Lecture 10: Miscellaneous Applications

1. Linear integral operators

Let $L^2 = L^2[a,b]$, the **Lebesgue square integrable functions** on the finite interval $[a,b]$. Let $K(s,t)$ be an L^2–**kernel** on $a \le s,t, \le b$,

$$\text{i.e., } \int_a^b \int_a^b |K(s,t)|^2 ds\, dt \quad \text{exists and is finite.}$$

Consider the two operators $T_1, T_2 \in \mathcal{B}(L^2, L^2)$ defined by

$$(T_1\mathbf{x})(s) = \int_a^b K(s,t)\mathbf{x}(t)dt\ , \quad a \le s \le b\ ,$$

$$(T_2\mathbf{x})(s) = \mathbf{x}(s) - \int_a^b K(s,t)\mathbf{x}(t)dt\ , \quad a \le s \le b\ ,$$

called **Fredholm integral operators of the first kind** and **the second kind**, respectively. Then

(a) $R(T_2)$ is **closed**,

(b) $R(T_1)$ is **nonclosed** unless it is finite dimensional.

The Fredholm integral equation of the 2nd kind

$$x(s) - \lambda \int_a^b K(s,t)\, x(t)\, dt = y(s)\ , \quad a \le s \le b\ , \tag{1}$$

is also written as

$$(I - \lambda K)\, \mathbf{x} = \mathbf{y}\ ,$$

where λ and all functions are complex, and $[a, b]$ is a bounded interval.

We need the following facts from the Fredholm theory of integral equations. For any λ, K as above

(a) $(I - \lambda K) \in \mathcal{B}(L^2, L^2)$,

(b) $(I - \lambda K)^* = I - \overline{\lambda} K^*$, where $K^*(s,t) = \overline{K(t,s)}$.

(c) The null spaces $N(I - \lambda K)$ and $N(I - \overline{\lambda} K^*)$ have equal finite dimensions,

$$\dim\, N(I - \lambda K) = \dim\, N(I - \overline{\lambda} K^*) = n(\lambda)\ ,\ \text{say}\ . \tag{2}$$

Fredholm (cont'd)

(d) A scalar λ is called a **regular value** of K if $n(\lambda) = 0$, in which case the operator $I - \lambda K$ has an **inverse** $(I - \lambda K)^{-1} \in \mathcal{B}(L^2, L^2)$ written as

$$(I - \lambda K)^{-1} = I + \lambda R \,, \tag{3}$$

where $R = R(s, t; \lambda)$ is an L^2–kernel called the **resolvent** of K.

(e) A scalar λ is called an **eigenvalue** of K if $n(\lambda) > 0$, in which case any nonzero $\mathbf{x} \in N(I - \lambda K)$ is called an **eigenfunction** of K corresponding to λ.

For any λ and, in particular, for any eigenvalue λ, both range spaces $R(I - \lambda K)$ and $R(I - \overline{\lambda} K^*)$ are closed and,

$$R(I - \lambda K) = N(I - \overline{\lambda} K^*)^{\perp} \,, \quad R(I - \overline{\lambda} K^*) = N(I - \lambda K)^{\perp} \,. \tag{4}$$

Fredholm (cont'd)

(f) If λ is a **regular value** of K then (1) has, for any $\mathbf{y} \in L^2$, a unique solution given by

$$\mathbf{x} = (I + \lambda R)\, \mathbf{y} \ ,$$

or,

$$x(s) = y(s) + \lambda \int_a^b R(s,t,\lambda)\, y(t)\, dt \ , \ \ a \le s \le b \ . \tag{5}$$

(g) If λ is an **eigenvalue** of K then (1) is consistent if and only if $\mathbf{y}$ is **orthogonal** to every $\mathbf{u} \in N(I - \overline{\lambda} K^*)$, in which case the general solution of (1) is

$$\mathbf{x} = \mathbf{x}_0 + \sum_{i=1}^{n(\lambda)} c_i \mathbf{x}_i \ , \quad c_i \text{ arbitrary scalars} \ ,$$

$\mathbf{x}_0$ a **particular solution**, $\{\mathbf{x}_1, \ldots, \mathbf{x}_{n(\lambda)}\}$ a **basis** of $N(I - \lambda K)$.

Pseudo resolvents

Let λ be an **eigenvalue** of K. Following **Hurwitz**, an L^2–kernel $R = R(s,t,\lambda)$ is called a **pseudo resolvent** of K if for any $\mathbf{y} \in R(I - \lambda K)$, the function

$$x(s) = y(s) + \lambda \int_a^b R(s,t,\lambda)\, y(t)\, dt \tag{5}$$

is a solution of (1).

Hurwitz constructed a **pseudo resolvent** as follows.

Let λ_0 be an **eigenvalue** of K, and let $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ and $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be **o.n. bases** of $N(I - \lambda_0 K)$ and $N(I - \overline{\lambda_0} K^*)$ respectively. Then λ_0 is a **regular value** of the kernel

$$K_0(s,t) = K(s,t) - \frac{1}{\lambda_0} \sum_{i=1}^{n} u_i(s)\, \overline{x_i(t)}\ , \tag{6}$$

Pseudo resolvents (cont'd)

The eigenvalue λ_0 is a regular value of

$$K_0(s,t) = K(s,t) - \frac{1}{\lambda_0} \sum_{i=1}^{n} u_i(s)\, \overline{x_i(t)} \ , \tag{6}$$

written for short as $\quad K_0 = K - \frac{1}{\lambda_0} \sum_{i=1}^{n} \mathbf{u}_i \mathbf{x}_i^*$

and the resolvent R_0 of K_0 is a pseudo resolvent of K, satisfying

$$\begin{aligned}
(I + \lambda_0 R_0)(I - \lambda_0 K)\, \mathbf{x} = \mathbf{x} \ , &\quad \text{for all } \mathbf{x} \in R(I - \overline{\lambda_0} K^*) \\
(I - \lambda_0 K)(I + \lambda_0 R_0)\, \mathbf{y} = \mathbf{y} \ , &\quad \text{for all } \mathbf{y} \in R(I - \lambda_0 K) \\
(I + \lambda_0 R_0)\, \mathbf{u}_i = \mathbf{x}_i \ , &\quad i = 1, \ldots, n \ .
\end{aligned} \tag{7}$$

If R is a pseudo resolvent of K, then $I + \lambda R$ is a $\{1\}$–**inverse** of $I - \lambda K$. As with $\{1\}$–inverses, the pseudo resolvent is **not unique**.

Characterization of pseudo resolvents

The pseudo resolvent is not unique: For $R_0, \mathbf{u}_i, \mathbf{x}_i$ as above, and any scalars c_{ij}, the kernel $R_0 + \sum_{i,j,=1}^n c_{ij}\mathbf{x}_i\mathbf{u}_j^*$ is a pseudo resolvent of K.

Theorem (**Hurwitz**). Let K be an L^2–kernel, λ_0 be an eigenvalue of K and $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ and $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be orthonormal bases of $N(I - \lambda_0 K)$ and $N(I - \overline{\lambda_0}K^*)$ respectively. An L^2–kernel R is a pseudo resolvent of K if and only if

$$R = K + \lambda_0 KR - \frac{1}{\lambda_0}\sum_{i=1}^n \boldsymbol{\beta}_i \mathbf{u}_i^* \;, \tag{8a}$$

$$R = K + \lambda_0 RK - \frac{1}{\lambda_0}\sum_{i=1}^n \mathbf{x}_i \boldsymbol{\alpha}_i^* \;, \tag{8b}$$

where $\boldsymbol{\alpha}_i$, $\boldsymbol{\beta}_i \in L^2$ satisfy

$$\langle \boldsymbol{\alpha}_i, \mathbf{x}_j \rangle = \delta_{ij} \;,\; \langle \boldsymbol{\beta}_i, \mathbf{u}_j \rangle = \delta_{ij} \;, \quad i,j = 1, \ldots, n \;. \tag{9}$$

Characterization (cont'd)

Here KR stands for the kernel

$$KR(s,t) = \int_a^b K(s,u)R(u,t)\,du$$

If λ is a regular value of K then (8a)–(8b) reduce to

$$R = K + \lambda KR\ ,\ R = K + \lambda RK\ ,$$

which uniquely determines the resolvent $R(s,t,\lambda)$.

Degenerate kernels

A kernel $K(s, t)$ is called **degenerate** if it is a finite sum of products of L^2 functions, as follows:

$$K(s,t) = \sum_{i=1}^{m} f_i(s)\,\overline{g_i(t)}\ . \tag{10}$$

Degenerate kernels are convenient because they reduce the integral equation (1) to a finite system of linear equations. Also, any L^2–kernel can be approximated, arbitrarily close, by a degenerate kernel.

Let $K(s, t)$ be given by (10). Then

(a) The scalar λ is an eigenvalue of (10) if and only if $1/\lambda$ is an eigenvalue of the $m \times m$ matrix

$$B = [b_{ij}]\ , \quad \text{where } b_{ij} = \int_a^b f_j(s)\,\overline{g_i(s)}\,ds\ .$$

Degenerate kernels (cont'd)

(b) Any eigenfunction of K $[K^*]$ corresponding to an eigenvalue λ $[\overline{\lambda}]$ is a linear combination of the m functions $f_1, \ldots, f_m$ $[g_1, \ldots, g_m]$.

(c) If λ is a regular value of (10), then the resolvent at λ is

$$R(s,t,;\lambda) = \frac{\det \begin{bmatrix} 0 & \vdots & f_1(s) & \cdots & f_m(s) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -\overline{g_1(t)} & \vdots & & & \\ \vdots & \vdots & & I - \lambda B & \\ -\overline{g_m(t)} & \vdots & & & \end{bmatrix}}{\det(I - \lambda B)} .$$

Example

Consider the equation

$$x(s) - \lambda \int_{-1}^{1} (1 + 3st)\, x(t)\, dt = y(s) \tag{11}$$

with $K(s,t) = 1 + 3st$. The resolvent is

$$R(s,t;\lambda) = \frac{1+3st}{1-2\lambda} \; .$$

K has a single eigenvalue $\lambda = \frac{1}{2}$ and an o.n. basis of $N(I - \frac{1}{2}K)$ is

$$\left\{ x_1(s) = \frac{1}{\sqrt{2}},\, x_2(s) = \frac{\sqrt{3}}{\sqrt{2}}\, s \right\}$$

which, by symmetry, is also an orthonormal basis of $N(I - \frac{1}{2}K^*)$.

Example (cont'd)

From (6) we get

$$\begin{aligned} K_0(s,t) &= K(s,t) - \frac{1}{\lambda_0}\sum u_i(s)\,\overline{x_i(t)} \\ &= (1+3st) - 2\left(\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} + \frac{\sqrt{3}}{\sqrt{2}}s\frac{\sqrt{3}}{\sqrt{2}}t\right) \\ &= 0\ , \end{aligned}$$

and the resolvent of $K_0(s,t)$ is therefore

$$R_0(s,t;\lambda) = 0\ .$$

(a) If $\lambda \neq \frac{1}{2}$, then for each $y \in L^2[-1,1]$ equation (11) has a unique solution,

$$x(s) = y(s) + \lambda \int_{-1}^{1} \frac{1+3st}{1-2\lambda} y(t)\,dt\ .$$

Example (cont'd)

(b) If $\lambda = \frac{1}{2}$, then (11) is consistent if and only if

$$\int_{-1}^{1} y(t)\,dt = 0 \ , \quad \int_{-1}^{1} t\,y(t)\,dt = 0 \ ,$$

in which case the general solution is

$$x(s) = y(s) + c_1 + c_2 s \ , \quad c_1, c_2 \text{ arbitrary} \ .$$

2. Linear systems theory

Systems modeled by linear differential equations call for symbolic computation of generalized inverses for matrices whose elements are rational functions.

As example, consider the homogeneous system

$$A(D)\mathbf{x}(t) = \mathbf{0} \tag{1}$$

where $\mathbf{x}(t) : [0-, \infty) \to \mathbb{R}^n$, $D := \dfrac{d}{dt}$,

$$A(D) = A_q D^q + \cdots + A_1 D + A_0 \ , \tag{2}$$

and $A_i \in \mathbb{R}^{m\times n}$, $i = 0, 1, \ldots, q$. Let $\mathcal{L}$ denote the **Laplace transform**, and let $\widehat{\mathbf{x}}(s) = \mathcal{L}(\mathbf{x}(t))$. The system (1) transforms to

$$A(s)\widehat{x}(s) = \widehat{\mathbf{b}}(s) \ ,$$

allowing algebraic solution.

Linear systems theory (cont'd)

Theorem (**Jones, Karampetakis and Pugh**) The system (1) has a solution if and only if

$$A(s)A(s)^{\dagger}\widehat{\mathbf{b}}(s) = \widehat{\mathbf{b}}(s) \tag{3}$$

in which case the general solution is

$$\mathbf{x}(t) = \mathcal{L}^{-1}(\widehat{\mathbf{x}}(s)) = \mathcal{L}^{-1}\left\{A(s)^{\dagger}\widehat{\mathbf{b}}(s) + (I_n - A(s)^{\dagger}A(s))\mathbf{y}(s)\right\} \tag{4}$$

where $\mathbf{y}(s) \in \mathbb{R}^n(s)$ is arbitrary. □

3. Tchebycheff approximation

A **Tchebycheff approximate solution** of the system

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

is a vector $\mathbf{x}$ minimizing the **Tchebycheff norm**

$$\|\mathbf{r}\|_\infty = \max_{i=1,\dots,m} \{|r_i|\}$$

of the residual vector

$$\mathbf{r} = \mathbf{b} - A\mathbf{x} \ . \tag{2}$$

Let $A \in \mathbb{C}_n^{(n+1)\times n}$ and $\mathbf{b} \in \mathbb{C}^{n+1}$ be such that (1) is inconsistent. Then (1) has a unique Tchebycheff approximate solution given by

$$\mathbf{x} = A^\dagger(\mathbf{b} + \mathbf{r}) \ , \tag{3}$$

Tchebycheff approximation (cont'd)

where the **residual** $\mathbf{r} = [r_i]$ is

$$r_i = \frac{\sum_{j=1}^{n+1} |(P_{N(A^*)}\mathbf{b})_j|^2}{\sum_{j=1}^{n+1} |(P_{N(A^*)}\mathbf{b})_j|} \frac{(P_{N(A^*)}\mathbf{b})_i}{|(P_{N(A^*)}\mathbf{b})_i|} \, , \quad i \in \overline{1, n+1} \, . \tag{4}$$

Proof. From

$$\mathbf{r}(\mathbf{x}) - \mathbf{b} = -A\mathbf{x} \in R(A)$$

it follows that any residual $\mathbf{r}$ satisfies

$$P_{N(A^*)}\mathbf{r} = P_{N(A^*)}\mathbf{b}$$

or equivalently

$$\langle P_{N(A^*)}\mathbf{b}, \mathbf{r} \rangle = \langle \mathbf{b}, P_{N(A^*)}\mathbf{b} \rangle \, , \tag{5}$$

since $\dim N(A^*) = 1$ and $\mathbf{b} \notin R(A)$.

Tchebycheff approximation (cont'd)

Equation (5) represents the **hyperplane of residuals**. A routine computation now shows, that among all residuals $\mathbf{r}$ satisfying (5) there is **a unique residual of minimum Tchebycheff norm** given by (4), from which (3) follows since $N(A) = \{\mathbf{0}\}$.

4. Interval linear programming

For two vectors $\mathbf{u} = (u_i), \mathbf{v} = (v_i) \in \mathbb{R}^m$ let

$$\mathbf{u} \leq \mathbf{v}$$

denote the fact that $u_i \leq v_i$ for $i \in \overline{1, m}$. A linear programming problem of the form

$$\text{maximize} \left\{ \mathbf{c}^T \mathbf{x} : \; \mathbf{a} \leq A\mathbf{x} \leq \mathbf{b} \right\} , \qquad (1)$$

with given $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$; $\mathbf{c} \in \mathbb{R}^n$; $A \in \mathbb{R}^{m \times n}$, is called an **interval linear program** and denoted by $IP(\mathbf{a}, \mathbf{b}, \mathbf{c}, A)$ or simply by IP.

The IP (1) is **consistent** (also **feasible**) if the set

$$F = \{ \mathbf{x} \in \mathbb{R}^n : \; \mathbf{a} \leq A\mathbf{x} \leq \mathbf{b} \} \neq \emptyset \qquad (2)$$

in which case the elements of F are the **feasible solutions** of $IP(\mathbf{a}, \mathbf{b}, \mathbf{c}, A)$.

Interval linear programming (cont'd)

A consistent $IP(\mathbf{a}, \mathbf{b}, \mathbf{c}, A)$ is **bounded** if

$$\max\{\mathbf{c}^T\mathbf{x} : x \in F\}$$

is **finite**, in which case the **optimal solutions** of $IP(\mathbf{a}, \mathbf{b}, \mathbf{c}, A)$ are its feasible solutions $\mathbf{x}_0$ which satisfy

$$\mathbf{c}^T\mathbf{x}_0 = \max\{\mathbf{c}^T\mathbf{x} : \mathbf{x} \in F\} .$$

Lemma. Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$; $\mathbf{c} \in \mathbb{R}^n$; $A \in \mathbb{R}^{m\times n}$ be such that $IP(\mathbf{a}, \mathbf{b}, \mathbf{c}, A)$ is consistent. Then $IP(\mathbf{a}, \mathbf{b}, \mathbf{c}, A)$ is bounded if and only if

$$\mathbf{c} \in N(A)^\perp . \tag{3}$$

Proof. $F = F + N(A)$, etc. □

Interval linear programming (cont'd)

Let $\eta : \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$ be defined for $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^m$ by

$$\eta_i = \begin{cases} u_i & \text{if } w_i < 0, \\ v_i & \text{if } w_i > 0, \\ \lambda_i u_i + (1-\lambda_i) v_i & \text{where } 0 \le \lambda_i \le 1\ ,\ \text{if } w_i = 0 \end{cases} \tag{4}$$

Theorem. Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$; $\mathbf{c} \in \mathbb{R}^n$; $A \in \mathbb{R}_m^{m \times n}$ (**full row-rank**) be such that $IP(\mathbf{a}, \mathbf{b}, \mathbf{c}, A)$ is consistent and bounded, and let $A^{(1)}$ be any $\{1\}$–inverse of A. Then the general optimal solution of $IP(\mathbf{a}, \mathbf{b}, \mathbf{c}, A)$ is

$$\mathbf{x} = A^{(1)} \eta(\mathbf{a}, \mathbf{b}, A^{(1)T}\mathbf{c}) + \mathbf{y}\ , \quad \mathbf{y} \in N(A)\ . \tag{5}$$

Proof. For $\mathbf{u} = A\mathbf{x}$, the problem (1) is

$$\max\{\mathbf{c}^T A^{(1)} \mathbf{u} :\ \mathbf{a} \le \mathbf{u} \le \mathbf{b}\}\ ,\ \text{etc.}$$

Note: The rank assumption is a severe restriction of usefulness.

5. Nonlinear least squares solutions

Let $\mathbf{f} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, and let

$$J_{\mathbf{f}}(\mathbf{x}) = \left(\frac{\partial f_i(\mathbf{x})}{\partial x_j} \right) .$$

If the **Newton method**

$$\mathbf{x}_+ := \mathbf{x} - J_{\mathbf{f}}(\mathbf{x})^\dagger \, \mathbf{f}(\mathbf{x})$$

converges to $\mathbf{x}_\infty$, plus 2 more **if**'s, then

$$J_{\mathbf{f}}(\mathbf{x}_\infty)^\dagger \, \mathbf{f}(\mathbf{x}_\infty) = \mathbf{0}$$

and x_∞ is a **stationary point** of $\|\mathbf{f}(\mathbf{x})\|^2$.

A **Maple** code for a **Newton method** using the Moore–Penrose inverse of the Jacobi matrix is available, contact the instructor or see **http://benisrael.net/Newton-MP.pdf**