Matrix Computations and the Secular Equation

Gene H. Golub

Stanford University

What is the secular equation?

"The term secular ('continuing through long ages' OED2) recalls that one of the origins of spectral theory was in the problem of the long-run behavior of the solar system investigated by Laplace and Lagrange. [...] The 1829 paper in which Cauchy established that the roots of a symmetric determinant are real has the title, 'Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des mouvements des planétes'; this signified only that Cauchy recognized that his problem, ***of choosing*** x ***to maximize*** $x^T Ax$ ***subject to*** $x^T x = 1$ *(to use modern notation), led to an equation like that studied in celestial mechanics. Sylvester's title 'On the Equation to the Secular Inequalities in the Planetary Theory' [...] was even more misleading as to content. In this tradition the 'Säkulärgleichung' of Courant and Hilbert's Methoden der Mathematischen Physik (1924) and the 'secular equation' of E. T. Browne's 'On the Separation Property of the Roots of the Secular Equation' American Journal of Mathematics, 52, (1930), 843-850 refer to the characteristic equation of a symmetric matrix."*

From `http://members.aol.com/jeff570/e.html`

Outline

Applications

Constrained Eigenvalue Problem

$$A = A^T$$

$$\max_{\mathbf{x}\neq\mathbf{0}} \quad \mathbf{x}^T A \mathbf{x}$$

$$\text{s.t.} \quad \mathbf{x}^T \mathbf{x} = 1$$

$$\mathbf{c}^T \mathbf{x} = 0$$

$$\phi(\mathbf{x}; \lambda, \mu) = \mathbf{x}^T A \mathbf{x} - \lambda(\mathbf{x}^T \mathbf{x} - 1) + 2\mu \mathbf{x}^T \mathbf{c}$$

$$\text{grad } \phi = 0 \implies A\mathbf{x} - \lambda \mathbf{x} + \mu \mathbf{c} = \mathbf{0}$$

$$\mathbf{x} = -\mu (A - \lambda I)^{-1} \mathbf{c}$$

$$\mathbf{c}^T \mathbf{x} = 0 \implies \mathbf{c}^T (A - \lambda I)^{-1} \mathbf{c} = 0$$

Constrained Eigenvalue Secular Equation

$$A = Q \Lambda Q^T, \mathbf{d} = Q^T \mathbf{c}$$

$$\sum_{i=1}^{n} \frac{d_i^2}{(\lambda_i - \lambda)} = 0$$

Rank One Change

$$A\mathbf{x} = \lambda\mathbf{x}$$

$$(A + \mathbf{c}\mathbf{c}^T)\mathbf{y} = \mu\mathbf{y}$$

Rank One Change Secular Equation

$$1 + \mathbf{c}^T(A - \mu I)^{-1}\mathbf{c} = 0$$

Rank k-change

$$(A + CC^T)\mathbf{y} = \mu\mathbf{y}$$

$$\det\big(I + C^T(A - \mu I)^{-1}C\big) = 0$$

Another secular equation

Consider

$$\begin{pmatrix} A & \mathbf{b} \\ \mathbf{b}^T & c \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ y \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{x} \\ y \end{pmatrix}.$$

Then

$$(A - \lambda I)\mathbf{x} = -y\mathbf{b}$$

and hence,

$$(c - \lambda - \mathbf{b}(A^T - \lambda I)^{-1}\mathbf{b})y = 0.$$

Hence we must solve another secular equation when the matrix is expanded.

Quadratic Constraint

$$A = A^T, \quad \text{positive definite}$$

$$\begin{array}{ll} \min_{\mathbf{x}} & \mathbf{x}^T A\mathbf{x} - 2\mathbf{c}^T\mathbf{x} \\ \text{s.t.} & \mathbf{x}^T\mathbf{x} = \alpha^2 \end{array}$$

$$\phi(\mathbf{x}; \lambda) = \mathbf{x}^T A\mathbf{x} - 2\mathbf{c}^T\mathbf{x} - \lambda(\mathbf{x}^T\mathbf{x} - \alpha^2)$$

$$\text{grad } \phi = 0 \implies$$

$$(A - \lambda I)\mathbf{x} - \mathbf{c} = \mathbf{0}$$

Quadratic Constraint Secular Equation

$$\mathbf{c}^T(A - \lambda I)^{-2}\mathbf{c} = \alpha^2$$

Least Squares with a Quadratic Constraint

$$\mathbf{b}^T A(A^T A - \lambda I)^{-2} A^T\mathbf{b} = \alpha^2$$

Total Least Squares (TLS)

$$(A + E)\mathbf{x} = \mathbf{b} + \mathbf{r}, \quad A : m \times n$$

$$\begin{pmatrix} A & \mathbf{b} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ -1 \end{pmatrix} + \begin{pmatrix} E & \mathbf{r} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ -1 \end{pmatrix} = \mathbf{0}$$

$$(C + F)\mathbf{z} = \mathbf{0}; C : m \times n + 1$$

Determine F and $\mathbf{z}$ so that

$$\text{rank } (C + F) \leq n \text{ and } ||F||_F = \min.$$

Equivalently, find

$$\min_{\mathbf{z}} \quad \frac{||C\mathbf{z}||_2}{||\mathbf{z}||_2} \equiv \sigma_{\min}(C)$$

Total Least Squares (cont.)

$$C^T C\mathbf{z} = \hat{\sigma}^2\mathbf{z}$$

Total Least Squares Secular Equation

$$\mathbf{b}^T A(A^T A - \hat{\sigma}^2 I)^{-1} A^T\mathbf{b} - \mathbf{b}^T\mathbf{b} - \hat{\sigma}^2 = 0$$

$$\hat{\sigma} < \sigma_{\min}(A)$$

$$\mathbf{x}_{TLS} = (A^T A - \hat{\sigma}^2 I)^{-1} A^T\mathbf{b}$$

Data Least Squares (DLS)

$$(A + E)\mathbf{x} = \mathbf{b}$$

$$\mathbf{b}^T A(A^T A - \hat{\tau}^2 I)^{-1} A^T\mathbf{b} - \mathbf{b}^T\mathbf{b} = 0$$

$$\hat{\tau} < \sigma_{\min}(A)$$

Regularized total least squares (Fischer/G.)

Note, that the TLS solution is equivalent to

$$\min \frac{\|\mathbf{b} - A\mathbf{x}\|_2^2}{1 + \|\mathbf{x}\|_2^2} = \min \frac{\|C\mathbf{z}\|_2^2}{\|\mathbf{z}\|_2^2} = \sigma_{\min}(C),$$

where

$$C = (A, \mathbf{b}) \quad \text{and} \quad z_{n+1} = -1.$$

For the **regularized TLS** we consider

$$\min \frac{\|\mathbf{b} - A\mathbf{x}\|_2^2}{1 + \mathbf{x}^T V \mathbf{x}}, \text{ subject to } \mathbf{x}^T V \mathbf{x} = \alpha^2,$$

where V is a given symmetric positive definite matrix. Now, let

$$W = \begin{pmatrix} V & 0 \\ 0 & 1 \end{pmatrix} = F^T F$$

and observe that

$$\min \frac{\|\mathbf{b} - A\mathbf{x}\|_2^2}{1 + \mathbf{x}^T V \mathbf{x}} = \min \frac{\|C\mathbf{z}\|_2^2}{\mathbf{z}^T W \mathbf{z}}$$

with $\|\mathbf{z}\|_2^2 = 1 + \alpha^2, \; z_{n+1} = -1.$

Least squares with linear and quadratic constraints

With

$$\mathbf{y} = F\mathbf{z}, B = F^{-T}C^TCF^{-1}, \mathbf{c} = \mathbf{e}_{n+1}^T F^{-1},$$

$$\gamma^2 = 1 + \alpha^2, \text{ and } \beta = -1$$

we may rewrite our regularized TLS problem in terms of a **least squares problem with linear and quadratic constraints**

$$\min \frac{\mathbf{y}^\mathbf{T} B\mathbf{y}}{\mathbf{y}^T\mathbf{y}}, \quad \text{s. t. } \|\mathbf{y}\|_2^2 = \gamma^2, \ \mathbf{c}^T\mathbf{y} = \beta.$$

where γ and β are non-zero.

Lagrange multipliers

$$\psi(\mathbf{y}; \lambda, \mu) = \mathbf{y}^\mathbf{T} B\mathbf{y} - \lambda(\mathbf{y}^\mathbf{T}\mathbf{y} - \gamma^2) - 2\mu(\mathbf{c}^T\mathbf{y} - \beta).$$

grad $\psi = 0$ when

$$B\mathbf{y} - \lambda\mathbf{y} - \mu\mathbf{c} = \mathbf{0}.$$

Introducing the projection matrix

$$P = I - \frac{\mathbf{c}\mathbf{c}^T}{\mathbf{c}^T\mathbf{c}} \text{ and } \mathbf{d} = \frac{\beta\mathbf{c}}{\mathbf{c}^T\mathbf{c}}$$

we arrive at

$$\begin{aligned} (PB - \lambda I)\mathbf{y} &= -\lambda\mathbf{d} \\ \mathbf{y}^T\mathbf{y} &= \gamma^2, \end{aligned}$$

which leads to the secular equation

$$\lambda^2\mathbf{d}^T(PB - \lambda I)^{-T}(PB - \lambda I)^{-1}\mathbf{d} = \gamma^2.$$

Instead, consider

$$\begin{pmatrix} (PB - \lambda I)(PB - \lambda I)^T & \lambda\mathbf{d} \\ \lambda\mathbf{d}^T & \gamma^2 \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \xi \end{pmatrix} = 0.$$

Note,

$$\left((PB - \lambda I)(PB - \lambda I)^T - \frac{\lambda^2}{\gamma^2}\mathbf{d}\mathbf{d}^{\mathbf{T}}\right)\mathbf{u} = 0.$$

Thus, λ can be found as an eigenvalue of a quadratic eigenvalue problem with $\hat{\mathbf{y}} = \mathbf{u}/\xi$.

Approximating the secular equation

How do we approximate the secular equation for large n?

The problems we have described are closely associated with estimating a quadratic form

$$\mathbf{u}^T F(A)\mathbf{u}$$

where $\mathbf{u}$ is a given vector and A is a symmetric matrix.

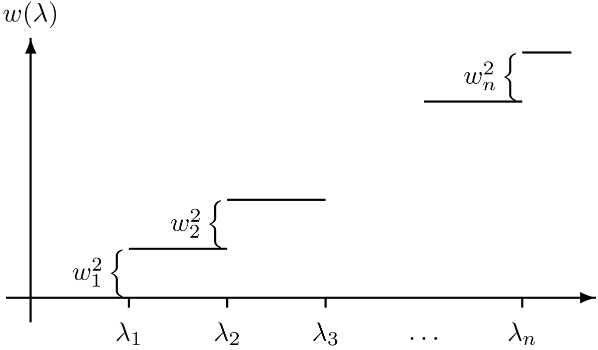
Matrix Function to Integral

$$A = Q \Lambda Q^T$$

$$\mathbf{u}^T F(A) \mathbf{u} = \mathbf{u}^T F(Q \Lambda Q^T) \mathbf{u} = \mathbf{u}^T Q F(\Lambda) Q^T \mathbf{u} = \mathbf{w}^T F(\Lambda) \mathbf{w}$$

$$\mathbf{w} = Q^T \mathbf{u}$$

$$\mathbf{u}^T F(A) \mathbf{u} = \sum_{i=1}^{n} F(\lambda_i) w_i^2 = \int_a^b F(\lambda) dw(\lambda)$$



Gauss-Radau Quadrature Rules

$$L \leq \int_a^b F(\lambda) dw(\lambda) \leq U$$

$$\mu_r = \int \lambda^r dw(\lambda) \qquad (r = 0, 1, \ldots, 2k + m - 1)$$

$$\int_a^b F(\lambda) dw(\lambda) = I[F] + R[F]$$

$$I[F] = \sum_{i=1}^{k} A_i F(t_i) + \sum_{j=1}^{m} B_j F(z_j)$$

$\{A_i, t_i\}_{i=1}^k$	*unknown* weights and nodes
$\{z_j\}_{j=1}^m$	prescribed nodes
$\{B_j\}_{j=1}^m$	calculated weights

Gauss-Radau Quadrature Rules (cont.)

$$I(\lambda^r) = \mu_r$$

$$\mu_r = \sum_{i=1}^{k} A_i t_i^r + \sum_{j=1}^{m} B_j z_j^r$$

System of non-linear equations.

$$R[F] = \frac{F^{(2k+m)}(\eta)}{(2k+m)!} \int_a^b \prod_{j=1}^{m}(\lambda - z_j) \left[\prod_{i=1}^{k}(\lambda - t_i)\right]^2 dw(\lambda)$$

$$a < \eta < b$$

$m = 1$

$$F^{(2k+1)}(\eta) \le 0 \text{ and } z_1 = a \qquad R[F] \le 0 \quad I[F] = U$$

$$F^{(2k+1)}(\eta) \le 0 \text{ and } z_1 = b \qquad R[F] \ge 0 \quad I[F] = L$$

Gauss Quadrature

$$\int p_r(\lambda)p_s(\lambda)d\alpha(\lambda) = 0, \qquad r \neq s, (r, s = 0, 1, \ldots, k)$$

$$p_{j+1}(\lambda) = (\lambda - \xi_{j+1})p_j(\lambda) - \eta_j^2 p_{j-1}(\lambda)$$

$$p_k(t_i) = 0, \qquad i = 1, 2, \ldots, k$$

$$J_k = \begin{pmatrix} \xi_1 & \eta_1 & & & \\ \eta_1 & \xi_2 & \eta_2 & & \\ & \eta_2 & \ddots & \ddots & \\ & & \ddots & \ddots & \eta_{k-1} \\ & & & \eta_{k-1} & \xi_k \end{pmatrix}$$

$$\mu_0 = 1$$

$$J_k \mathbf{v}_j = t_j \mathbf{v}_j, \quad j = 1, 2, \ldots, k$$

$$A_j = v_{1j}^2, \quad j = 1, 2, \ldots, k$$

Gauss-Radau (Inverse Eigenvalue Problem)

$$\bar{J}_{k+1} = \begin{pmatrix} & & & 0 \\ & J_k & & \vdots \\ & & & \eta_k \\ 0 & \cdots & \eta_k & \bar{\xi}_{k+1} \end{pmatrix}$$

$$0 = p_{k+1}(t_0) = (t_0 - \bar{\xi}_{k+1})p_k(t_0) - \eta_k^2 p_{k-1}(t_0)$$

$$\bar{\xi}_{k+1} = t_0 - \eta_k^2 \frac{p_{k-1}(t_0)}{p_k(t_0)}$$

or

$$(J_k - t_0 I)\delta = \eta_k^2 \mathbf{e}_k$$

$$\bar{\xi}_{k+1} = t_0 + \delta_k$$

Evaluate $I[F]$

$$I[F] = \sum_{i=0}^{k} v_{1i}^2 F(t_i)$$

$$\bar{J}_{k+1} = VTV^T$$

$$V^T \mathbf{e}_1 = \langle \text{ first component of } V \ \rangle$$

$$\begin{aligned} I[F] &= \mathbf{e}_1^T V F(T) V^T \mathbf{e}_1 \\ &= \mathbf{e}_1^T F(VTV^T) \mathbf{e}_1 \\ &= \mathbf{e}_1^T F(\bar{J}_{k+1}) \mathbf{e}_1 \end{aligned}$$

Orthonormal polynomials w.r.t the measure $w(\lambda)$

How do we build these polynomials?

$$p_{j+1}(\lambda) = (\lambda - \xi_{j+1})p_j(\lambda) - \eta_j^2 p_{j-1}(\lambda)$$
$$p_{j+1}(A) = (A - \xi_{j+1}I)p_j(A) - \eta_j^2 p_{j-1}(A)$$
$$p_{j+1}(A)\mathbf{u} = (A - \xi_{j+1}I)p_j(A)\mathbf{u} - \eta_j^2 p_{j-1}(A)\mathbf{u}$$

Set $\mathbf{w}_j = p_j(A)\mathbf{u}$.
We define ξ_{j+1} and η_j^2 so that

$$\mathbf{w}_{j+1}^T \mathbf{w}_j = 0$$
$$\mathbf{w}_{j+1}^T \mathbf{w}_{j-1} = 0,$$

and then

$$\mathbf{w}_{j+1}^T \mathbf{w}_r = 0 \qquad \text{for } r < j-1$$

$$\xi_{j+1} = \frac{(\mathbf{w}_j, A\mathbf{w}_j)}{(\mathbf{w}_j, \mathbf{w}_j)} \text{ and } \eta_j^2 = \frac{(\mathbf{w}_j, \mathbf{w}_j)}{(\mathbf{w}_{j-1}, \mathbf{w}_{j-1})}$$

Orthonormal polynomials w.r.t the measure $w(\lambda)$

$$\mathbf{w}_{j+1}^T \mathbf{w}_r = 0 \qquad \text{for } r < j-1$$

$$\xi_{j+1} = \frac{(\mathbf{w}_j, A\mathbf{w}_j)}{(\mathbf{w}_j, \mathbf{w}_j)} \text{ and } \eta_j^2 = \frac{(\mathbf{w}_j, \mathbf{w}_j)}{(\mathbf{w}_{j-1}, \mathbf{w}_{j-1})}$$

The Lanczos Process! To construct J_k, begin the Lanczos process with $\mathbf{u}$, then

$$\begin{aligned}
(\mathbf{w}_j, \mathbf{w}_k) &= 0 \\
&= (p_j(A)\mathbf{u}, p_k(A)\mathbf{u}) \\
&= \mathbf{u}^T Q p_j(\Lambda) Q^T Q p_k(\Lambda) Q^T \mathbf{u} \\
&= \mathbf{w}^T p_j(\Lambda) p_k(\Lambda) \mathbf{w} \\
&= \int p_j(\lambda) p_k(\lambda) dw(\lambda)
\end{aligned}$$

Examples

An example

We need to solve

$$\mathbf{b}^T(A+\mu I)^{-2}\mathbf{b} = \alpha^2$$

Algorithm

1. Begin Lanczos process with $\mathbf{u} = \mathbf{b}$
2. Construct $\bar{J}_{k+1}$
3. Solve $\mathbf{e}_1^T(\bar{J}_{k+1} + \mu I)^{-2}\mathbf{e}_1 = \alpha^2$.

Numerical Comparison

Numerical Comparison with Total Least Squares

$$\begin{aligned} \min_{E,\mathbf{r}} \quad & \| \begin{pmatrix} E & \mathbf{r} \end{pmatrix} \|_F \\ \text{s.t.} \quad & (A+E)\mathbf{x} = \mathbf{b} + \mathbf{r} \end{aligned}$$

$$\psi(\hat{\sigma}^2) = b^T A(A^T A - \hat{\sigma}^2 I)^{-1} A^T b - b^T b - \hat{\sigma}^2 = 0$$

Algorithms

- Approximate $b^T A(A^T A - \hat{\sigma}^2 I)^{-1} A^T b$ using moment theory and Lanczos on $A^T A$
- Approximate $b^T A(A^T A - \hat{\sigma}^2 I)^{-1} A^T b$ using moment theory and Lanczos bidiagonalization on A
- Solve a set of non-linear equations derived from the normal equations. (Björck's algorithm)

Solving secular equations with moments

Given a current approximation to the value $\hat{\sigma}_k^2$, we consider updates of the form

$$\hat{\sigma}_{k+1}^2 = \hat{\sigma}_k^2 - \frac{\psi(\hat{\sigma}_k^2)}{\psi'(\hat{\sigma}_k^2)} C_k.$$

Method	$\mathbf{C_k}$	**Interp. func.**
Newton's	1	$c_0 + \hat{\sigma} c_1$
SRA[1]	$\frac{\|\|b\|\|^2 - \psi(\hat{\sigma}_k^2)}{\|\|b\|\|^2}$	$\|\|b\|\| - \frac{c_1}{c_2 - \hat{\sigma}^2}$
Halley's	$1 / \left(1 - \frac{\psi(\hat{\sigma}_k^2)\psi''(\hat{\sigma}_k^2)}{2(\psi'(\hat{\sigma}_k^2)^2)}\right)$	$c_0 - \frac{c_1}{c_2 - \hat{\sigma}^2}$

The derivatives in this equation are secular equations themselves. We can use the same procedure to compute estimates of the derivatives by changing the function f applied to the matrix $\bar{J}$.

[1] Simple Rational Approximation

Solving secular equations with moments

Recall that we need an estimate of $\lambda_{\min}$ and $\lambda_{\max}$ of $A^T A$ to use for the upper and lower bounds in the quadrature rules. We set

$$b = ||A||_1 ||A||_\infty > \lambda_{\max} \text{ and } a = 10^{-9} \overset{?}{<} \lambda_{\min}.$$

In the TLS problem, we need $\hat{\sigma} < \sigma_{\min}(A)$ and employ bisection to guarantee this condition.

Algorithm

1. $\hat{\sigma}^2_{\min} = \min |a_{ij}|^2$
2. While not converged...
3. Compute an approximation to the secular function $\phi(\hat{\sigma}^2_k), \phi'(\hat{\sigma}^2_k), \phi''(\hat{\sigma}^2_k)$
4. If the approximation failed because the bounds on the secular function are not monotone, set $\hat{\sigma}^2_{\min} = \hat{\sigma}^2_k$ and $\hat{\sigma}^2_{k+1} = (1/2)\hat{\sigma}^2_k$
5. Otherwise, set $\hat{\sigma}^2_{k+1} = \hat{\sigma}^2_k - \frac{\psi(\hat{\sigma}^2_k)}{\psi'(\hat{\sigma}^2_k)} C_k$, repeat.

Björck's algorithm

Solve the system of nonlinear equations

$$\begin{pmatrix} A^TA & A^Tb \\ b^TA & b^Tb \end{pmatrix} \begin{pmatrix} x \\ -1 \end{pmatrix} = \lambda \begin{pmatrix} x \\ -1 \end{pmatrix},$$

or equivalently, the system

$$\begin{pmatrix} f(x,\lambda) \\ g(x,\lambda) \end{pmatrix} = \begin{pmatrix} -A^Tr - \lambda x \\ -b^Tr + \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

with $r = b - Ax$ using a Rayleigh-quotient iteration (RQI). (Note, λ is used in place of $\hat{\sigma}^2$ in this derivation.)

This algorithm will always converge to a singular value/vector pair, but we might not get $\lambda = \hat{\sigma}^2$. Björck suggested one initial inverse iteration (i.e. $\lambda = 0$) to move closer to the desired λ, and then apply the RQI procedure.

Details of the matrix moments based algorithm

- Algorithm 2 uses the Golub-Kahan bidiagonalization of A and applies the moment algorithm to $T = B^T B$ instead of computing T directly from the Lanczos process on $A^T A$.
- Algorithm 1 restarts the Lanczos process at each iteration.
- Algorithm 2 never restarts the bidiagonalization process and simply continues the process at each iteration.

Problems

- Jo's problems, 15×8 and 750×400
- Björck's problem 1: 30×15 matrix
- Large scale problems with 10000×5000 and 100000×60000 matrices.

The large scale problems were generated using random Householder matrices to build the SVD of $\begin{bmatrix} A & b \end{bmatrix}$ in product form. Each large-scale matrix was available solely as an operator to all of the algorithms. The singular values of

$$\begin{bmatrix} A & b \end{bmatrix}$$

are

$$\sigma_i = \log(i) + |N(0,1)|,$$

where $N(0,1)$ is a standard normal random variable.

Parameter choices

		Algorithm 1 (Tridiag...)			Algorithm 2 (Bidiag...)		
		$\lambda^{(0)} = 0$	$\lambda^{(0)} = 1$	$\lambda^{(0)} = \rho$	$\lambda^{(0)} = 0$	$\lambda^{(0)} = 1$	$\lambda^{(0)} = \rho$
1	newton	6	4	5	6	4	5
	sra	5	5	5	5	5	5
	halley	5	5	6	5	5	6
2	newton	++	++	–	*8	*8	*7
	sra	++	–	++	*12	*24	*7
	halley	++	++	++	*14	*23	*6
3	newton	–	–	–	*20	*7	*10
	sra	–	–	–	*20	25	*64
	halley	–	–	–	*55	55	*12
4	newton	–	–	–	*15	*11	*11
	sra	++	–	–	*15	*25	*14
	halley	–	–	–	*20	*57	*11
5	newton	100	–	–	++	++	++
	sra	100	-5	–	++	++	–
	halley	100	–	–	++	++	–

* wrong root; ++ correct w/o convergence; – no convergence

$\rho = ||A * \mathbf{x}_{ls}||^2 / (||\mathbf{x}_{ls}||^2 + 1)$

Problem	1	2	3	4	5
Size	(15,8)	(750,400)	(10000,5000)	(100000,60000)	(30,15)

Convergence

TEST	ALG	ITERS	ERROR	TIME	LANZ.
`jo`	`björck`	6	1.0×10^{-14}	0	
$(15, 8)$	`Alg 1`	5	4.4×10^{-16}	0	
$\sigma^2 = 5.6 \times 10^{-1}$	`Alg 2`	5	3.3×10^{-14}	0	12
`jo`	`björck`	7	8.5×10^{0}	0.2	
$(750, 400)$	`Alg 1`	>100	8.5×10^{-14}	52.5	
$\sigma^2 = 1.8 \times 10^{1}$	`Alg 2`	23	5.0×10^{-1}	0.7	163
`large-scale`	`björck`	8	1.1×10^{-16}	0.5	
$(10000, 5000)$	`Alg 1`	>100	1.0×10^{-3}	36.1	
$\sigma^2 = 1.9 \times 10^{-1}$	`Alg 2`	55	8.3×10^{-16}	1.5	152
`large-scale`	`björck`	5	3.9×10^{-17}	5.1	
$(100000, 60000)$	`Alg 1`	>100	5.5×10^{-7}	324.9	
$\sigma^2 = 3.5 \times 10^{-3}$	`Alg 2`	57	5.3×10^{-8}	14.6	155
`björck`	`björck`	7	2.6×10^{-19}	0	
$(30, 15)$	`Alg 1`	>100	σ_2	0.3	
$\sigma^2 = 9.9 \times 10^{-12}$	`Alg 2`	18	2.9×10^{4}	0	33

Conclusions

- The secular equation unifies many problems in matrix theory.
- We can approximate the secular equation using Gaussian quadrature and derive upper and lower bounds.
- When combined with robust root-finding procedures, we can use these bounds in algorithms to solve large scale problems.
- Finding zeros is difficult!

Thanks to colleagues and co-authors

- Zhajun Bai
- David Gleich
- Sung-Eun Jo
- Gerard Meurant
- Bernd Fischer