Moment Problems: old and new

Mihai Putinar

UCSB and NTU

November 26, 2013

Classical problem

Find a positive Borel measure μ on $\mathbf{R}$, with prescribed moments

$$s_k = \int x^k d\mu(x), \quad 0 \leq k < n.$$

With $n \leq \infty$.

Abstract version

X structured vector space, $f_1, \ldots, f_n$ linearly independent elements

Find $L \in X^*$ with prescribed values

$$L(f_k) = s_k, \; 0 \leq k < n,$$

and additional properties such as

minimal norm, L is positive on a specified cone, extremal among all solutions,...

Statisticians first

M. G. Kendall: *The Advanced Theory of Statistics*, London, 1945.

In prectice numerical moments of order higher than fourth are rarely required, being so sensitive to sampling fluctuations that values computed from moderate numbers of observations are subject to a large margin of error.

Cumulants

$$\sum_{k=0}^{\infty} \kappa_k \frac{z^n}{n!} = \log \sum_{\ell=0}^{\infty} s_\ell \frac{z^\ell}{\ell!}.$$

are invariants under the translation group.

Extremely simple for the classical probability distributions, additive for sums of independent random variables. The coefficients (Bell polynomials) have a high combinatorial significance.

Hald, Anders. *A History of Mathematical Statistics from 1750 to 1930*. New York: Wiley., 1998.

Best approximation

Chebyshev's problem (around 1860):

Given a C^1-function $F(x_1, ..., x_d; p_1, ...p_n)$ on a domain $\Omega \subset \mathbf{R}^d$, depending on parameters $p_1, ..., p_n$, find

$$\min_{p_1,\dots,p_n} \max_{x_1,\dots,x_d} |F(x_1, ..., x_d; p_1, ...p_n)|.$$

inspired by Poncelet's extremal problem for $\frac{p_1 x + p_2}{\sqrt{1+x^2}} - 1$.
Both Chebyshev and A. A. Markov have extensively worked on this subject, becoming masters of continued fractions, as a necessary technical tool.

Chebyshev polynomials

Consider the polynomial of minimal variation from zero, on an interval $[-h, h]$, of the form

$$f(x) = x^n + p_1 x^{n-1} + \ldots + p_n,$$

with $p_1, ..., p_n$ parameters to be determined.

Necessarily

$$f(x)^2 - L^2 = (x^2 - h^2)\frac{f'(x)^2}{n^2}$$

where L is the optimal value.

The division algorithm

$$f(x) - \sqrt{x^2 - h^2}\frac{f'(x)}{n} = \frac{L^2}{f(x) + \sqrt{x^2 - h^2}\frac{f'(x)}{n}}$$

hence

$$\frac{1}{\sqrt{x^2 - h^2}} - \frac{f'(x)}{nf(x)} = \frac{L^2}{\sqrt{x^2 - h^2}f(x)[f(x) + \sqrt{x^2 - h^2}f'(x)/n}]$$

that is

$$\frac{1}{\sqrt{x^2 - h^2}} - \frac{f'(x)}{nf(x)} = O(\frac{1}{x^{2n+1}}).$$

There is only one choice: $\frac{f'(x)}{nf(x)}$ is the n-th convergent of the coninued fraction expansion of $\frac{f'(x)}{nf(x)}$:

$$\frac{f'(x)}{nf(x)} = \cfrac{1}{x - \cfrac{h^2}{2x - \cfrac{h^2}{\ddots - \frac{h^2}{2x}}}}.$$

That is

$$f(x) = \frac{(x+\sqrt{x^2-h^2})^n + (x-\sqrt{x^2-h^2})^n}{2^n}.$$

Polynomial extrapolation

Chebyshev again

The real points $x_0, ..., x_n, X$ are given. Knowing the approximative values of a polynomial F of degree n at $n+1$ points x_k, find the errors in the values $F(x_k)$ which have minimal influence of the value $F(X)$.

As a matter of fact an extremal problem in square mean.

$$F(x) = \mu_0 q_0(x) f(x_0) + \ldots + \mu_m q_n(x) f(x_n)$$

with $q_k(x)$ unknown polynomials, such that

$$\mu_0 q_o(x)^2 + \ldots + \mu_n q_n(x)^2$$

is minimal.

Continued fractions

The solution exposed by Chebyshev reduces to a continued fraction argument. Given

$$\sum_{k=0}^{n} \frac{\mu_k}{x_k - z} = -\frac{s_0}{z} - \frac{s_1}{z^2} - \dots$$

find a polynomial $P(x)$ of degree m such that

$$P(x)(\frac{s_0}{z} + \frac{s_1}{z^2} + \dots)$$

begins with a term of highest order.

Bounds for integrals

Chebyshev 1833

Find bounds for $\int_0^a f(x)dx$*, from the known values*

$$s_0 = \int_0^A f(x)dx, s_1 = \int_0^A xf(x)dx, \dots, s_n = \int_0^A x^n f(x)dx$$

where $A > a$ *and* $f(x) \geq 0$.

Major observation: if $q_m(x)/p_m(x)$ is the continued fraction convergent of the expansion of

$$\int_0^A \frac{f(x)dx}{x-z} = -\frac{s_0}{z} - \frac{s_1}{z^2} - \dots$$

and $\lambda_1 < \lambda_2 < \dots$ are the (simple) zeros of $p_n(x)$, then

$$\sum_{k=\ell+1}^{n-1} \frac{q_m(\lambda_k)}{p'_n(\lambda_k)} < \int_{\lambda_\ell}^{\lambda_n} f(x)dx < \sum_{k=\ell}^{n} \frac{q_m(\lambda_k)}{p'_n(\lambda_k)}$$

Limit Theorems in Probability Theory

Chebyshev 1887, Markov 1899:

Assume that the functions $f_n \geq 0$ satisfy

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} x^k f_n(x)dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^k e^{-x^2/2} dx, \quad k \geq 0.$$

Then, for every $\alpha < \beta$

$$\lim_{n\to\infty} \int_{\alpha}^{\beta} f_n(x)dx = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-x^2/2} dx.$$

Stieltjes memoir, 1894-95

The sequence $(s_k)_{k=0}^{\infty}$ represents the moments of a positive measure supported on $[0,\infty)$ if and only if the continued fraction expansion

$$\frac{s_0}{z}+\frac{s_1}{z^2}+\ldots=\cfrac{1}{c_0 z-\cfrac{1}{c_1-\cfrac{1}{c_2 z-\cfrac{1}{c_3-\ldots}}}}$$

contains only non-negative terms $c_k \geq 0$.

Determinateness

If $\sum_{k=0}^{\infty} c_k = \infty$, then the convergents $\frac{P_m(z)}{Q_m(z)}$ converge in the upper-half plane to $\int_0^\infty \frac{d\sigma(x)}{z-x}$, and the measure $\sigma \geq 0$ is unique.

If $\sum_{k=0}^{\infty} c_k < \infty$, then the convergents satisfy in the upper-half plane:

$$P_{2k}(z) \to p(z), \quad Q_{2k}(z) \to q(z),$$

$$P_{2k+1}(z) \to p_1(z), \quad Q_{2k+1}(z) \to q_1(z),$$

where p, q, p_1, q_1 are entire functions of genus zero, satisfying

$$q(z)p_1(z) - q_1(z)p(z) = 1.$$

In that case the problem has infinitely many solutions, with distinct Cauchy transforms. Among these:

$$\frac{p(z)}{q(z)} = \sum_{k=1}^{\infty} \frac{\mu_k}{z - \lambda_k}, \quad \frac{p_1(z)}{q_1(z)} = \sum_{k=1}^{\infty} \frac{\mu'_k}{z - \lambda'_k}.$$

More on continued fractions

Hamburger 1919-1921

solves the power moment problem on the real axis, remarking (following Stieltjes) the positivity of the Hankel determinants

$$\det \begin{pmatrix} s_0 & s_1 & \dots & s_n \\ s_1 & s_2 & \dots & s_{n+1} \\ \vdots & & \ddots & \vdots \\ s_n & s_{n+1} & \dots & s_{2n} \end{pmatrix} \geq 0$$

as necessary and sufficient conditions for solvability.

Recent references

Iosifescu, Kraaikamp: *Metrical Theory of Continued Fractions*, Springer 2002
Cuyt, Brevik Petersen, Verdonk, Waadeland, Jones: *Handbook of Continued Fractions*, Springer 2008
Khrushchev: *Orthogonal Polynomials and Continued Fractions*, Cambridge U. Press, 2008.

Bounded analytic interpolation

The moment problem as a tangential (i.e. boundary) interpolation problem for analytic functions of the form

$$f(z) = \int_{\mathbf{R}} \frac{d\sigma(x)}{x - z}, \quad f(z) \approx -\frac{s_0}{z} - \frac{s_1}{z^2} - \dots .$$

In this case $\Im z > 0 \;\Rightarrow\; \Im f(z) > 0$ and $\sup_{t \geq 1} |tf(it)| < \infty$.

The interpolation problem: *Find and describe* $f : \Omega \to \omega$ *analytic such that* $f(\lambda) = c_\lambda$ *for* $\lambda \in \Delta$*, discrete subset of* Ω.

has a glorious past and present:

Google citations, Nov. 2013

Caratheodory-Fejér interpolation	10,200
Nevanlinna-Pick interpolation	29,500
Schur algorithm	407,000
H^∞-control	7,160,000

Nevanlinna parametrization

Obtained around 1922 by Hamburger and Nevanlinna:
free parametrization of all solutions to the power moment problem, in the indeterminate case

$$f(z) = \int_{\mathbf{R}} \frac{d\sigma(x)}{x-z} = -\frac{p(z)\phi(z) - p_1(z)}{q(z)\phi(z) - q_1(z)},$$

where $\phi(z)$ is any analytic function satisfying $\Im\phi(z) \geq 0$ whenever $\Im z > 0$.

Berg: *Indeterminate moment problems and the theory of entire functions*, J. Comput. Appl. Math. 65: 13(1995), 27 - 55.

Fourier transform

Known as the method of characteristic function in Probability

$$\hat{\sigma}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ix\xi} d\sigma(x)$$

uniquely determines σ (by known inversion formulae).

Bochner: $\sigma \geq 0$ if and only if $\hat{\sigma}(\xi_1 - \xi_2)$ is a positive semi-definite kernel.

If the moments exist $\hat{\sigma} \in C^{\infty}$ and

$$s_k = \sqrt{2\pi}(-i)^k \hat{\sigma}^{(k)}(0), \quad k \geq 0.$$

Carleman's approach

Quasi-analytic Functions (1926):

$u \in C^\infty[0,1]$ is fully determined by $(u^{(k)})_{k=0}^\infty$ if and only if

$$\sum_{k=0}^{\infty} \frac{1}{L_k} = \infty$$

where

$$|u^{(k)}(x)| \leq K^{k+1} M_k, \quad 0 \leq x \leq 1, \quad k \geq 0,$$

and

$$L_k = \inf_{j \geq k} M_j^{1/j}.$$

Determinateness

via quasi-analyticity

$$\sum_{k=0}^{\infty} \frac{1}{s_{2k}^{1/(2k)}} = \infty$$

implies uniqueness of the positive measure on $\mathbf{R}$ with moments (s_k).

$$\sum_{k=0}^{\infty} \frac{1}{s_{k}^{1/(2k)}} = \infty$$

implies uniqueness of the positive measure on $[0, \infty)$ with moments (s_k).

Reconstruction

Carleman's reconstruction formula of a quasi-analytic function $u \in C_M^\infty$:

$$u(x) = \lim_{n\to\infty} \sum_{k=0}^{n-1} \omega_{n,k} u^{(k)}(0) x^k,$$

where the coefficients $\omega_{n,k}$ depend only on the class C_M^∞.

Laplace transform

$$g(t) = \mathcal{L}(\sigma)(t) = \int_0^\infty e^{-xt} d\sigma(x),$$

also satisfies

$$s_k = (-1)^k \mathcal{L}(\sigma)^{(k)}(0), \quad k \geq 0$$

whenever the moments exist.
Analytic extension

$$\mathcal{L}(\sigma)(z) = \int_0^\infty e^{-xz} d\sigma(x),$$

for $Rez > 0$ with known inversion formulae...

S. Bernstein (1929): characterization of transforms of positive measures

$$g \in C^\infty[0, \infty), \quad (-1)^k g^{(k)} \geq 0, \quad k \geq 0.$$

Abolutely continuous measures

$$d\sigma(x) = \tau(x)dx, \quad \tau \in L^p([0,\infty))$$

has Laplace transform in a Hardy class of the right half-plane, with an arsenal of function theory of a complex variable available.

Widder (1934) inversion: $g(t) = \mathcal{L}(\tau dx)(t)$ reads

$$\lim_{k\to\infty} (-1)^k g^{(k)}(\frac{k}{t})(\frac{k}{t})^{k+1} = \tau(t), \quad t > 0.$$

Summation of divergent series

Improving the convergence of series by average methods, induced by a regular summation scheme

$$t_n = \sum_{k=0}^{n} a_{nk} u_k,$$

so that

$$\lim u_n = u \quad \Rightarrow \lim t_n = u.$$

Hausdorff (1921) A sequence $(s_n)_{n=0}^{\infty}$ represents the moments of a probability measure on $[0,1]$ if and only if the matrix $D\mathrm{diag}(s_0, s_1, \ldots)D$ induces a regular summation scheme, where

$$D = ((-1)^n \binom{k}{n}).$$

Extension of linear functionals

A positive Borel measure σ on $\mathbf{R}$, admitting all moments, is given by a positive linear functional $L : C_p(\mathbf{R}) \longrightarrow \mathbf{R}$, where C_p is the space of continuous functions of polynomial growth:

$$\int f d\sigma = L(f), \quad f \in C_p(\mathbf{R}).$$

M. Riesz (1922-23) Solving the moment problem by **linear extension** of positive linear functionals

$$L_d : \mathbf{R}_d[x] \longrightarrow \mathbf{R}$$

where

$$\mathbf{R}_d[x] = \{p \in \mathbf{R}[x];\ \deg p \leq d\}.$$

Convexity

M. Riesz method prompted to study the convex cones

$$\{p \in \mathbf{R}_d[x], \quad p(x) \geq 0, \quad x \in K\}$$

and

$$\{\sigma \in M_+(\mathbf{R}), \int x^k d\sigma(x) = s_k, \quad 0 \leq k \leq n\}.$$

Carathéodory (1911) carried earlier a similar analysis on the unit circle $\mathbf{T}$.

Classical Harmonic Analysis

On the unit circle $\mathbf{T}$, the moments of a measure are exactly its Fourier coefficients:

$$\hat{\sigma}(n) = \int_{-\pi}^{\pi} e^{-inx} d\sigma(x), \ \ n \in \mathbf{Z}.$$

Toeplitz and Carathéodory (1911): $\sigma \geq 0$ if and only if the kernel $\hat{\sigma}(n - m)$ is positive semi-definite.

F. Riesz and Herglotz (1911) establish a fundamental correspondence between non-negative harmonic functions in the unit disk and (moments of) positive measures on the circle.

Hilbert space

The moments of a positive measure σ are structured in a non-negative definite Gramm matrix

$$s_{k+m} = \int x^k x^m d\sigma(x) = \langle x^k, x^m \rangle_{2,\sigma}.$$

Hence an associated system of orthogonal polynomials

$$P_k(x) = \gamma_k z^k + O(z^{k-1}), \quad \langle P_k, P_m \rangle_{2,\sigma} = \delta_{km}.$$

The leading coefficient solves an extremal problem, à la Chebyshev-Markov:

$$\gamma_k^{-1} = \inf_{\deg q \leq k-1} \|z^k - q(z)\|_{2,\sigma}.$$

Christoffel functions

Given orthonormal polynomials P_n, the Christoffel function

$$C_n(z,z) = \sum_{k=0}^{n} |P_k(z)|^2$$

and its polarization

$$C_n(z,w) = \sum_{k=0}^{n} P_k(z)\overline{P_k(w)}$$

are essential in the study of the asymptotics of the polynomials P_n.

M. Riesz (1923): the solution to the extremal problem

$$\rho_n(\lambda) = \min\{\|q\|_{2,\sigma}^2, \quad \deg q \le n, \quad q(\lambda) = 1\}$$

is attained by the polynomial

$$\rho_n(z) = \frac{C(z,\lambda)}{C(\lambda,\lambda)}.$$

Weyl circle

Another parametrization of all solutions, obtained via the values of the Cauchy transforms

$$C(\sigma_n)(\lambda) = \int \frac{d\sigma_n(x)}{x-\lambda}, \quad \int x^k d\sigma_n(x) = s_k, \quad k \leq n.$$

They form a disk of radius ρ_n.

Parallel theory to the continuous case (Sturm-Liouville problem) studied by H. Weyl a decade before Riesz.

Jacobi matrices

The sequence of OP P_n satisfies a finite difference equation

$$zP_n(z) = b_{n+1}P_{n+1}(z) + a_nP_n(z) + b_nP_{n-1}(z), \quad P_{-1} = 0.$$

Hence the associated tri-diagonal symmetric matrix

$$J = \begin{pmatrix} a_0 & b_1 & 0 & 0 & \dots \\ b_1 & a_1 & b_2 & 0 & \dots \\ 0 & b_2 & a_2 & b_3 & \dots \\ \vdots & \ddots & \ddots & & \ddots \end{pmatrix}.$$

Just another remarkable free parametrization of a moment sequence, with uncountably many ramifications and applications.

OP literature

Herbert Stahl, Vilmos Totik: *General orthogonal polynomials*, Encyclopedia of Mathematics and its Applications, 43 Cambridge University Press, Cambridge, 1992. xii+250 pp.

Barry Simon: *Orthogonal poynomials on the unit circle*, 2. vol., Amer. Math. Soc. 2004.

The Oxford Handbook of Random Matrix Theory (Oxford Handbooks in Mathematics), (G. Akemann et al, eds.), Oxford Univ. Press, 2011.

Google citations (Nov. 2013):
OP 2,690,000
RM 17,000,000

The spectral theorem

Hilbert, Hellinger, Hahn, F. Riesz (around 1910)
Let U be a unitary transformation on a Hilbert space and let ξ be a fixed vector. The sequence $\langle U^n\xi, \xi\rangle, \quad n \in \mathbf{Z}$ is represented by a positive measure μ on the unit circle, because the kernel

$$s_{k-m} = \langle U^k\xi, U^m\xi\rangle$$

is positive semi-definite. Hence

$$\langle q(U)\xi, \xi\rangle = \int_{-\pi}^{\pi} q(e^{it})d\mu(t), \quad q \in \mathbf{C}[x].$$

With the known techniques of integration theory, one can define $f(U)$ for any bounded Borel function, so that

$$\langle f(U)\xi, \xi\rangle = \int_{-\pi}^{\pi} f(e^{it})d\mu(t).$$

Similarly

Given a bounded symmetric linear operator S, densely defined on a Hilbert space, and a vector ξ, the sequence

$$s_k = \langle S^k \xi, \xi \rangle, \quad k \geq 0$$

is a moment sequence, because the kernel $s_{k+m} = \langle S^k \xi, S^m \xi \rangle$ is positive semi-definite. Therefore there exists a positive measure on the line, such that

$$\langle f(A)\xi, \xi \rangle = \int f(x) d\sigma(x),$$

for every bounded Borel function on the line.

Self-adjointness

von Neumann (1929), Stone (1930)

The case of a symmetric linear operator S, densely defined but unbounded: even when all powers $S^n\xi$ are well defined and $s_{k+m} = \langle S^k\xi, S^m\xi\rangle$ is a moment sequence, the Borel functional calculus $f(S)$ may exist only on a larger Hilbert space. The self-adjoint condition

$$S = S^*$$

is necessary for having a spectral theorem/decomposition in the original Hilbert space.

Since then, this is the natural theoretical framework for quantum mechanics.

M. H. Stone: *Linear transformation in Hilbert space*, Amer. Math. Soc., 1932.

B. Simon: *The classical moment problem as a self-adjoint finite difference operator*, Adv. Math. , 137 (1998) pp. 82203.

Mark G. Krein (1907-1989)

Has incorporated the classical problem of moments in modern analysis, with great, original contributions. Master of the geometry of Hilbert space and function theory of a complex variable. A few of his topics of research:

Strings and spectral functions
Self-adjoint operators on Hilbert spaces of entire functions
Extremal problems related to the truncated moment problem
Convex analysis and duality
Prediction theory of stochastic processes
Stability and control of systems of differential equations
Representation theory of locally compact groups
Spectral analysis in spaces with an indefinite metric

Had 50 students, 805 descendants (many working today on moment problems) and was the uncontested mentor of the Ukrainian school of Functional Analysis.

Moment method in numerical mathematics

Given $f_0, \ldots, f_n$ linearly independent vectors in a Hilbert space, analyse the linear transformation $A_n : H_n \longrightarrow H_n$ such that

$$A_n f_k = \pi_n f_{k+1}, \quad 0 \leq k \leq n-1,$$

where $H_n = \mathrm{lin.span}(f_0, \ldots, f_{n-1})$ and π_n is the orthogonal projection onto H_n.

Yu. V. Vorobiov, *The method of moments in applied mathematics*, Fizmatgiz, Moscow, 1958.
Roger F. Harrington, *Field Computation by Moment Methods*, Wiley 1993.
L. Trefethen, M. Embree: *Spectra and pseudospectra*, Princeton, 2005.

Multivariate moments

Around for a century or so, still intriguing and occupying the recent generations. Challenges:

Algebraic structure of positive polynomials
Convex analysis of moment data
Matrix completion and extension of linear functionals
Function theory of several complex variables
Commuting systems of symmetric linear operatos
Determinateness

Some recent applications and ramifications

Global polynomial optimization
Geometric tomography
Bolzmann equations and max. entropy
Signal processing via wavelet transforms
Elliptic growth
Free probability theory

Free cumulants

"...free independence can be characterized by the vanishing of mixed cumulants. An important technical tool for deriving this characterization is a formula for free cumulants where the arguments are products of random variables. This formula is actually at the basis of many of our forthcoming results in later lectures and allows elegant proofs of many statements."

A. Nica, R. Speicher: *Lectures on the Combinatorics of Free Probability*, Cambridge Univ. Press, 2006.