Chapter 3

Homotopy Method for λ-Matrix Problem

- Overview
- Basic ideas
- Preliminary facts
- Homotopy method
- Numerical experiment

Overview

- The problem:
  - ◇ Given $A_0, A_1, \ldots, A_k \in C^{n\times n}$, define
    $$P(\lambda) := A_k\lambda^k + A_{k-1}\lambda^{k-1} + \ldots + A_1\lambda + A_0.$$
  - ◇ Find $\lambda \in C$ and $x \in C^n$ such that
    $$P(\lambda)x = 0.$$
- Special cases:
  - ◇ Regular eigenvalue problem:
    $$\lambda x = Ax.$$
  - ◇ Generalized eigenvalue problem:
    $$\lambda Bx = Ax.$$

Numerical Methods

- Solving the linearized problem:
  - ◇ Can make use of existing software.
  - ◇ Increase the size considerably.
- Direct iteration:
  - ◇ Subspace iteration — global but slow
  - ◇ Newton-type iteration — fast but local
- Reducing to the canonical form:
  - ◇ Need polynomial root solver.
  - ◇ Ill-conditioned.
- Homotopy method:
  - ◇ Maybe costly in tracing curves.
  - ◇ Can follow curves simultaneously.
  - ◇ Guarantee to reach all isolated eigenpairs.
  - ◇ Matrix structure is respected.

Basic Ideas

- Homotopy equation:

$$H(x,t) = (1-t)g(x) + tf(x) = 0.$$

- Zero set:

$$H^{-1}(0) = \{(x,t) | H(x,t) = 0\}.$$

- Practical concerns:
  - ◇ Need to ensure $H^{-1}(0)$ is a 1-dimensional manifold.
  - ◇ Need to ensure the curve extens from $t = 0$ to $t = 1$.

- Curve tracing:

$$\begin{aligned} \begin{bmatrix} D_xH & D_tH \end{bmatrix} \begin{bmatrix} \frac{dx}{ds} \\ \frac{dt}{ds} \end{bmatrix} &= 0 \\ x(0) &= \text{zero(s) of } g(x) \\ t(0) &= 0. \end{aligned}$$

A Simple Example

- The problem:

$$\begin{aligned} \lambda x &= Ax \\ x^T x &= 1 \\ A &:= \text{real, symmetric and tridiagonal.} \end{aligned}$$

  ◇ A nonlinear (polynomial) system in $n+1$ unknowns x and λ.

  ◇ # of solutions $\leq$ Bezout number $= 2^{n+1}$.

- The homotopy:

$$\begin{aligned} H &: R^n \times R \times R \longrightarrow R^n \times R \\ H(x,\lambda,t) &:= \left([D + t(A-D)]x - \lambda x, \frac{1 - x^T x}{2} \right) \\ D &:= \text{diagonal with distinct elements.} \end{aligned}$$

- Existence of the curve:
  - ◇ Rank of

$$D_{(x,\lambda)} = \begin{bmatrix} D + t(A - D) - \lambda, & -x \\ -x^T, & 0 \end{bmatrix}$$

    is of rank $n + 1$.
  - ◇ Implicit function theorem $\Rightarrow$ (x, λ) is a function of t.
- Extension of the curve:
  - ◇ Gershgorin's theorem $\Rightarrow$ Boundedness of the curves.
  - ◇ Curves must extend to $t = 1$.

Preliminary Facts

- Regularity and canonical form
- Resultant theorem
- Perturbation theorem
- Rank property

Regularity and Canonical Form

- A_k is nonsingular $\Rightarrow$ $P(\lambda)$ is regular, i.e., has nk eigenvalues.
- $P(\lambda)$ is regular $\Rightarrow$ There exist $E(\lambda)$ and $F(\lambda)$ such that

$$E(\lambda)P(\lambda)F(\lambda) = \text{diag}(a_1(\lambda), \ldots, a_n(\lambda)).$$

  - ◇ $\det(E(\lambda)), \det(F(\lambda))$ = nonzero constants.
  - ◇ $a_i(\lambda)$ is a monic polynomial.
  - ◇ $a_i(\lambda)$ is unique.
  - ◇ $a_i(\lambda) | a_{i+1}(\lambda)$.

Resultant Theorem

- Resultant of two polynomials:

$$\begin{aligned} R &= R(a_0, \ldots, a_n, b_0, \ldots, b_m) \\ &= \det \begin{bmatrix} a_0 & a_1 & & \ldots & a_n & 0 & \ldots & 0 \\ 0 & a_0 & a_1 & & & a_n & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & & \ddots & \\ & & & & a_1 & & & \\ 0 & \ldots & & & a_0 & a_1 & \ldots & a_n \\ b_0 & b_1 & \ldots & & 0 & \ldots & & \\ 0 & b_0 & b_1 & \ldots & & & & 0 \\ \vdots & \ddots & \ddots & & \vdots & & & \\ 0 & \ldots & & b_0 & b_1 & \ldots & & b_m \end{bmatrix}. \end{aligned}$$

  - ◇ $f(x) = a_0 + a_1 x + \ldots\ldots + a_n x^n$.
  - ◇ $g(x) = b_0 + b_1 x + \ldots + b_m x^m$.

- f and g have common non-constant factor $\Leftrightarrow R = 0$.

Perturbation Theorem

- There exist real numbers $d_1, \ldots, d_n$ such that $p(\lambda) := \det(P(\lambda) - D)$ has no multiple roots.
  - ⋄ $D := \mathrm{diag}(d_1, \ldots, d_n)$.
- The set

  $$\{(d_1, \ldots, d_n) | P(\lambda) - D \text{ has multiple eigenvalues}\}$$

  is of complex codimension 1.
- With probability one the matrix $P(\lambda) - D$ has distinct eigenvalues.

Rank Property

- $P(\lambda)$ has nk distinct eigenvalues and $P(\lambda_j)x_j = 0$ $\Rightarrow$

$$Q(x_j, \lambda_j) := [P(\lambda_j), P'(\lambda_j)x_j]$$

  is of complex rank n.

- Identify a linear transformation $C^{n+1} \longrightarrow C^n$ as a transformation $R^{2n+2} \longrightarrow R^{2n}$:
  - ◇ Replace each component, say $x_j = a_j + ib_j$, of the vector by $[a_1, b_1]^T$.
  - ◇ Replace each component, say $z = a + ib$, of the transformation matrix by $\begin{bmatrix} a, & -b \\ b, & a \end{bmatrix}$.

$$Q \in C^{n\times(n+1)} \hookrightarrow \hat{Q} \in R^{2n\times(2n+2)}.$$

- If $P(\lambda)x = 0$, then

$$M(x,\lambda) := \begin{bmatrix} \hat{Q}(x_j, \lambda_j) \\ a_1, b_1, a_2, \ldots, a_n, b_n, 0, 0 \end{bmatrix}$$

  is of real rank $2n+1$ in $R^{(2n+1)\times(2n+2)}$.

Homotopy Method

- Notation:

$$\begin{aligned} P(\lambda) &= A_k\lambda^k + A_{k-1}\lambda^{k-1} + \ldots + A_0 \\ Q(\lambda) &= cI\lambda^k - D \\ R(\lambda, t, c, D) &= (1-t)Q(\lambda) + tP(\lambda) \\ D &= \text{diag}(d_1, \ldots, d_n) \\ c, d_1, \ldots, d_n &= \text{ complex numbers} \end{aligned}$$

- Control of regularity:

$$R(\lambda, t, c, D) = [(1-t)cI + tA_k]\lambda^k + \ldots$$

  ◇ Choose d from an open dense set such that $[(1-t)cI + tA_k]$ is nonsingular for $0 \leq t < 1$.

- The homotopy:

$$\begin{aligned} H &: C^n \times C \times [0,1) \longrightarrow C^n \times C \\ H(x,\lambda,t) &:= \begin{bmatrix} R(\lambda,t,c,D)x \\ \frac{x^*x-1}{2} \end{bmatrix}. \end{aligned}$$

- Initial values:

$$\begin{aligned} e_i &= \text{The standard } i\text{-th unit vector}, \\ \lambda_{ij} &= \text{The } j\text{-th complex root of } (\frac{d_i}{d})^{1/k}. \end{aligned}$$

◇ $i = 1, \ldots, n$ and $j = 1, \ldots, k$.

◇ $H(e_i, \lambda_{ij}, 0) = 0$.

Major Theorem

There exists an open, dense, full measure subset $U \subset C^n$ such that, for $(d_1, \ldots, d_n) \in U$ and each initial point $y_{ij} := (e_i, \lambda_{ij}, 0)$, the connected component $C(y_{ij})$ of y_{ij} in $H^{-1}(0)$, when identified as a subset in $R^{2n} \times R^2 \times R$, has the following properties:

- $C(y_{ij})$ is a real analytic submanifold with real dimension 2.
- The intersection of $C(y_{ij})$ with each hyperplane $t \equiv$ constant $\in [0, 1)$ is a unit circle centered $(0, \lambda) \in R^{2n} \times R$ for some λ.
- Manifolds $C(y_{ij})$ corresponding to different initial points do not intersect for $t \in [0, 1)$.
- Each manifold $C(y_{ij})$ is bounded for $t \in [0, 1)$.

Computation

- It is a tube:

$$H : R^{2n} \times R^2 \times R \longrightarrow R^{2n} \times R.$$

  - ◇ $C(y_{ij})$ is a 2-dimensional tube with unit radius at each cross-section $t \equiv$ constant.

- Need a vector field on the tube:
  - ◇ The path can be parametrized by the variable t.
  - ◇ Define vector field $(\dot{x}, \dot{\lambda}, 1)$ on $H^{-1}(0)$ by

$$M(x, \lambda, t) \begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} (Q(\lambda) - P(\lambda))x \\ 0 \end{bmatrix},$$

$$\left[ix^T, 0\right] \begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} = 0.$$

    - ▷ First equation is necessary for being tangent.
    - ▷ Second equation means the vector field is perpendicular to the circle of intersection.

- Observe:

$$\begin{bmatrix} M(x,\lambda,t) \\ ix^T, 0 \end{bmatrix} = \begin{bmatrix} & \hat{Q}(x,\lambda,t) & & & & \\ a_1, & b_1, \ldots, a_n, & b_n, & 0, & 0 \\ -b_1, & a_1, \ldots, -b_n, & a_n, & 0, & 0 \end{bmatrix}$$
$$\hookrightarrow \begin{bmatrix} R(\lambda,t,c,D), & R'(\lambda,t,c,D)x \\ x^*, & 0 \end{bmatrix}.$$

- The initial value problem in $C^n \times C$:

$$\begin{bmatrix} R(\lambda,t,c,D), & R'(\lambda,t,c,D)x \\ x^*, & 0 \end{bmatrix} \begin{bmatrix} \frac{dx}{dt} \\ \frac{d\lambda}{dt} \end{bmatrix} = \begin{bmatrix} (Q(\lambda) - P(\lambda))x \\ 0 \end{bmatrix},$$
$$x(0) = e_i$$
$$\lambda(0) = \lambda_{ij}$$

  - ◇ $i = 1, \ldots, n$ and $j = 1, \ldots, k$.

- Parallel computation:
  - ◇ Homotopy curves are independent of each other.
  - ◇ Can integrate simultaneously.

- Sparsity preservation: $R(\lambda, t, c, D)$ does not destroy the matrix structure of the coefficients in $P(\lambda)$.