Chapter 12
Solving secular equations

Gérard MEURANT

January-February, 2012

Examples of secular equations

What are secular equations?

The term “secular” comes from the latin “saecularis” which is related to “saeculum”, which means “century”

So secular refers to something that is done or happens every century. It is also used to refer to something that is several centuries old

It appeared in mathematics to denote equations related to the motion of planets and celestial mechanics

For instance, it appears in the title of a 1829 paper of A. L. Cauchy (1789–1857) *“Sur l’équation à l’aide de laquelle on détermine les inégalités séculaires des mouvements des planètes”* (Oeuvres Complètes (IIème Série), v 9 (1891), pp 174–195). There is also a paper by J. J. Sylvester (1814–1897) whose title is *“On the equation to the secular inequalities in the planetary theory”* (Phil. Mag., v 5 n 16 (1883), pp 267-269)

Eigenvalues of a Tridiagonal Matrix

We look for an eigenvalue λ and an eigenvector $x = \begin{pmatrix} y & \zeta \end{pmatrix}^T$ of J_{k+1} where y is a vector of dimension k and ζ is a real number. Then

$$J_k y + \eta_k \zeta e^k = \lambda y$$

$$\eta_k y_k + \alpha_{k+1} \zeta = \lambda \zeta$$

where y_k is the last component of y, α_j, $j = 1, \ldots, k+1$ are the diagonal entries of J_{k+1} and η_j, $j = 1, \ldots, k$ are the subdiagonal entries

By eliminating the vector y from these two equations we have

$$(\alpha_{k+1} - \eta_k^2 (e^k)^T (J_k - \lambda I)^{-1} e^k) \zeta = \lambda \zeta$$

$$\alpha_{k+1} - \eta_k^2 \sum_{j=1}^{k} \frac{(\xi_j)^2}{\theta_j - \lambda} = \lambda$$

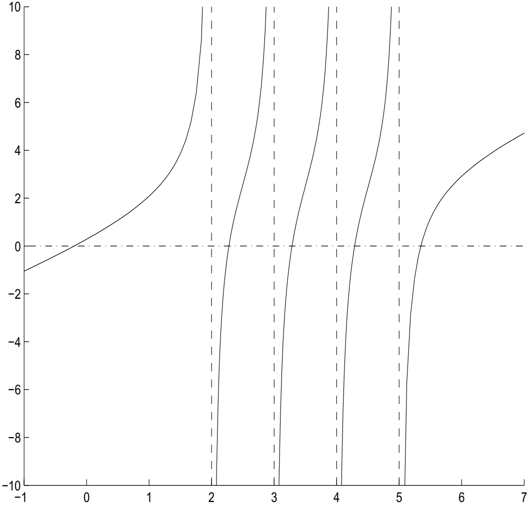
where $\xi_j = z_k^j$ is the kth (i.e., last) component of the jth eigenvector of J_k

the θ_j's are the eigenvalues of J_k, that is, the Ritz values

Too obtain the eigenvalues of J_{k+1} from those of J_k we have to solve

$$f(\lambda) = \lambda - \alpha_{k+1} + \eta_k^2 \sum_{j=1}^{k} \frac{\xi_j^2}{\theta_j - \lambda} = 0$$

The secular function f has poles at the eigenvalues (Ritz values) of J_k for $\lambda = \theta_j = \theta_j^{(k)}, j = 1 \dots, k$



Example of secular function with $k = 4$

Modification by a Rank-One Matrix

Assume that we know the eigenvalues of a matrix A and we would like to compute the eigenvalues of a rank-one modification of A

We have

$$Ax = \lambda x$$

where we know the eigenvalues λ and we want to compute μ such that

$$(A + cc^T)y = \mu y$$

where c is a given vector (not orthogonal to an eigenvector of A)

Clearly μ is not an eigenvalue of A

Therefore $A - \mu I$ is nonsingular and we obtain an equation for μ

$$y = -(A - \mu I)^{-1} cc^T y$$

Multiplying by c^T to the left,

$$c^T y = -c^T (A - \mu I)^{-1} cc^T y$$

The secular equation is

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

Using the spectral decomposition of $A = Q\Lambda Q^T$ with Q orthogonal and Λ diagonal and $z = Q^T c$

$$1 + \sum_{j=1}^{n} \frac{(z_j)^2}{\lambda_j - \mu} = 0$$

where λ_j are the eigenvalues of A

Constrained Eigenvalue Problem

We wish to find a vector x of norm one which is the solution of

$$\max_x x^T A x$$

satisfying the constraint $c^T x = 0$ where c is a given vector
We introduce a functional φ with two Lagrange multipliers λ and μ corresponding to the two constraints

$$\varphi(x, \lambda, \mu) = x^T A x - \lambda(x^T x - 1) + 2\mu x^T c$$

Computing the gradient of φ with respect to x, which must be zero at the solution,

$$Ax - \lambda x + \mu c = 0$$

from which we have $x = -\mu(A - \lambda I)^{-1} c$
If λ is not an eigenvalue of A and using $c^T x = 0$ we have

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

Using the spectral decomposition of $A = Q\Lambda Q^T$ and $d = Q^T c$

$$f(\lambda) = \sum_{j=1}^{n} \frac{d_j^2}{\lambda_j - \lambda} = 0$$

There are $n-1$ solutions to the secular equation
When we have the values of λ that are solutions, we use the constraint $x^T x = 1$ to remark that

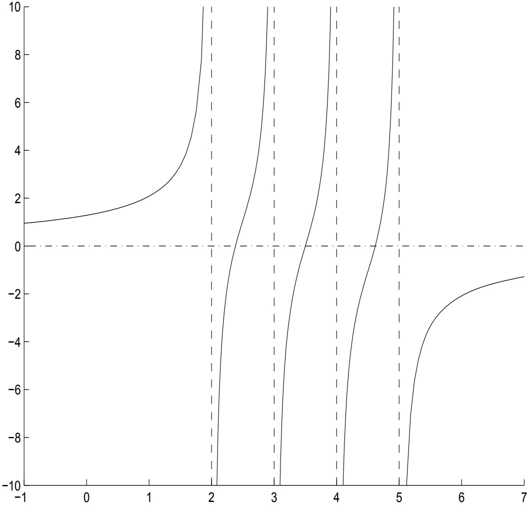
$$x^T x = \mu^2 c^T (A - \lambda I)^{-2} c = 1$$

Therefore,

$$\mu^2 = \frac{1}{c^T (A - \lambda I)^{-2} c}$$

and

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



Example of secular function

Eigenvalue Problem with a Quadratic Constraint

We consider the problem

$$\min_x x^T A x - 2c^T x$$

with the constraint $x^T x = \alpha^2$

Introducing a Lagrange multiplier and using the stationary values of the Lagrange functional

$$(A - \lambda I)x = c, \quad x^T x = \alpha^2$$

Let $A = Q \Lambda Q^T$

The Lagrange equations can be written as

$$\Lambda Q^T x - \lambda Q^T x = Q^T c, \quad x^T Q Q^T x = \alpha^2$$

Introducing $y = Q^T x$ and $d = Q^T c$

$$\Lambda y - \lambda y = d, \quad y^T y = \alpha^2$$

Assume that all the eigenvalues λ_i of A are simple

If λ is equal to one of the eigenvalues, say λ_j, we must have $d_j = 0$

For all $i \neq j$ we have

$$y_i = \frac{d_i}{\lambda_i - \lambda}$$

Then, there is a solution or not, whether we have

$$\sum_{i \neq j} \left(\frac{d_i}{\lambda_i - \lambda_j} \right)^2 = \alpha^2$$

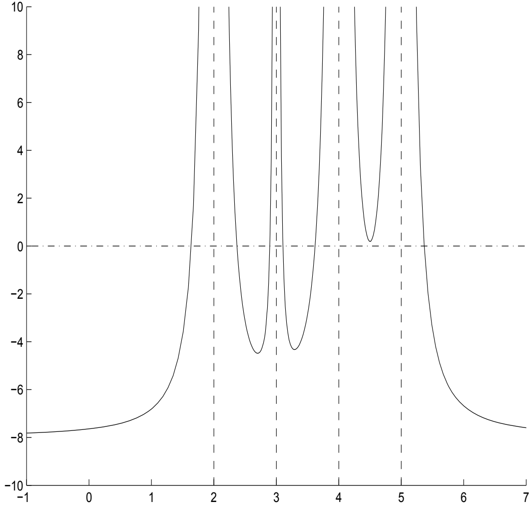
or not

If λ is not an eigenvalue of A, the inverse of $A - \lambda I$ exists and we obtain the secular equation

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

which is written as

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



Example of secular function

Secular equation solvers

Consider solving the equation

$$1+\rho\sum_{j=1}^{n}\frac{(c_j)^2}{d_j-\lambda}=0$$

with $\rho>0$

When we look for the solution in the interval $]d_i,d_{i+1}[$, we make a change of variable $\lambda=d_i+\rho t$

Denoting $\delta_j=(d_j-d_i)/\rho$, the secular equation is

$$f(t)=1+\sum_{j=1}^{i}\frac{c_j^2}{\delta_j-t}+\sum_{j=i+1}^{n}\frac{c_j^2}{\delta_j-t}=1+\psi(t)+\phi(t)=0$$

Note that

$$\psi(t)=\sum_{j=1}^{i-1}\frac{c_j^2}{\delta_j-t}-\frac{c_i^2}{t}$$

since $\delta_i=0$

The function ψ has poles $\delta_1,\ldots,\delta_{i-1},0$ with $\delta_j<0$, $j=1,\ldots,i-1$

The solution is sought in the interval $]0, \delta_{i+1}[$ with $\delta_{i+1} > 0$

Let us denote $\delta = \delta_{i+1}$

In $]0, \delta[$ we have $\psi(t) < 0$ and $\phi(t) > 0$

Assume we know all the poles of the secular function f and we are able to compute values of ψ and ϕ and their derivatives

BNS methods

Bunch, Nielsen and Sorensen interpolated ψ to first order by a rational function $p/(q-t)$ and ϕ by $r+s/(\delta-t)$
This is called osculatory interpolation
The parameters p, q, r, s are determined by matching the exact values of the function and the first derivative of ψ or ϕ at some given point $\bar{t}$ (to the right of the exact solution) where f has a negative value

$$q = \bar{t} + \psi(\bar{t})/\psi'(\bar{t})$$

$$p = \psi(\bar{t})^2/\psi'(\bar{t})$$

$$r = \phi(\bar{t}) - (\delta - \bar{t})\phi'(\bar{t})$$

$$s = (\delta - \bar{t})^2\phi'(\bar{t})$$

For computing them we have

$$\psi'(t) = \sum_{j=1}^{i} \frac{c_j^2}{(\delta_j - t)^2}$$

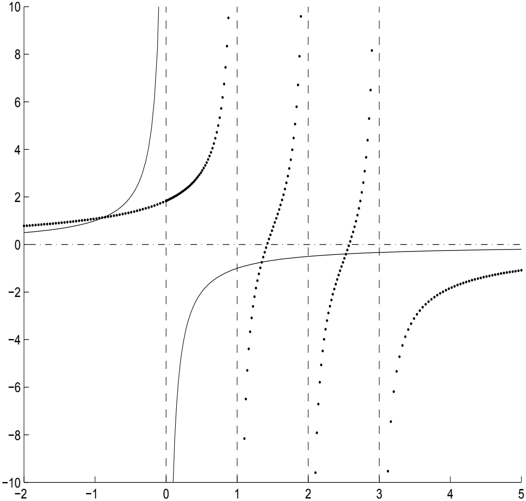
and

$$\phi'(t) = \sum_{j=i+1}^{n} \frac{c_j^2}{(\delta_j - t)^2}$$

Then the new iterate is obtained by solving the quadratic equation

$$1 + \frac{p}{q-t} + r + \frac{s}{\delta - t} = 0$$

This equation has two roots, of which only one is of interest
This method is called BNS1



Functions ψ (solid) and ϕ (dots) as function of t

To obtain an efficient method it remains to find a good initial guess

Melman's strategy is to look for a zero of an interpolant of the function $tf(t)$
When seeking the ith root, this function is written as

$$tf(t) = t\left(1 - \frac{c_i^2}{t} + \frac{c_{i+1}^2}{\delta - t} + h(t)\right)$$

The function h is written in two parts $h = h_1 + h_2$

$$h_1(t) = \sum_{j=1}^{i-1} \frac{c_j^2}{\delta_j - t}, \quad h_2(t) = \sum_{j=i+2}^{n} \frac{c_j^2}{\delta_j - t}.$$

The functions h_1 and h_2 do not have poles in the interval of concern

They are interpolated by rational functions of the form $p/(q-t)$

The parameters are found by interpolating the function values at points 0 and δ

This gives a function $\bar{h} = \bar{h}_1 + \bar{h}_2$

The starting point is obtained by computing the zero (in the interval $]0, \delta[$) of the function

$$t\left(1 - \frac{c_i^2}{t} + \frac{c_{i+1}^2}{\delta - t} + \bar{h}(t)\right)$$

This can be done with a standard zero finder

BNS2

Interpolating ψ by $\bar{r} + \bar{s}/t$ and ϕ by $\bar{p}/(\bar{q} - t)$, one obtains another method called BNS2
Then,

$$\bar{r} = \psi(\bar{t}) - (\delta - \bar{t})\psi'(\bar{t}),$$

$$\bar{s} = (\delta - \bar{t})^2\psi'(\bar{t})$$

$$\bar{q} = \bar{t} + \phi(\bar{t})/\phi'(\bar{t})$$

$$\bar{p} = \phi(\bar{t})^2/\phi'(\bar{t})$$

MW

R. C. Li proposed to use $r + s/(\delta - t)$ to interpolate both functions ψ and ϕ

In fact, we interpolate ψ, which has a pole at 0 by $\bar{r} + \bar{s}/t$ and ϕ, which has a pole at δ by $r + s/(\delta - t)$

This method is not monotonic, but has quadratic convergence

Gragg's method

Gragg interpolates f at $\bar{t}$ to second order with the function

$$a + b/t + c/(\delta - t)$$

This gives

$$c = \frac{(\delta - \bar{t})^3}{\delta} f' + \frac{\bar{t}(\delta - \bar{t})^3}{2\delta} f''$$

$$b = \frac{\bar{t}^3}{2\delta} \left[f''(\delta - \bar{t}) - 2f' \right]$$

$$a = f - \frac{f''\bar{t}}{2}(\delta - \bar{t}) + (2\bar{t} - \delta) f'$$

Then we solve the quadratic equation

$$at^2 - t(a\delta - b + c) - b\delta = 0$$

The convergence is cubic

Numerical experiments

$$1 + \rho \sum_{j=1}^{n} \frac{(c_j)^2}{d_j - \lambda} = 0$$

The dimension is $n = 4$ and $d = [1,\ 1+\beta,\ 3,\ 4]$
Defining $v^T = [\gamma,\ \omega,\ \ 1,\ 1]$, we have $c^T = v^T/\|v\|$

$$\beta = 1, \gamma = 10^{-2}, \omega = 1$$

Method	No. it.	Root−1	f(Root)
BNS1	2	$2.068910657999406\ 10^{-5}$	$-4.2294\ 10^{-12}$
BNS2	2	$2.068910658015177\ 10^{-5}$	$8.0522\ 10^{-12}$
MW	2	$2.068910657999406\ 10^{-5}$	$-4.2294\ 10^{-12}$
GR	2	$2.068910657999392\ 10^{-5}$	$-4.2398\ 10^{-12}$

$$\beta = 1, \gamma = 1, \omega = 10^{-2}$$

Method	No. it.	Root-1	f(Root)
BNS1	3	$2.539918603315181\ 10^{-1}$	$-3.3307\ 10^{-16}$
BNS2	2	$2.539918603315182\ 10^{-1}$	$-1.1102\ 10^{-16}$
MW	3	$2.539918603315181\ 10^{-1}$	$-3.3307\ 10^{-16}$
GR	3	$2.539918603315181\ 10^{-1}$	$-3.3307\ 10^{-16}$

$$\beta = 10^{-2}, \gamma = 1, \omega = 1$$

Method	No. it.	Root−1	f(Root)
BNS1	2	$4.939569815595898\ 10^{-3}$	$7.8160\ 10^{-14}$
BNS2	2	$4.939569815595901\ 10^{-3}$	$1.4921\ 10^{-13}$
MW	2	$4.939569815595898\ 10^{-3}$	$7.8160\ 10^{-14}$
GR	2	$4.939569815595873\ 10^{-3}$	$-4.0501\ 10^{-13}$

J.R. Bunch, C.P. Nielsen and D.C. Sorensen, *Rank-one modification of the symmetric eigenproblem*, Numer. Math., v 31 (1978), pp 31–48

Ren-Cang Li, *Solving secular equations stably and efficiently*, Report UCB CSD-94-851, University of California, Berkeley (1994)

A. Melman, *A unifying convergence analysis of second-order methods for secular equation*, Math. Comput., v 66 n 217 (1997), pp 333–344

A. Melman, *A numerical comparison of methods for solving secular equations*, J. Comput. Appl. Math., v 86 (1997), pp 237–249