On the Vorobyev method of moments

Zdeněk Strakoš
Charles University in Prague and Czech Academy of Sciences

http://www.karlin.mff.cuni.cz/˜strakos

Conference in honor of Volker Mehrmann
Berlin, May 2015

Thanks, bounds for 1955

Gene Golub, for pushing me to moments
Bernd Fischer, for the beautiful book and much more
Gérard Meurant, for many moment related joint interests
Claude Brezinski, for pointing out the work of Vorobyev
Jǒrg Liesen, for sharing interests and many years of collaboration

Volker Mehrmann, for lasting inspiration and support in many ways.

1954: Operator orthogonal polynomials and approximation methods
for determination of the spectrum of linear operators.

1958 (1965): Method of moments in applied mathematics.

Broader context of 1955

- Euclid (300BC), Hippassus from Metapontum (before 400BC), ... ,
- Bhascara II (around 1150), Brouncker and Wallis (1655-56): Three term recurences (for numbers)
- Euler (1737, 1748), ... , Brezinski (1991), Khrushchev (2008)
- Gauss (1814), Jacobi (1826), Christoffel (1858, 1857), ... , Chebyshev (1855, 1859), Markov (1884), Stieltjes (1884, 1893-94): Orthogonal polynomials, quadrature, analytic theory of continued fractions, problem of moments, minimal partial realization, Riemann-Stieltjes integral
  Gautschi (1981, 2004), Brezinski (1991), Van Assche (1993), Kjeldsen (1993),
- Hilbert (1906, 1912), ... , Von Neumann (1927, 1932), Wintner (1929) resolution of unity, integral representation of operator functions in quantum mechanics

Broader context of 1955

- Krylov (1931), Lanczos (1950, 1952, 1952c), Hestenes and Stiefel (1952), Rutishauser (1953), Henrici (1958), Stiefel (1958), Rutishauser (1959), ... , Vorobyev (1954, 1958, 1965), Golub and Welsh (1968), ... , Laurie (1991 - 2001), ...
- Gordon (1968), Schlesinger and Schwartz (1966), Steen (1973), Reinhard (1979), ... , Horáček (1983-...), Simon (2007)
- Paige (1971), Reid (1971), Greenbaum (1989), ...
- Magnus (1962a,b), Gragg (1974), Kalman (1979), Gragg, Lindquist (1983), Gallivan, Grimme, Van Dooren (1994), ...

Who is Yu. V. Vorobyev?

All what we know can be found in Liesen, S, *Krylov subspace methods,* OUP, 2013, Section 3.7.

Russian Monographs and Texts on Advanced Mathematics and Physics

Volume I
A. I. KHINCHIN, *A Course of Mathematical Analysis*
Volume II
S. G. MIKHLIN, *Linear Integral Equations*
Volume III
P. P. KOROVKIN, *Linear Operators and Approximation Theory*
Volume IV
L. E. EL'SGOL'TS, *Differential Equations*
Volume V
L. A. LUSTERNIK and V. J. SOBOLEV, *Elements of Functional Analysis*
Volume VI
V. L. GINZBURG, *Propagation of Electromagnetic Waves in Plasma*
Volume VII
A. A. VLASOV, *Many-Particle Theory and its Application to Plasma*
Volume VIII
M. A. KRASNOSEL'SKI, *Convex Functions and Orlicz Spaces*
Volume IX
YU. V. VOROBYEV, *Method of Moments in Applied Mathematics*
Volume X
N. N. BOGOLIUBOV and Y. A. MITROPOLSKY, *Asymptotic Methods in the Theory of Non-Linear Oscillations*
Volume XI
S. D. VOLKOV, *Statistical Strength Theory*
Volume XII
Z. S. AGRANOVICH and V. A. MARCHENKO, *The Inverse Problem of Scattering Theory*
Volume XIII
V. F. NOZDREV, *Application of Ultrasonics in Molecular Physics*
Volume XIV
S. TARG, *Theoretical Mechanics: A Short Course*
Volume XV
A. S. KOMPANEYETS, *Theoretical Physics*
Volume XVI
P. STEPIN, *Strength of Materials*
Volume XVII
L. F. PONTRYAGIN, *Topological Groups*, Second Edition
Volume XVIII
S. B. PIKELNER, *Principles of Cosmic Electrodynamics*
Volume XIX
O. V. KOVALEV, *Irreducible Representation of Space Groups*

Additional Volumes in Preparation

Method of Moments in Applied Mathematics

by YU V. VOROBYEV

Translated from the Russian

by BERNARD SECKLER

GORDON AND BREACH SCIENCE PUBLISHERS
New York · London · Paris

The problem of moments in Hilbert space

Let $z_0, z_1, \ldots, z_n$ be n+1 linearly independent elements of Hilbert space V. Consider the subspace V_n generated by all possible linear combinations of $z_0, z_1, \ldots, z_{n-1}$ and construct a linear operator $\mathcal{B}_n$ defined on V_n such that

$$
\begin{aligned}
z_1 &= \mathcal{B}_n z_0 , \\
z_2 &= \mathcal{B}_n z_1 , \\
&\vdots \\
z_{n-1} &= \mathcal{B}_n z_{n-2} , \\
E_n z_n &= \mathcal{B}_n z_{n-1} ,
\end{aligned}
$$

where $E_n z_n$ is the projection of z_n onto V_n .

Approximation of bounded linear operators

Let $\mathcal{B}$ be a bounded linear operator in Hilbert space V . Choosing an element z_0 , we first form a sequence of elements $z_1, \ldots, z_n, \ldots$

$$z_0,\ z_1 = \mathcal{B} z_0,\ z_2 = \mathcal{B} z_1 = \mathcal{B}^2 z_0,\ \ldots\ , z_n = \mathcal{B} z_{n-1} = \mathcal{B}^n z_{n-1},\ \ldots$$

For the present $z_1, \ldots, z_n$ are assumed to be linearly independent. By solving the moment problem we determine a sequence of operators $\mathcal{B}_n$ defined on the sequence of nested subspaces V_n such that

$$\begin{aligned} z_1 &= \mathcal{B} z_0 = \mathcal{B}_n z_0, \\ z_2 &= \mathcal{B}^2 z_0 = (\mathcal{B}_n)^2 z_0, \\ &\vdots \\ z_{n-1} &= \mathcal{B}^{n-1} z_0 = (\mathcal{B}_n)^{n-1} z_0, \\ E_n z_n &= E_n \mathcal{B}^n z_0 = (\mathcal{B}_n)^n z_0. \end{aligned}$$

Approximation of bounded linear operators

Using the projection E_n onto V_n we can write for the operators constructed above (here we need the linearity of $\mathcal{B}$)

$$\mathcal{B}_n = E_n \, \mathcal{B} \, E_n \, .$$

The finite dimensional operators $\mathcal{B}_n$ can be used to obtain approximate solutions to various linear problems. The choice of the elements $z_0, \ldots, z_n, \ldots$ as above gives Krylov subspaces that are closely connected with the application (described, e.g. by partial differential equations).

Challenges: 1. convergence, 2. computational efficiency.

The most important classes of operators to study:

- completely continuous (compact),
- self-adjoint.

Inner product and Riesz map

Let V be a real (infinite dimensional) Hilbert space with the inner product

$$(\cdot,\cdot)_V : V \times V \to \mathbb{R}, \quad \text{the associated norm } \|\cdot\|_V ,$$

$V^\#$ be the dual space of bounded (continuous) linear functionals on V with the duality pairing

$$\langle \cdot,\cdot \rangle : V^\# \times V \to \mathbb{R} .$$

For each $f \in V^\#$ there exists a unique $\tau f \in V$ such that

$$\langle f, v \rangle = (\tau f, v)_V \quad \text{for all } v \in V .$$

In this way the inner product $(\cdot,\cdot)_V$ determines the Riesz map

$$\tau : V^\# \to V .$$

Operator formulation of the PDE BVP

Consider a PDE problem described in the form of the functional equation

$$\mathcal{A}x = b, \quad \mathcal{A} : V \to V^{\#}, \quad x \in V, \quad b \in V^{\#},$$

where the linear, bounded, and coercive operator $\mathcal{A}$ is self-adjoint with respect to the duality pairing $\langle \cdot, \cdot \rangle$.

Standard approach to solving boundary-value problems using the preconditioned conjugate gradient method (PCG) preconditions the algebraic problem,

$$\mathcal{A},\ \langle b, \cdot \rangle \ \to\ \mathbf{A}, \mathbf{b} \ \to\ \text{preconditioning} \ \to\ \text{PCG applied to } \mathbf{Ax} = \mathbf{b}\,,$$

i.e., discretization and preconditioning are often considered separately.

2 Krylov subspaces in Hilbert spaces

Using the Riesz map $\tau\mathcal{A} : V \to V$, one can form for $g \in V$ the Krylov sequence

$$g, \ \tau\mathcal{A}g\,, \ (\tau\mathcal{A})^2 g\,, \ \ldots \qquad \text{in } V$$

and define Krylov subspace methods in the Hilbert space operator setting (here CG) such that with $r_0 = b - \mathcal{A}x_0 \in V^{\#}$ the approximations x_n to the solution x , $n = 1, 2, \ldots$ belong to the Krylov subspaces in V

$$x_n \in x_0 + K_n(\tau\mathcal{A}, \tau r_0) \equiv$$

$$x_0 + \text{span}\{\tau r_0, \tau\mathcal{A}(\tau r_0), (\tau\mathcal{A})^2(\tau r_0), \ldots, (\tau\mathcal{A})^{n-1}(\tau r_0)\}\,.$$

Approximating the solution $x = (\tau\mathcal{A})^{-1}\tau b$ using Krylov subspaces is not the same as approximating the operator inverse $(\tau\mathcal{A})^{-1}$ by the operators $I, \tau\mathcal{A}, (\tau\mathcal{A})^2, \ldots$ Vorobyev moment problem depends on τb!

Vorobyev moment problem

Using the orthogonal projection E_n onto K_n with respect to the inner product $(\cdot,\cdot)_V$, consider the orthogonally restricted operator

$$\tau\mathcal{A}_n : K_n \to K_n \, , \quad \tau\mathcal{A}_n \equiv E_n \, (\tau\mathcal{A}) \, E_n \, ,$$

by formulating the following equalities

$$\begin{aligned}
\tau\mathcal{A}_n \, (\tau r_0) &= \tau\mathcal{A} \, (\tau r_0) \, , \\
(\tau\mathcal{A}_n)^2 \, \tau r_0 = \tau\mathcal{A}_n \, (\tau\mathcal{A} \, (\tau r_0)) &= (\tau\mathcal{A})^2 \, \tau r_0 \, , \\
&\vdots \\
(\tau\mathcal{A}_n)^{n-1} \, \tau r_0 = \tau\mathcal{A}_n \, ((\tau\mathcal{A})^{n-2} \, \tau r_0) &= (\tau\mathcal{A})^{n-1} \, \tau r_0 \, , \\
(\tau\mathcal{A}_n)^{n} \, \tau r_0 = \tau\mathcal{A}_n \, ((\tau\mathcal{A})^{n-1} \, \tau r_0) &= E_n \, (\tau\mathcal{A})^{n} \, \tau r_0 \, .
\end{aligned}$$

Lanczos process and Jacobi matrices

The n-dimensional approximation $\tau\mathcal{A}_n$ of $\tau\mathcal{A}$ matches the first $2n$ moments

$$((\tau\mathcal{A}_n)^\ell \tau r_0, \tau r_0)_V \; = \; ((\tau\mathcal{A})^\ell \tau r_0, \tau r_0)_V \, , \quad \ell = 0, 1, \ldots, 2n-1 \, .$$

Denote symbolically $Q_n = (q_1, \ldots, q_n)$ a matrix composed of the columns $q_1, \ldots, q_n$ forming an orthonormal basis of K_n determined by the Lanczos process

$$\tau\mathcal{A}\, Q_n = Q_n\, \mathbf{T}_n + \delta_{n+1}\, q_{n+1}\, \mathbf{e}_n^T$$

with $q_1 = \tau r_0 / \|\tau r_0\|_V$. We get $(\tau\mathcal{A}_n)^\ell = Q_n\, \mathbf{T}_n^\ell\, Q_n^*$, $\ell = 0, 1, \ldots$ and the matching moments condition

$$\mathbf{e}_1^*\, \mathbf{T}_n^\ell\, \mathbf{e}_1 = q_1^* (\tau\mathcal{A})^\ell q_1 , \quad \ell = 0, 1, \ldots, 2n-1 \, ,$$

Conjugate gradient method - first n steps

$$\mathbf{T}_n = \begin{pmatrix} \gamma_1 & \delta_2 & & & \\ \delta_2 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \delta_n \\ & & & \delta_n & \gamma_n \end{pmatrix}$$

is the Jacobi matrix of the orthogonalization coefficients and the CG method is formulated by

$$\mathbf{T}_n \mathbf{y}_n = \|\tau r_0\|_V \, \mathbf{e}_1 , \qquad x_n = x_0 + Q_n \mathbf{y}_n \, , \quad x_n \in V \, .$$

Spectral representation

Since $\tau\mathcal{A}$ is bounded and self-adjoint, its spectral representation is

$$\tau\mathcal{A} = \int_{\lambda_L}^{\lambda_U} \lambda \, d\mathcal{E}_\lambda \, .$$

The spectral function $\mathcal{E}_\lambda$ of $\tau\mathcal{A}$ represents a family of orthogonal projections which is

- non-decreasing, i.e., if $\mu > \nu$, then the subspace onto which $\mathcal{E}_\mu$ projects contains the subspace into which $\mathcal{E}_\nu$ projects;
- $\mathcal{E}_{\lambda_L} = 0$, $\mathcal{E}_{\lambda_U} = I$;
- $\mathcal{E}_\lambda$ is right continuous, i.e. $\lim_{\lambda' \rightarrow \lambda_+} \mathcal{E}_{\lambda'} = \mathcal{E}_\lambda$.

The values of λ where $\mathcal{E}_\lambda$ increases by jumps represent the eigenvalues of $\tau\mathcal{A}$, $\tau\mathcal{A}z = \lambda z, \quad z \in V$.

Representation of the moment problem

For the (finite) Jacobi matrix $\mathbf{T}_n$ we can analogously write

$$\mathbf{T}_n = \sum_{j=1}^{n} \theta_j^{(n)} \mathbf{s}_j^{(n)} (\mathbf{s}_j^{(n)})^* \,, \quad \lambda_L < \theta_1^{(n)} < \theta_2^{(n)} < \cdots < \theta_n^{(n)} < \lambda_U \,,$$

and the operator moment problem turns into the $2n$ equations for the $2n$ unknowns $\theta_j^{(n)}, \omega_j^{(n)}$

$$\sum_{j=1}^{n} \omega_j^{(n)} \, \{\theta_j^{(n)}\}^{\ell} \; = \; m_\ell \; \equiv \; \int_{\lambda_L}^{\lambda_U} \lambda^\ell \, d\omega(\lambda) \,, \qquad \ell = 0, 1, \ldots, 2n-1 \,,$$

where $d\omega(\lambda) = q_1^* d\mathcal{E}_\lambda q_1$ represents the Riemann-Stieltjes distribution function associated with $\tau\mathcal{A}$ and q_1 . The distribution function $\omega^{(n)}(\lambda)$ approximates $\omega(\lambda)$ in the sense of the nth Gauss-Christoffel quadrature; Gauss (1814), Jacobi (1826), Christoffel (1858).

Gauss-Christoffel quadrature

$$\tau\mathcal{A}, q_1 = \tau r_0/\|\tau r_0\|_V \quad \longleftrightarrow \quad \omega(\lambda), \ \int f(\lambda)\, d\omega(\lambda)$$

$$\uparrow \qquad\qquad\qquad \uparrow$$

$$\mathbf{T}_n, \mathbf{e}_1 \quad \longleftrightarrow \quad \omega^{(n)}(\lambda), \ \sum_{i=1}^{n} \omega_i^{(n)} f\left(\theta_i^{(n)}\right)$$

Using $f(\lambda) = \lambda^{-1}$ gives

$$\int_{\lambda_L}^{\lambda_U} \lambda^{-1}\, d\omega(\lambda) \ = \ \sum_{i=1}^{n} \omega_i^{(n)} \left(\theta_i^{(n)}\right)^{-1} + \frac{\|x - x_n\|_a^2}{\|\tau r_0\|_V^2}$$

Continued fraction representation, minimal partial realization etc.

References

- J. Málek and Z.S., Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs. SIAM Spolight Series, SIAM (2015)
- J. Liesen and Z.S., Krylov Subspace Methods, Principles and Analysis. Oxford University Press (2013)
- Z.S. and P. Tichý, On efficient numerical approximation of the bilinear form $c^* A^{-1} b$, SIAM J. Sci. Comput., 33 (2011), pp. 565-587

Non self-adjoint compact operators?

Gauss quadrature in complex plane?

Vorobyev moment problem can be based on generalization of the Lanczos process to non self-adjoint operators with starting elements z_0, w_0 . Then, however, the tridiagonal matrix of the recurrence coefficients for the properly normalized formal orthogonal polynomials (assuming, for the present, their existence) is complex symmetric but not (in general) Hermitian.

Generalization of the n-weight Gauss quadrature representation of the Vorobyev moment problem that eliminates restrictive assumptions on diagonalizability can be based on quasi-definite functionals; see the poster of Stefano Pozza and

S. Pozza, M. Pranic and Z.S., Gauss quadrature for quasi-definite linear functionals, submitted (2015).

Conclusions

- Vorobyev work was built on the deep knowledge of the previous results.
- It is amazingly thorough and as to the coverage and references.
- Published in 1958 (1965), it was much ahead of time. It stimulates new developments for the future.

Volker, Many Thanks and Congratulations!

Whatever we try, does not work