distribution.

LANCZOS ALGORITHM

C.C. Paige *Computational variants of the Lanczos method for the eigenproblem, JIMA 10, 1972*

Y. Saad *On the rates of convergence of the Lanczos and block Lanczos Methods* , SIAM An. Num . 17, 1980

```
Algorithm Lanczos
 v_1 ← random vector with norm 1.
 v_0 ← 0
 β_1 ← 0
 Iteration: for j = 1,2,···,m
      w_j ← A v_j − β_j v_{j−1}
      α_j ← (w_j, v_j)
      w_j ← w_j − α_j v_j
      β_{j+1} ← ||w_j||
      v_{j+1} ← w_j / β_{j+1}
 return
```

$$T_{mm} = \begin{pmatrix} \alpha_1 & \beta_2 & & & & 0 \\ \beta_2 & \alpha_2 & \beta_3 & & & \\ & \beta_3 & \alpha_3 & \ddots & & \\ & & \ddots & \ddots & \beta_{m-1} & \\ & & & \beta_{m-1} & \alpha_{m-1} & \beta_m \\ 0 & & & & \beta_m & \alpha_m \end{pmatrix}$$

The vectors v_j (Lanczos vectors) generated on the fly defines the transformation matrix *Vmm*, with *Tmm=Vmm' A Vmm*

LANCZOS ALGORITHM

The vectors v_j (Lanczos vectors) generated on the fly defines the transformation matrix *Vmm*, with *Tmm=Vmm' A Vmm*

The eigenvalues and eigenvectors of *Tmm* are approximations of the largest and smallest eigen elements of *A*

Usefull for very large sparce matrix problems:

Ranking algorithms (web)
Latent Semantic Indexing (text queries and mining)
(based on SVD A=USV', AA'=VS^2V')
Dimension reduction
Molecular dynamics

LANCZOS ALGORITHM

* *Ranking algorithms (web)*
Latent Semantic Indexing (text queries and mining)
(based on SVD A=USV', AA'=VS^2V')
Dimension Reduction
Molecular dynamics

Latent Semantic Indexing called such because of its ability to correlate semantically related terms that are latent in a collection of text, it was first applied to text at Bell Laboratories in the late 1980s. The method, also called Latent Semantic Analysis (LSA), uncovers the underlying latent semantic structure in the usage of words in a body of text and how it can be used to extract the meaning of the text in response to user queries, commonly referred to as concept searches

LSI uses common linear algebra techniques to learn the conceptual correlations in a collection of text. In general, the process involves constructing a weighted term-document matrix, performing a Singular Value Decomposition on the matrix, and using the matrix to identify the concepts contained in the text.

LANCZOS ALGORITHM

** Ranking algorithms (web)*
Latent Semantic Indexing (text queries and mining)
(based on SVD A=USV', AA'=VS^2V')
Dimension Reduction
Molecular dynamics

Saad 2009 (Lanczos Vectors versus Singular Vectors for Effective Dimension Reduction IEEE Trans.on Knowledge and Data Engineering.)
takes an in-depth look at a technique for computing filtered matrix-vector (mat-vec) products which are required in many data analysis applications.

In these applications, the data matrix is multiplied by a vector and we wish to perform this product accurately in the space spanned by a few of the major singular vectors of the matrix.

The Lanczos-based approach achieves this goal by using a small number of Lanczos vectors, but it does not explicitly compute singular values/vectors of the matrix. The main advantage of the Lanczos-based technique is its low cost when compared with that of the truncated SVD.

LANCZOS ALGORITHM

* *Ranking algorithms (web)*

The basic idea is that hub and authority style algorithms are intimately related to eigenvector or singular value decompositions (depending on whether the links are symmetrical). This also means that there is a close relationship to asymptotic beahavior of random walks on the graph.

>

If you represent the linkage in the web by a matrix that has columns representing source page and rows representing the target page and with a 1 where-ever the source page has a link pointing to the target page, then if you start with a vector with a single non-zero element equal to 1 as a representation of a single page, then multiplying by the linkage matrix will give you a vector with 1 in the positions corresponding to the pages the original page linked to. If you multiply again, you get all the pages That you can get to in two steps from the original page etc…

$$A^n = U S^n U$$

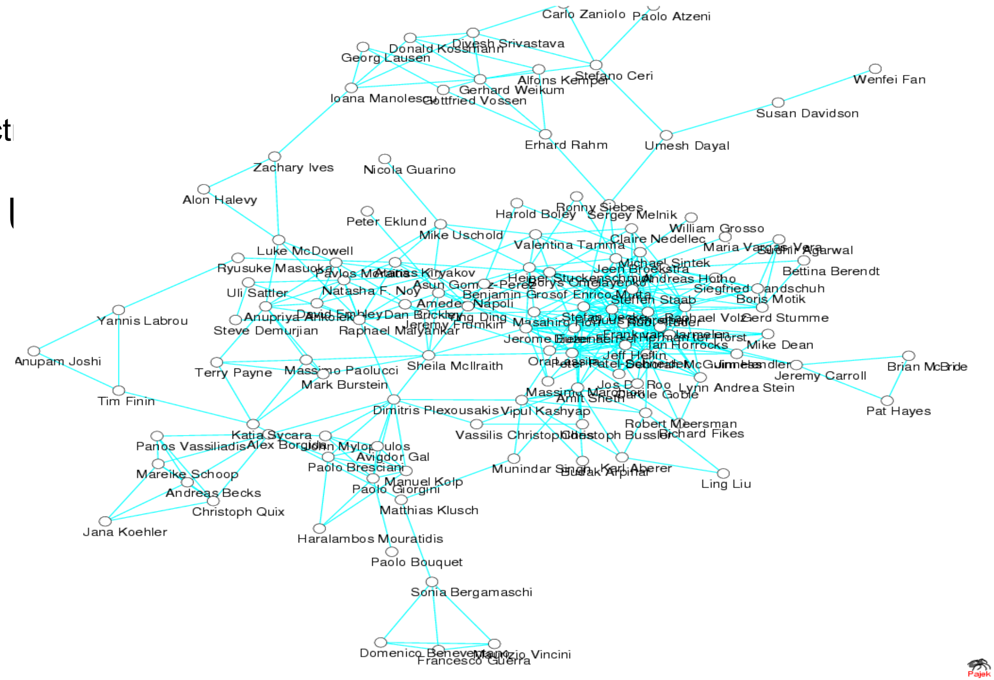
LANCZOS ALGORITHM

* *Ranking algorithms (web)*

> Connected (or nearly connected) clusters of pages can also be derived from

- the eigenvector
- decomposition.
- This is the basis
- of so-called spect
- > clustering.

$A^n = U S^n$



LANCZOS ALGORITHM

* *Ranking algorithms (web)*

> Connected (or nearly connected) clus

- the eigenvector
- decomposition.
- This is the basis
- of so-called spectral

> clustering.

$$A^n = U S^n U$$

LANCZOS ALGORITHM

workshop on large-scale eigenvalue problems held at Argonne National Laboratory. I was stunned by the progress made in the area since the late sixties and early seventies, when I had worked on the problem. The details of these new and improved algorithms are the subject of the second half of this volume. But any unified exposition makes it easy to forget that the foundations for these advances had been laid by people who were willing to work in an unexplored field, where confusion and uncertainty were the norm. Naming names carries the risk of omitting someone, but any list would surely include Vel Kahan, Shmuel Kaniel, Chris Paige, Beresford Parlett, and Yousef Saad. This book is dedicated to these pioneers and their coworkers.

G. W. Stewart
College Park, MD

Citations are not limited to Mathematics

Kaniel v. Minister of Justice (1973) 27 PD I 794 high on citation index
-The Israeli Supreme Court ruled on three separate occasions that these Basic Laws could not be used to overrule laws passed by the Knesset (Kaniel v. Minister of Justice (1973) 27 PD I 794)
These law and court rulings reflected the belief …that a true democracy was dominated by the elected parliament.
(in some other democracies) the government consisted of a balance between the parliament , and the judicial branch, and the executive branch.
Overtime these ideas have also been adopted in Israel by the new generation of leaders, and thus the Israeli Supreme Court has become more empowered to oversee laws passed by the Knesset
http://www.israellawresourcecenter.org/israelcourtrulings/essays/israelicourtrulingsessay.htm
----
HCJ 9333/03 Professor Shmuel Kaniel et al. v. Government of Israel [2005]
www.wcl.american.edu/journal/lawrev/56/edrey.pdf

The Israeli Supreme Court has adopted an interesting approach. In Kaniel v. Minister of Justice, (K) argued that the tax reform of 2003,57 which introduced significant tax preferences for capital gains (15% as opposed to 50% for top marginal tax rate for ordinary income), violated basic constitutional rights and principles……
The Israeli Supreme Court did not reject any of the petitioner's arguments.
Headed by President Barak, the court acknowledged that tax laws, … are subject to judicial- constitutional review.
Under this theory, the court could then consider tax preferences for capital gains to be a violation of constitutional rights and principles.
Yet, under Israeli constitutional law, a mere violation of a constitutional right or interest is not enough to declare a law void. In Kaniel, the court found that "for the time being," the apparent violation was justified
According to the court, the tax reforms served important public purposes, to both the states, the community and the taxpayers, …It is possible that as time passes a new reality will be created where the violation of equality to further societal goals will no longer be appropriate. Under the current circumstances, however, I cannot conclude that this balancing act is incorrect.(Judge Barak)
