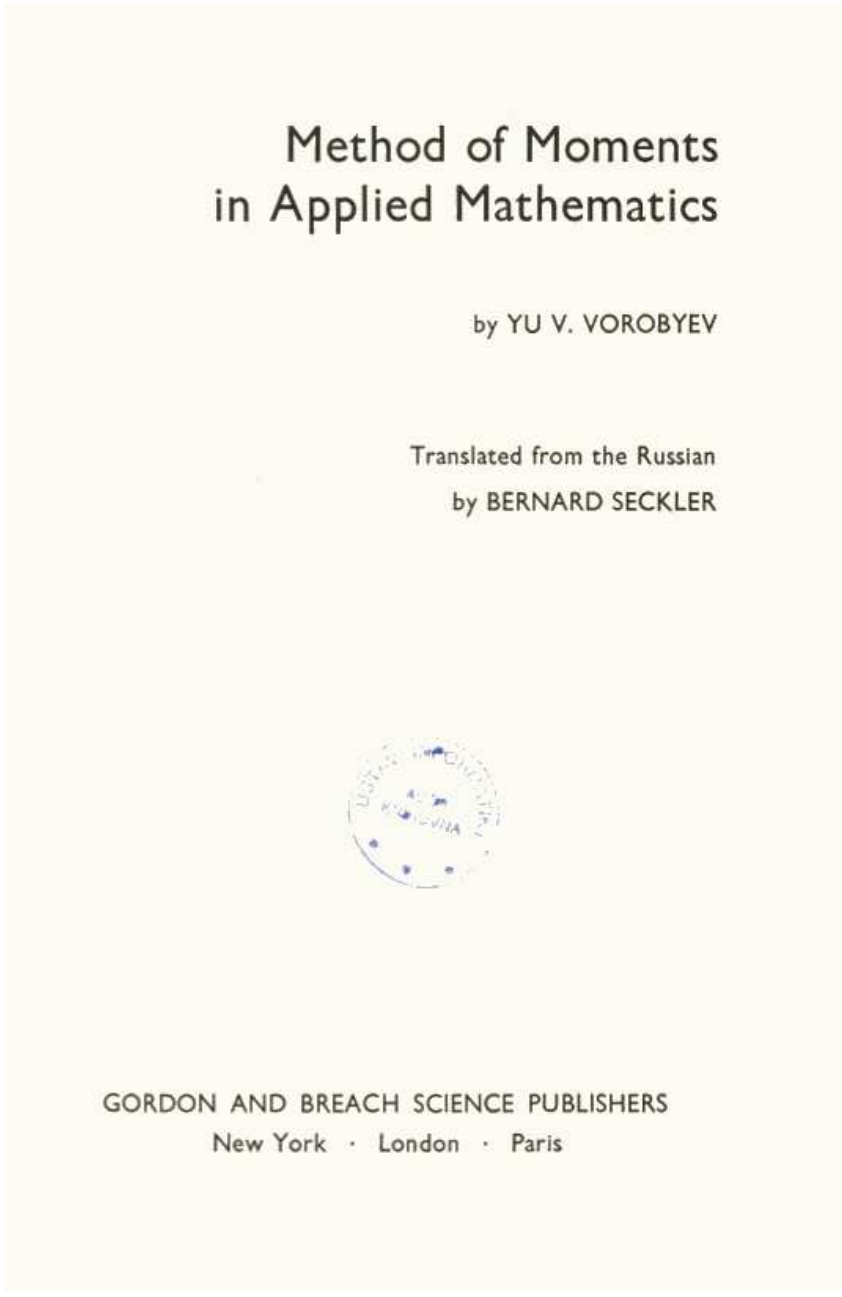
Nečas Center for Mathematical Modeling

MOdelling REvisited + MOdel REduction
ERC-CZ project LL1202 - MORE

MINISTRY OF EDUCATION, YOUTH AND SPORTS

Vorobyev's Method of Moments in Applied Mathematics - Part I (Self-adjoint Operators)

MORE seminar, May 4, 2015



Yu. V. Vorobyev: Method of Moments in Applied Mathematics; Translated from the Russian original (1958) by Bernard Seckler; Gordon and Breach Science Publishers, New York, 1965.

Outline

Part I introduction, self-adjoint operators

Part II generalizations, classification

Part III moment problem for completely continuous (compact) operators

Part IV partial realization, examples, mathematical physics, miscellaneous, . . .

Other references

J. Liesen, Z. Strakoš:
Krylov Subspace Methods, Principles and Analysis, Chapter 3.
Oxford University Press, (2013).

J. Málek and Z. Strakoš:
Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs.
SIAM Spotlight Series, SIAM, Philadelphia, (2015).

Outline of Part I

1. Problem of moments in Hilbert space
2. Method of moments for self-adjoint operators

I. Preliminaries

- We consider a (complete) Hilbert space H with the inner product $(\cdot, \cdot)$ and the induced norm $\|\cdot\|$;
- operators act on H, i.e.,

$$A, A_n : H \longrightarrow H;$$

- (strong) convergence of operators in H, $A_n \to A$ means

$$\lim_{n\to\infty} \|A_n x - Ax\| = 0 \qquad \text{for all } x \in H;$$

- for a subspace $H_n \subset H$, E_n denotes the orthogonal projection onto H_n.

I. Problem of moments in Hilbert space

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

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

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

I. Eigenvalues of operator A_n

Since $E_n z_n$ is an element of H_n, we can find numbers $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ such that

$$E_n z_n = -\alpha_0 z_0 - \alpha_1 z_1 - \cdots - \alpha_{n-1} z_{n-1} .$$

Using the definition of A_n,

$$P_n(A_n) z_0 := \left(A_n^n + \alpha_{n-1} A_n^{n-1} + \cdots + \alpha_0 I\right) z_0 = 0 .$$

It can be shown that the eigenvalues of A_n are the roots of the (monic) polynomial

$$P_n(\lambda) = \lambda^n + \alpha_{n-1} \lambda^{n-1} + \cdots + \alpha_0 .$$

I. Approximation of bounded linear operators

Let A be a bounded linear operator in Hilbert space H.
Choosing an element z_0, we first form a sequence of elements $z_1, \ldots, z_n, \ldots$

$$z_0,\ z_1 = Az_0,\ z_2 = Az_1 = A^2 z_0,\ \ldots\ , z_n = Az_{n-1} = A^n z_0,\ \ldots$$

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 A_n defined on the sequence of nested subspaces H_n such that

$$\begin{aligned} z_1 &= Az_0 &&= A_n z_0, \\ z_2 &= A^2 z_0 &&= (A_n)^2 z_0, \\ &\vdots \\ z_{n-1} &= A^{n-1} z_0 &&= (A_n)^{n-1} z_0, \\ E_n z_n &= E_n A^n z_0 &&= (A_n)^n z_0. \end{aligned}$$

I. Approximation of bounded linear operators

Using the projection E_n onto H_n we can write for the operators constructed above (here we need the linearity of A)

$$A_n = E_n A E_n .$$

The orthogonality of E_n gives

$$\|A_n\| \leq \|A\| .$$

Matching moments property

There holds

$$(A_n^\ell z_0, z_0) = (A^\ell z_0, z_0) , \qquad \ell = 0, 1, \ldots, n-1, n .$$

I. Convergence $A_n \to A$

Given z_0, consider the subspace H_z as the closure, w.r.t. the norm $\|\cdot\|$, of the linear manifold

$$\{x \in H \mid x = Q(A)z_0, \quad Q(\lambda) \text{ is an arbitrary polynomial}\}$$

Then A can be considered to be an operator on H_z.

Theorem II, p. 21

If A is a bounded linear operator and A_n a sequence of solutions of the problem of moments, then the sequence A_n converges strongly to A in the subspace H_z.

I. Comments

Linear (in)dependence of $z_0, z_1, \ldots, z_n, \ldots$

If z_n is linearly dependent on $z_0, z_1, \ldots, z_{n-1}$, then $H_z = H_n$ and A_n coincides with A on H_z.

Krylov subspaces

The choice of the elements $z_0, z_1, \ldots, z_n, \ldots$ as above gives Krylov subspaces that are closely connected with the application (described, e.g. by partial differential equations).

Method of moments falls within the general framework of the Galerkin's method

The approximate operator

$$A_n = E_n A E_n .$$

Section 2

Method of moments for self-adjoint operators

II. Self-adjoint operators

Let A be a bounded self-adjoint operator, i.e.

$$(Ax, y) = (x, Ay) \qquad \text{for all } x, y \in H\,.$$

Consequently (Ax, x) is a real number and the eigenvalues of A are also real.
Bounds of a self-adjoint operator

$$m\,(x, x) \leq (Ax, x) \leq M\,(x, x) \qquad \text{for all } x \in H\,.$$

II. Matching moments property for self-adjoint operators

The problem of moments for a self-adjoint operator A and an element z_0 yields the sequence of (nested) subspaces H_n and the self-adjoint operators $A_n = E_n A E_n$ that matches the first $2n$ moments of the operator A, i.e.

$$(A_n^\ell z_0, z_0) = (A^\ell z_0, z_0)\,, \qquad \ell = 0, 1, \ldots, n, \ldots, 2n-1\,.$$

II. Outline

- spectral representation of self-adjoint operators
- Stieltjes' moment problem
- Gauss-Christoffel quadrature
- determining the spectrum of A_n
- orthogonal polynomials
- Lanczos algorithm, Conjugate Gradient Method

II. Finite-dimensional self-adjoint operator

Consider a self-adjoint operator A_n in some n-dimensional space H_n and assume for the moment that it has distinct eigenvalues

$$\lambda_1 < \lambda_2 < \ldots < \lambda_n \,.$$

The corresponding eigenelements (eigenvectors) u_k are mutually orthogonal, i.e.

$$(u_i, u_k) = 0 \qquad i \neq k \,.$$

We consider eigenelements u_k normalized with $\|u_k\| = 1$, so that $\{u_k\}$ form an orthonormal basis of H_n and

$$x = (x, u_1)\, u_1 + (x, u_2)\, u_2 + \ldots + (x, u_n)\, u_n \,, \qquad x \in H_n$$

II. Spectral function, n-dimensional case

The spectral function $\mathcal{E}_\lambda^{(n)}$ of A_n is the family of projections

$$\begin{aligned}
&\mathcal{E}_\lambda^{(n)} x = 0 && && \lambda &< \lambda_1 \\
&\mathcal{E}_\lambda^{(n)} x = (x, u_1)\, u_1 && && \lambda_1 \le \lambda &< \lambda_2 \\
&\cdots \\
&\mathcal{E}_\lambda^{(n)} x = (x, u_1)\, u_1 + \ldots + (x, u_{n-1})\, u_{n-1} && && \lambda_{n-1} \le \lambda &< \lambda_n \\
&\mathcal{E}_\lambda^{(n)} x = x && && \lambda_n \le \lambda &
\end{aligned}$$

Denoting the value of the jump by $\Delta_k \mathcal{E}_\lambda^{(n)} x \equiv (x, u_k)\, u_k$, $k = 1, \ldots, n$, we can write

$$\begin{aligned}
A_n x &= \lambda_1 (x, u_1)\, u_1 + \lambda_2 (x, u_2)\, u_2 + \ldots + \lambda_n (x, u_n)\, u_n \\
&= \sum_{k=1}^{n} \lambda_k \Delta_k \mathcal{E}_\lambda^{(n)} x
\end{aligned}$$

II. Spectral function, general case

The spectral representation for any bounded self-adjoint operator A

$$Ax = \int_m^M \lambda d\mathcal{E}_\lambda x\,,$$

where the spectral function $\mathcal{E}_\lambda$ of 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;
- $\mathcal{E}_m = 0,\ \mathcal{E}_M = I$;
- $\mathcal{E}_\lambda$ is right continuous, i.e. $\lim_{\lambda' \to \lambda_+} \mathcal{E}_{\lambda'} = \mathcal{E}_\lambda$.

The values of λ where $\mathcal{E}_\lambda$ increases by jumps represent the eigenvalues of A.

II. Convergence of spectral functions

Let A_n be the solution of problem of moments for A, z_0 and $\mathcal{E}_\lambda^{(n)}$ be the spectral function of A_n.

As a result of (strong) convergence $A_n \to A$ in H_z, we have

Theorem IX, p. 61

If A is a bounded self-adjoint operator and A_n a sequence of solutions of the problem of moments, then the sequence of spectral functions $\mathcal{E}_\lambda^{(n)}$ of the operators A_n converges strongly to the spectral function of A

$$\mathcal{E}_\lambda^{(n)} \to \mathcal{E}_\lambda$$

in the subspace H_z for all λ not belonging to the discrete spectrum of A.

II. Distribution function

The spectral function $\mathcal{E}_\lambda$ and z_0 determine the real function $(\mathcal{E}_\lambda z_0, z_0)$ that satisfies

- $(\mathcal{E}_\lambda z_0, z_0) = 0 \qquad$ for $\lambda < m$,
- $(\mathcal{E}_\lambda z_0, z_0) = \|z_0\|^2 \quad$ for $M \leq \lambda$,
- $(\mathcal{E}_\lambda z_0, z_0)$ is non-decreasing .

The function $(\mathcal{E}_\lambda z_0, z_0)$ can be considered as a distribution function for a Riemann-Stieltjes integral.

Hereafter, we will for simplicity of notation consider normalized distribution functions, i.e., we consider

$$\|z_0\| = 1\,.$$

II. Simplified Stieltjes' moment problem

Given a distribution function $\omega(\lambda)$ on $[m, M]$, find a distribution function $\omega^{(n)}(\lambda)$, with n distinct points of increase in $[m, M]$, such that the first $2n$ moments are matched, i.e. such that

$$\int_m^M \lambda^\ell d\omega(\lambda) = \int_m^M \lambda^\ell d\omega^{(n)}(\lambda), \qquad \ell = 0, 1, \ldots, 2n-1.$$

Equivalent formulation: find an n-node (mechanical) quadrature with positive weights for the given Riemann-Stieltjes integral with distribution function $\omega(\lambda)$, that has algebraic degree of exactness $2n-1$, i.e. that is exact for polynomials of degree at most $2n-1$.

[Liesen & Strakoš, §3.1]

II. Solution of Stieltjes' moment problem

Gauss-Christoffel quadrature

The n-node (interpolatory) quadrature that has algebraic degree of exactness $2n-1$.

The (distinct) quadrature nodes are given as the roots of the n-th monic orthogonal polynomial w.r.t. the Riemann-Stieltjes integral

$$\int_m^M f(\lambda) d\omega(\lambda) .$$

[Liesen & Strakoš, §3.2]

Properties of orthogonal polynomials and the relation to continued fractions: [Liesen & Strakoš, §3.3]

II. Orthogonal basis of H_n

Using Lanczos method of successive orthogonalization we construct an orthogonal basis of the subspace $H_n = \mathrm{span}\{z_0, z_1, \ldots, z_{n-1}\}$:

$$\begin{aligned} p_0 &= z_0\,, \\ p_1 &= Ap_0 - a_0 p_0\,, \qquad a_0 \text{ is such that } (p_0, p_1) = 0 \\ p_2 &= Ap_1 - a_1 p_1 - b_0 p_0\,, \\ &\ldots \\ p_{k+1} &= (A - a_k I)\, p_k - b_{k-1} p_{k-1}\,. \end{aligned}$$

Operator A is self-adjoint $\Rightarrow$ three-term recurrence.

We can write

$$p_k = P_k(A)\, z_0\,,$$

where $P_k(\lambda)$ is a monic polynomial of degree k.

II. Orthogonal basis of H_n

Since A_n is the solution of problem of moments, we have

$$P_n(A_n)\, z_0 = E_n P_n(A)\, z_0 = E_n p_n\,.$$

From the orthogonality of p_n to each of the elements $p_0, \ldots, p_{n-1}$,

$$E_n p_n = 0\,.$$

The spectrum of A_n is equal to the roots of $P_n(\lambda)$.

II. Orthogonal polynomials

Using the spectral function $\mathcal{E}_\lambda$ of the operator A,

$$p_k = P_k(A)\, z_0 = \int_m^M P_k(\lambda) d\mathcal{E}_\lambda z_0 \,.$$

The orthogonality of the elements p_k gives (for $i \neq k$)

$$0 = (p_i, p_k) = \int_m^M P_i(\lambda) P_k(\lambda)\; d(\mathcal{E}_\lambda z_0, z_0)\,,$$

Therefore the (monic) polynomials $P_k(\lambda)$ are orthogonal over interval $[m, M]$ w.r.t. the distribution function $(\mathcal{E}_\lambda z_0, z_0)$.

The roots of the polynomial $P_n(\lambda)$ are the nodes of the Gauss-Christoffel quadrature for solving the Stieltjes moment problem with distribution function $(\mathcal{E}_\lambda z_0, z_0)$.

II. Lanczos algorithm

Denote symbolically $Q_n = (q_1, \ldots, q_n)$, where the columns $q_1, \ldots, q_n$ form an orthonormal basis of H_n determined by the Lanczos process

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

with $q_1 = z_0$ (we assume $\|z_0\| = 1$). We get

$$A_n^\ell = Q_n \mathbf{T}_n^\ell Q_n^* , \qquad \ell = 0, 1, \ldots$$

and the moment matching property reads

$$\mathbf{e}_1^T \mathbf{T}_n^\ell \mathbf{e}_1 = q_1^* A^\ell q_1 = (A^\ell z_0, z_0) , \qquad \ell = 0, 1, \ldots, 2n-1 .$$

[Liesen & Strakoš, §3.5, §3.7]

II. Lanczos algorithm, Jacobi matrix

The matrix $\mathbf{T}_n$ of the orthogonalization coefficients has the form

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

where $\delta_j > 0$. Eigenvalues of $\mathbf{T}_n$ are equal to the eigenvalues of A_n.

Properties of Jacobi matrices: [Liesen & Strakoš, §3.4]

II. Conjugate Gradient Method

Let A be self-adjoint positive definite and the right-hand side f an element of H_z. We are solving

$$Ax = f\,, \quad \text{i.e.} \quad x = A^{-1}f = \int_m^M \lambda^{-1}\, d\mathcal{E}_\lambda f\,.$$

Approximation using the spectral function $\mathcal{E}_\lambda^{(n)}$

$$x_n = \int_m^M \lambda^{-1}\, d\mathcal{E}_\lambda^{(n)} f = \int_m^M \lambda^{-1}\, d\mathcal{E}_\lambda^{(n)} f_n = A_n^{-1} f_n\,,$$

where f_n denotes the projection of f on H_n. Since all the eigenvalues of A_n lie in the interval $[m, M]$ with $m > 0$, the approximation x_n is well-defined for any n.

II. Conjugate Gradient Method

The approximation $x_n = A_n^{-1} f_n$ can be computed by the Conjugate Gradient Method (CG) by [Hestenes & Stiefel].

The matrix representation of CG using the Jacobi matrix $\mathbf{T}_n$ and the orthonormal basis Q_n of H_n is

$$\mathbf{T}_n \mathbf{y}_n = \|f\| \mathbf{e}_1 , \quad x_n = Q_n \mathbf{y}_n \in H_n .$$

Provided that $f \in H_z$, convergence $A_n \to A$ in H_z gives $x_n \to x$.

II. Model reductions matching $2n$ moments

Given a self-adjoint operator A and $z_0 \in H$:

Vorobyev's problem of moments

$A : H_z \longrightarrow H_z\,, \quad A_n : H_n \longrightarrow H_n\,, \quad \dim(H_n) = n\,.$

Gauss-Christoffel quadrature

from the distribution function $\omega(\lambda)$ to the distribution function $\omega^{(n)}(\lambda)$ with n points of increase.

Lanczos algorithm

from A, z_0 to $\mathbf{T}_n, \mathbf{e}_1\,, \quad \mathbf{T}_n \in \mathbb{R}^{n\times n}\,.$

Conjugate Gradient Method

from $Ax = f$ to $\mathbf{T}_n\,\mathbf{y}_n = \|f\|\,\mathbf{e}_1$ (for positive definite self-adjoint operator A).

Questions for the next talks

- Can we match more than $n+1$ moments if A is not self-adjoint?
- Can we consider an oblique (i.e. not-orthogonal) projection E_n?
- Do eigenvalues of A_n converge to eigenvalues of A?
- Can we generalize Gauss-Christoffel quadrature for non positive definite linear functionals?
- ...

Thank you for your attention!