Linear Inverse Problems with discrete data

Ref.: M. Bertero, C. deMol and E.R. Pike

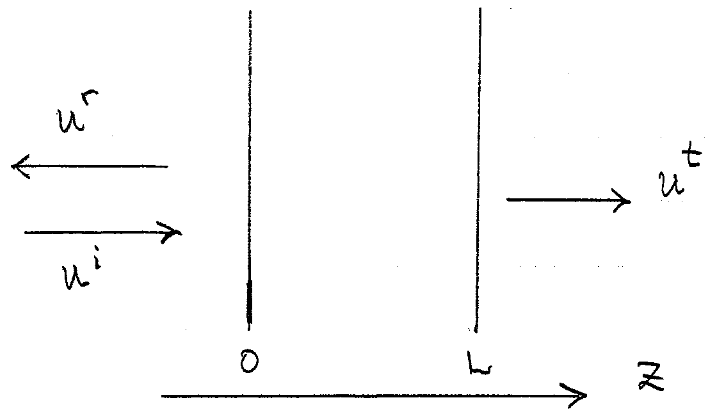
Linear Inverse Problems with discrete data

Inverse Problems 1, 301-330 (1985)

Inverse Problems 4, 573-594 (1988)

Table of contents

1. Example



$u^i(t)$ incident pulse
$u^r(t)$ reflected pulse
$u^t(t)$ transmitted pulse

Relation between u^i and u^r, u^t for a linear response

$$\begin{cases} u^r(t) = \int_0^t R(t-t')\, u^i(t')\, dt' \\ u^t(t) = a\left[u^i(t) + \int_0^t T(t-t')\, u^i(t')\, dt' \right] \end{cases}$$

$R(t)$ reflection kernel, impulse response
$T(t)$ transmission kernel
a constant related to phase shift and attenuation through the slab.

<u>Measure:</u> u^r, u^i and u^t

<u>Find:</u> $R(t)$, $T(t)$ (and a)

2. Fourier method

As a generic case take

$$u^r(t) = \int_0^t R(t-t')\, u^i(t')\, dt'$$

Fourier transform (denoted by carat ^)

$$\hat{u}^r(\omega) = \hat{R}(\omega)\, \hat{u}^i(\omega)$$

Solution: $\boxed{\hat{R}(\omega) = \dfrac{\hat{u}^r(\omega)}{\hat{u}^i(\omega)}}$

Problems: i) at discontinuities: Gibbs phenomenon

ii) amplification of noise where $\hat{u}^i(\omega) \approx 0$

iii) ...

iv) ...

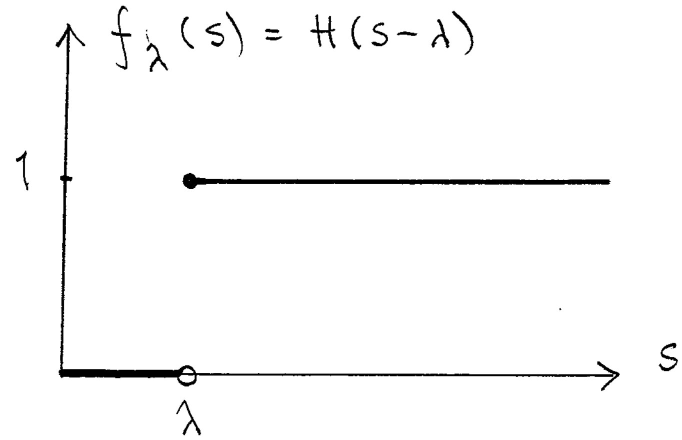
To avoid amplification of noise filter techniques have to be used.

Define $$\hat{R}_\lambda(\omega) := f_\lambda(|\hat{u}^i(\omega)|)\,\hat{R}(\omega)$$

$$= f_\lambda(|\hat{u}^i(\omega)|)\,\frac{\hat{u}^r(\omega)}{\hat{u}^i(\omega)}$$

$f_\lambda(s)$ is the filter function

Example of $f_\lambda(s)$:



3. Singular Value Decomposition method

This technique resides in the time domain

- Assume measurements of

 the reflected pulse at $t = t_n$, $n = 1, \dots, N$

 the incident pulse at $t = t_m$, $m = 1, \dots, M$

- Assume quadrature weights w_m, $m = 1, \dots, M$

Discretize and approximate

$$u^r(t) = \int_0^t R(t-t')\, u^i(t')\, dt' = \int_0^t u^i(t-t')\, R(t')\, dt'$$

$$\boxed{g_n = \sum_{m=1}^{M} L_{nm} f_m \qquad , n = 1, \dots, N}$$

where
$$\begin{cases} g_n := u^r(t_n) & , n = 1, \dots, N \\ L_{nm} := u^i(t_n - t_m)\, w_m & , \begin{cases} n = 1, \dots, N \\ m = 1, \dots, M \end{cases} \\ f_m := R(t_m) & , m = 1, \dots, M \end{cases}$$

Note! f_m, g_n and L_{nm} are real

SVD decomposition

Any real $N \times M$ matrix L_{nm} can be decomposed as

$$\boxed{L = U \Sigma V^T}$$

where $U := [\bar{u}_1, \dots, \bar{u}_N]$ $N \times N$ orthogonal matrix

$V := [\bar{v}_1, \dots, \bar{v}_M]$ $M \times M$ orthogonal matrix

$\Sigma := \operatorname{diag}(\sigma_i)$, $i = 1, \dots, p = \min(N, M)$

diagonal real $N \times M$ matrix of the form

$$N > M \quad \begin{pmatrix} \ddots & 0 \\ 0 & \end{pmatrix} \qquad \text{or} \qquad N < M \quad \begin{pmatrix} \ddots & & 0 \\ 0 & & \end{pmatrix}$$

and the positive real numbers σ_i are ordered:

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_K \geq 0$$

———— o ————

The real vectors $\bar{u}_i$ and $\bar{v}_j$ satisfy

$$\begin{cases} L \bar{v}_j = \sigma_j \bar{u}_j \\ L^T \bar{u}_j = \sigma_j \bar{v}_j \end{cases} \qquad j = 1, \dots, p = \min(N, M)$$

σ_i are called the singular values of L

$\bar{u}_i$ are called the left singular vectors of L

$\bar{v}_j$ — " — right — " —

- The number of singular values $\sigma_i > 0$ is the rank r of the matrix, i.e.

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > \sigma_{r+1} = \ldots = \sigma_p = 0$$

- The null space of L is the span of the vectors $\bar{v}_{r+1}, \bar{v}_{r+2}, \ldots, \bar{v}_M$
- The range of L is the span of the vectors $\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_r$

<u>Summary</u>

$$\boxed{\begin{aligned} &\text{rank}(L) = r \\ &N(L) = \text{span}\{\bar{v}_{r+1}, \ldots, \bar{v}_M\} \\ &R(L) = \text{span}\{\bar{u}_1, \ldots, \bar{u}_r\} \\ &L = \sum_{i=1}^{r} \bar{u}_i \sigma_i \bar{v}_i^T \end{aligned}}$$

Apply the SVD to our equation

$$g_n = \sum_{m=1}^{M} L_{nm} f_m$$

$$\bar{g} = L\bar{f} = U\Sigma V^T \bar{f} = \sum_{i=1}^{r} \bar{u}_i \sigma_i \underbrace{\bar{v}_i^T \bar{f}}$$

$= (\ ,\ )(\ ) = (\bar{v}_i, \bar{f})$

= scalar product i $\mathbb{R}^M$

or

$$(\bar{u}_i, \bar{g}) = \sigma_i (\bar{v}_i, \bar{f}) \qquad \text{since } \bar{u}_i \text{ i ON-bas i } \mathbb{R}^N$$

Define

$$\Sigma^+ := \text{diag}(\sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_r^{-1}, 0, \dots, 0)$$

a diagonal real $M \times N$ matrix

$\Sigma^+\Sigma = 1$ (unit matrix in $\mathbb{R}^M$)

$\Sigma\Sigma^+ = 1$ (— " — $\mathbb{R}^N$)

Define the generalized inverse L^+ to L

(Moore – Penrose inverse)

$$L^+ := V \Sigma^+ U^T$$

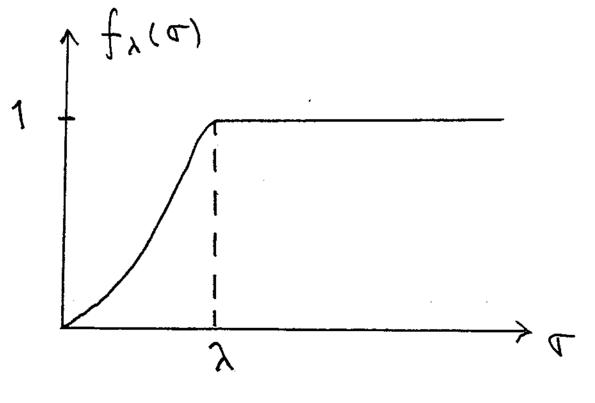
and the generalized solution $\bar{f}^+ := L^+ \bar{g}$

$$\bar{f}^+ = V \Sigma^+ U^T \bar{g} = \sum_{i=1}^{r} \frac{(\bar{u}_i, \bar{g})}{\sigma_i} \bar{v}_i$$

In practical situations introduce a filter function $f_\lambda(\sigma)$ to stabilize noise corrupted data.

$$\bar{f}_\lambda^+ = \sum_{i=1}^{r} f_\lambda(\sigma_i) \frac{(\bar{u}_i, \bar{g})}{\sigma_i} \bar{v}_i$$

Example of $f_\lambda(\sigma)$:



How to pick λ ?

(Some ways of picking λ will be introduced later)

4. Linear inverse problem (discrete data) (deeper look)

4.1 Formulation of the problem

The general linear problem we want to solve is

$$g_n = \int K_n(y)\, f(y)\, dy \quad , \quad i = 1, \dots, N$$

<u>Given:</u> i) the real (or complex) numbers $\{g_n\}_{n=1}^{N}$.

ii) the functions $\{K_n(y)\}_{n=1}^{N}$.

<u>Find:</u> The function $f(y)$ in some function space

Or more generally

<u>Given:</u> i) a class of functions $\underline{X}$ (function space)

ii) a set $\{F_n\}_{n=1}^{N}$ of linear functionals on $\underline{X}$

iii) a set of numbers $\{g_n\}_{n=1}^{N}$

<u>Find:</u> an object f in $\underline{X}$ such that

$$F_n(f) = g_n \quad , \quad i = 1, \dots, N$$

<u>Comments</u>: i) This formulation is more general than the one we studied in section 3, where also the space $\underline{X}$ was finite dimensional

ii) No attention is paid to the effects of noise at the moment.

To simplify the analysis we will assume $\underline{X}$ to be a Hilbert space, so that an inner product is defined. $\underline{X}$ is also called <u>the object space</u>.

Riesz representation theorem (N & S. p. 345, Th. 5.21.1) implies that there exist unique elements $\{\phi_n\}_{n=1}^{N}$ in $\underline{X}$ such that

$$F_n(f) = (f, \phi_n)_{\underline{X}} \quad , \forall f \in \underline{X} \; , \; n = 1, \ldots, N$$

A third formulation of the problem would then be

> **Given:** i) Hilbert space $\underline{X}$
>
> ii) a set of functions $\{\phi_n\}_{n=1}^{N}$ in $\underline{X}$
>
> iii) a set of numbers $\{g_n\}_{n=1}^{N}$
>
> **Find:** a function $f \in \underline{X}$ such that
>
> $$(f, \phi_n)_{\underline{X}} = g_n \quad , n=1,\dots,N$$

We call this problem the LIP problem

Problems: i) The information about $f \in \underline{X}$ is incomplete, since only a finite number of components of f is known. This implies lack of uniqueness.

ii) When the number of data is large ($N \to \infty$) lack of numerical stability is possible.

The origin of the lack of numerical stability derives from the fact that in most cases our problem is a projection on a finite-dimensional space of an ill-posed problem.

4.2 The normal solution

Assume that the functions $\{\phi_n\}_{n=1}^{N}$ are linearly independent (not necessarily orthogonal).

Denote $\underline{X}_N = \text{span}\{\phi_n\}$

The subspace $\underline{X}_N$ has dimension N

Lemma: With the assumption on $\{\phi_n\}_{n=1}^{N}$ above the mapping $F: \underline{X} \to R^N$ defined by

$$F(f) = (f, \phi_n)_{\underline{X}} \quad \text{is onto } R^N$$

Proof: Suppose the mapping is not onto
The range of F is then a subspace of R^N and there exists a vector $R^N \ni \bar{c} \neq \bar{0}$:

$$c_1 (f, \phi_1)_{\underline{X}} + \dots + c_N (f, \phi_N)_{\underline{X}} = 0 \quad , \quad \forall f \in \underline{X}$$

or

$$(f, \bar{c}_1 \phi_1 + \dots + \bar{c}_N \phi_N)_{\underline{X}} = 0 \quad , \quad \forall f \in \underline{X}$$

$$\Rightarrow \quad \bar{c}_1 \phi_1 + \dots + \bar{c}_N \phi_N = 0$$

Contradiction to the linear independence of $\{\phi_n\}_{n=1}^{N}$.

QED

- This lemma implies that <u>a solution to the LIP always exists.</u>

- The solution, however, is <u>not unique.</u>

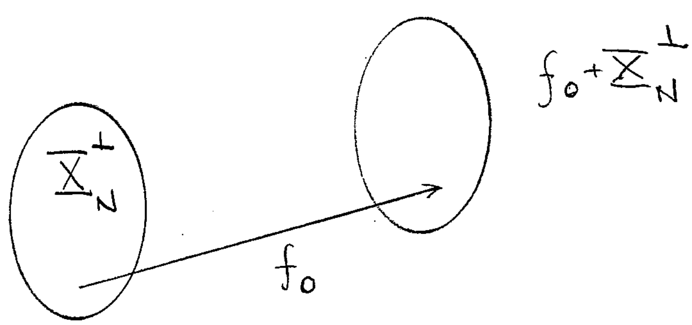
  Let $f_1 \in \overline{X}_N^{\perp}$, i.e.

  $$(f_1, \phi_n) = 0 \quad , \quad n = 1, \dots, N$$

  The set of solutions S to the LIP problem is:

  All functions $f := f_1 + f_0$, where f_0 is a given solution, and $f_1 \in \overline{X}_N^{\perp}$.

- The set of solutions S to LIP is therefore a translation of $\overline{X}_N^{\perp}$



Since $\overline{X}_N^{\perp}$ is a closed linear subspace (N.& S. p. 294, Th. 5.15

<u>S is a closed linear subspace of $\overline{X}$.</u>

Furthermore, <u>S is convex</u>, since

$$\begin{cases} x_1 = f_1 + f_0 \in S \\ x_2 = f_2 + f_0 \in S \end{cases} \qquad f_1, f_2 \in \overline{X}_N^{\perp}$$

Then $\alpha x_1 + (1-\alpha) x_2 = \underbrace{\alpha(f_1 - f_2) + f_2}_{\in \overline{X}_N^{\perp}} + f_0 \in S$

for all $\alpha \in [0,1]$

<u>Definition</u> <u>The normal solution f^{+}</u> to the LIP is the element in S with minimal norm.

This definition is welldefined, since for every closed convex set there exists an element of minimal norm.

(N. & S. p. 290, Exercise 2)

<u>Lemma:</u> $f^{+} \in \underline{X}_N$

<u>Proof:</u> If not $f^{+} = f_0 + f_1$, where $f_0 \in \underline{X}_N$, $f_1 \in \underline{X}_N^{\perp}$

This decomposition is unique (N. & S. p. 297, Th. 5.15.6)

and $\|f^{+}\|^2 = \|f_0\|^2 + \|f_1\|^2$

On the other hand $(f_0, \phi_n)_{\underline{X}} = (f^{+} - f_1, \phi_n)_{\underline{X}} = (f^{+}, \phi_n)_{\underline{X}} = g_n$

so f_0 solves the LIP. This contradicts f^{+} having minimal norm.

The result in this section 4.2 can be summarized in

<u>Theorem:</u>	If the functions $\{\phi_n\}_{n=1}^{N}$ are linearly independent, there always exists a unique solution of minimal norm f^{+} to the LIP. f^{+} is the unique solution orthogonal to $\underline{X}_N^{\perp}$

An explicit expression for f^{+} can be obtained as follows:

$$f^{+} \in \underline{X}_N \Rightarrow f^{+} = \sum_{n=1}^{N} a_n \phi_n$$

Determine the coefficients $\{a_n\}_{n=1}^{N}$

$$f^{+} \text{ is a solution } \Rightarrow (f^{+}, \phi_n)_{\underline{X}} = g_n \, , \quad n = 1, \ldots, N$$

$$\sum_{n'=1}^{N} G_{n'n} \, a_{n'} = g_n \, , \quad n = 1, \ldots, N$$

compl. conjugation

where $\boxed{G_{n'n} := (\phi_{n'}, \phi_n)_{\underline{X}}} = G^{*}_{nn'}$

the <u>Gram matrix</u>

Since $\{\phi_n\}_{n=1}^{N}$ are linearly independent G^{-1} exists.

$$a_n = \sum_{n'=1}^{N} G^{-1}_{n'n} \, g_{n'} \, , \quad n = 1, \ldots, N$$

and

$$\boxed{f^{+} = \sum_{n,n'=1}^{N} G^{-1}_{n'n} \, g_{n'} \, \phi_n}$$

This last representation of the normal solution f^+ shows that f^+ depends continuously on the data $\{g_n\}_{n=1}^N$

If $\{\delta g_n\}_{n=1}^N$ are variations in the data the corresponding variation δf^+ of f^+ is

$$\delta f^+ = \sum_{n,n'=1}^{N} G^{-1}_{n'n} \, \delta g_{n'} \, \phi_n$$

and $\|\delta f^+\|_X \to 0$, as $\{\delta g\}_{n=1}^N \to 0$

Thus the problem of the computation of f^+ is always <u>well-posed</u> in the strict mathematical sense.

However, it can be strongly <u>ill-conditioned</u> and therefore <u>numerically unstable.</u>

The normal solution defines a mapping A of $\underline{X}$ into itself, $f^{+} = Af$ defined by

$$f^{+} = \sum_{n,n'=1}^{N} G^{-1}_{n'n}\, g_{n'}\, \phi_n = \sum_{n,n'=1}^{N} G^{-1}_{n'n}\, (f, \phi_{n'})_{\underline{X}}\, \phi_n$$

The mapping A is simply the orthogonal projection of f onto $\underline{X}_N$

<u>Example</u>: $\underline{X} = L^2$, $(f,h)_{\underline{X}} = \int f(x)\, h^*(x)\, \underbrace{w(x)}_{\text{weights} > 0}\, dx$

Then the mapping A is an integral operator

$$(Af)(x) = \int A(x,x')\, f(x')\, dx'$$

with kernel $A(x,x')$

$$A(x,x') = \sum_{n,n'=1}^{N} G^{-1}_{n'n}\, w(x')\, \phi^*_{n'}(x')\, \phi_n(x)$$

4.3 The generalized solution or normal pseudosolution

- If the functions $\{\phi_n\}_{n=1}^{N}$ are not linearly independent and the data $\{g_n\}_{n=1}^{N}$ are affected by experimental errors, the normal solution does not exist. (Data $\{g_n\}_{n=1}^{N}$ and $\{\phi_n\}_{n=1}^{N}$ do not satisfy the same linear relationships.)

- A solution in the subspace of the linearly independent $\{\phi_n\}_{n=1}^{N}$ is possible, but impractical.

- Find a procedure that produces a solution in any case!

  The least-squares method is the most popular one!

Introduce the following definitions:

- The numbers $\{g_n\}_{n=1}^{N}$ are assumed to belong to a space Y, <u>the data space</u>, with euclidean structure, and with the inner product

$$(\bar{g}, \bar{h})_{\mathrm{Y}} := \sum_{n,n'=1}^{N} W_{nn'} g_{n'} h_n^*$$

where the weights $W_{nn'}$ are the elements of a positive definite matrix W and $W_{nn'} = W_{n'n}^*$ (Ordinary R^N have $W_{nn'} = \delta_{nn'}$)

- Also define <u>the linear operator $L : \mathrm{X} \to \mathrm{Y}$</u> as

$$\boxed{g_n = (Lf)_n = (f, \phi_n)_{\mathrm{X}} \quad , \quad n = 1, 2, \ldots, N}$$

- L is a bounded operator with norm $\|L\|$

$$\|L\| \leq \sqrt{\sum_{n,n'=1}^{N} |W_{nn'}| \, \|\phi_n\|_{\mathrm{X}} \, \|\phi_{n'}\|_{\mathrm{X}}} \qquad \text{(Prove this!)}$$

From section 4.2 it is seen that the mapping L is onto Y if $\{\phi_n\}_{n=1}^{N}$ are linearly independent. Otherwise the range of L is a subspace $Y_{N'}$ of Y, with dimension $N' < N$.

$$\boxed{\begin{aligned} R(L) &= Y_{N'} \\ \mathrm{Dim}\, Y_{N'} &= N' \end{aligned}} \qquad \boxed{\begin{aligned} \mathbb{X}_{N'} &= \mathrm{span}\{\phi_n\} \\ \mathrm{Dim}\, \mathbb{X}_{N'} &= N' \end{aligned}}$$

Notice that $\mathbb{X}_{N'} = \mathrm{span}\{\phi_n\}$ has the same dimension N'.

The orthogonal complement of $\mathbb{X}_{N'}$, called $\mathbb{X}_{N'}^{\perp}$, is just the null space of L

$$\boxed{N(L) = \mathbb{X}_{N'}^{\perp}}$$

This space contains the "invisibe" components of f, those components that give zero data $(f, \phi_n)_{\mathbb{X}} = 0$

Define the <u>adjoint operator L^*</u> to L by

$$(Lf, \bar{g})_Y = (f, L^*\bar{g})_{\bar{X}} \quad , \quad \forall f \in \bar{X} \, , \, \forall \bar{g} \in Y$$

L^* maps Y into $\bar{X}_{N'}$, since

for all $f \in \bar{X}_{N'}^{\perp}$, i.e. $f \in N(L)$,

$$Lf = 0 \quad \Rightarrow \quad (f, L^* g)_{\bar{X}} = 0$$

$$\Rightarrow \quad L^* g \in (\bar{X}_{N'}^{\perp})^{\perp} = \bar{X}_{N'} \qquad (\bar{X}_{N'} \text{ is closed})$$

An explicit expression for L^* is

$$L^*\bar{g} = \sum_{n,n'=1}^{N} W_{nn'} g_{n'} \phi_n \qquad , \qquad R(L^*) = \bar{X}_{N'}$$

<u>Check:</u>

$$(Lf, \bar{g})_Y = \sum_{n,n'=1}^{N} W_{nn'} (f, \phi_{n'})_{\bar{X}} \, g_n^*$$

$$(f, L^*\bar{g})_{\bar{X}} = \sum_{n,n'=1}^{N} W_{nn'}^* g_{n'}^* (f, \phi_n)_{\bar{X}} = \sum_{n,n'=1}^{N} W_{n'n} g_{n'}^* (f, \phi_n)_{\bar{X}}$$

The LIP

given $\bar{g} \in Y$ find $f \in \underline{X}$: $Lf = g$

- If $\{\phi_n\}$ are not linearly independent then there exists a solution if and only if $\bar{g} \in Y_N' = R(L)$

- If $\bar{g} \notin Y_{N'}$ define a <u>least-squares solution or pseudosolution</u> of $Lf = \bar{g}$ as any $\tilde{f} \in \underline{X}$:

$$\boxed{\| L\tilde{f} - \bar{g} \|_Y = \text{minimum}} \qquad (*)$$

Such an object $\tilde{f} \in \underline{X}$ does always exist, since $R(L) = Y_N'$ is a closed subspace, and there is always a unique $\bar{g}_0 \in Y_N'$ such that

$$\| \bar{g}_0 - \bar{g} \|_{\underline{Y}} = \inf \{ \| \bar{g}_1 - \bar{g} \|_Y \, , \; \bar{g}_1 \in Y_N' \}$$

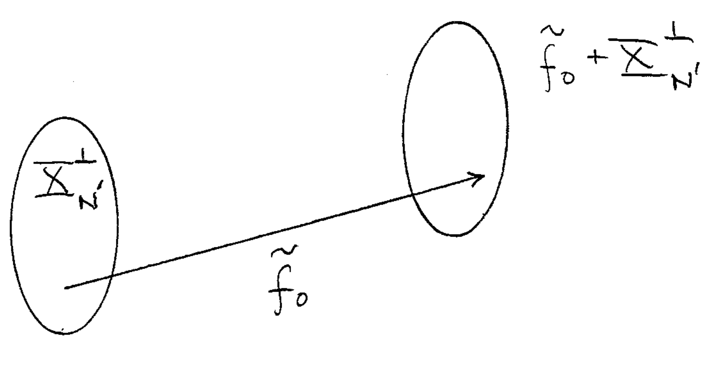
(see N. & S. p. 286, Th. 5.14.4)

<u>Remark:</u> The introduction of a pseudosolution makes a change in the concept of a solution to the LIP. Strict equality in $Lf = \bar{g}$ is no longer required. The "distance" between Lf and $\bar{g}$ is related to the numerical errors.

The set of pseudosolutions $\tilde{S}$ to $(*)$ is:

All functions $\tilde{f} = \tilde{f}_0 + \tilde{f}_1$, where $\tilde{f}_0$ is a given pseudosolution and $\tilde{f}_1 \in \overline{X}_{N'}^{\perp}$

- The set $\tilde{S}$ is therefore a translation of $\overline{X}_N^{\perp}$



The following propositions hold

i) if $\bar{g} \in Y_{N'}$ then the set of pseudosolutions, $\tilde{S}$ coincides with the set of solution S to $Lf = \bar{g}$.

ii) The set $\tilde{S}$ coincides with the solutions of the projected equation

$$L\tilde{f} = P\bar{g}$$

where $P\bar{g}$ is the orthogonal projection of $\bar{g}$ onto $Y_{N'}$

iii) The set $\tilde{S}$ coincides with the solutions of the normal equations (Euler's equation)

$$L^*L\tilde{f} = L^*\bar{g}$$

Similar to the normal solution case, $\tilde{S}$ is a closed, convex, linear subspace of $\underline{X}$ and therefore there exists an element f^+ with minimal norm in $\tilde{S}$

Definition The solution f^+ in $\tilde{S}$ with minimal norm is called the <u>normal pseudosolution</u> or <u>generalized solution</u>.

Remark: If $\{\phi_n\}_{n=1}^{N}$ are linearly independent
the normal solution = the generalized solution

Theorem: For any $\bar{g} \in Y$ there exists a unique generalized solution f^+. f^+ depends continuously on the data in the sense that if $\delta\bar{g}$ are data errors and δf^+ induced error

$$\|\delta\bar{g}\|_Y \to 0 \Rightarrow \|\delta f^+\|_{\underline{X}} \to 0$$

Thus the problem of the computation of f^+ is always <u>well-posed</u>. However, it can be strongly <u>ill-conditioned.</u>

— o —

The generalized solution defines a mapping $L^+ : Y \to \underline{X}$, the <u>generalized inverse</u> or <u>Moore-Penrose inverse</u> of L by

$$\boxed{f^+ = L^+ \bar{g}}$$

This operator L^+ is continuous (see theorem p. 27) and $f^+ \in \underline{X}_{N'}$ (If not, $f^+ = f_0 + f_1$, where $f_0 \in \underline{X}_{N'}$, $f_1 \in \underline{X}_{N'}^{\perp}$ (unique decomposition, see N. & S., p. 297, Th. 5.15.6) and $\|f^+\|_{\underline{X}}^2 = \|f_0\|_{\underline{X}}^2 + \|f_1\|_{\underline{X}}^2$

On the other hand

$$\|Lf_0 - \bar{g}\|_Y = \|Lf^+ - \bar{g} - \underbrace{Lf_1}_{=\bar{0}}\|_Y = \|Lf^+ - \bar{g}\|_Y$$

This contradicts f^+ having minimal norm)

<u>Error analysis</u> :

$$\left.\begin{aligned} L(f^+ + \delta f^+) &= \bar{g} + \delta\bar{g} \\ Lf^+ &= \bar{g} \end{aligned}\right\} \Rightarrow L\delta f^+ = \delta\bar{g}$$

$$\|\delta f^+\|_{\underline{X}} = \|L^+\delta\bar{g}\|_{\underline{X}} \leq \|L^+\|\,\|\delta\bar{g}\|_Y$$

$$\|\bar{g}\|_Y = \|Lf\|_Y \underset{\substack{\uparrow \\ f^+ - f \,\in\, X_{N'}^{\perp}}}{=} \|Lf^+\|_Y \leq \|L\|\,\|f^+\|_{\underline{X}}$$

Thus $\quad \|\delta f^+\|_{\underline{X}} / \|f^+\|_{\underline{X}} \leq \|L^+\|\,\|\delta\bar{g}\|_Y \|L\| / \|\bar{g}\|_Y$

$$\|\delta f^+\|_{\underline{X}} / \|f^+\|_{\underline{X}} \leq C(L)\,\|\delta\bar{g}\|_Y / \|\bar{g}\|_Y$$

where the condition number $C(L) := \|L\|\,\|L^+\|$

- Note that when the ordinary inverse L^{-1} of L exists then this definition coincides with the usual definition of condition number $C(L) = \|L\|\,\|L^{-1}\|$

4.4 The singular system (SVD)

From previous section 4.3, it is shown that the generized solution $f^{+} \in \underline{X}_{N'} = \text{span}\{\phi_n\}$ (if $\{\phi_n\}$ linearly independent, $N' = N$)

f^{+} can be expressed in any ON-basis of $\underline{X}_{N'}$. A natural choice whould be the orthogonal eigenvectors of $L^{*}L$ (cf. p. 26)

$$\boxed{L^{*}L\, u_n = \lambda_n u_n \quad , \quad n = 1, \dots, N'}$$

These orthogonol vectors $\{u_n\}_{n=1}^{N'}$ transform, by the operator L, into a set of orthogonal vectors $\{L u_n\}_{n=1}^{N'}$ of Y, since

$$(L u_n, L u_{n'})_Y = (u_n, L^{*} L u_{n'})_{\underline{X}} = \lambda_{n'}^{*} (u_n, u_{n'})_{\underline{X}} = \lambda_{n'}^{*} \delta_{nn'}$$

Introduce the notation

$$\boxed{\begin{aligned}\tilde{L} &= L^* L \\ \hat{L} &= L L^*\end{aligned}}$$

<u>Lemma:</u> The operator $\tilde{L} : \underline{X} \to \underline{X}_{N'}$ is a finite rank, self-adjoint, non-negative operator on $\underline{X}$ with explicit expression

$$\tilde{L} f = \sum_{n,n'=1}^{N'} W_{nn'} (f, \phi_{n'})_{\underline{X}} \phi_n$$

<u>Proof:</u> i) Since $R(L^*) = \underline{X}_{N'}$ is finite dimensional the rank of $\tilde{L}$ is finite (cf. p. 23)

ii) $(f, \tilde{L} f)_{\underline{X}} = (f, L^* L f)_{\underline{X}} = (Lf, Lf)_Y \geq 0, \ \forall f \in \underline{X}$

$\Rightarrow$ The operator $\tilde{L}$ is non-negative.

iii/ $\tilde{L}$ is defined on all of $\underline{X}$ and

for any $f_1, f_2 \in \underline{X}$

$$(f_1, \tilde{L} f_2)_{\underline{X}} = (f_1, L^* L f_2)_{\underline{X}} = (Lf_1, Lf_2)_Y$$

$$= (Lf_2, Lf_1)^*_Y = (f_2, L^* L f_1)^*_{\underline{X}} = (\tilde{L} f_1, f_2)_{\underline{X}}$$

$\Rightarrow \tilde{L}$ self-adjoint on $\underline{X}$

iv/ From p. 23, $L^* \bar{g} = \sum_{nn'=1}^{N'} W_{nn'} g_{n'} \phi_n$

and from p. 21 $(Lf)_n = (f, \phi_n)_{\underline{X}}$

it follows

$$\tilde{L} f = L^* L f = \sum_{n,n'=1}^{N'} W_{nn'} (f, \phi_{n'})_{\underline{X}} \phi_n$$

QED.

<u>Lemma:</u> The operator $\hat{L} : Y \to Y_{N'}$ is

self-adjoint, non-negative operator on Y

with components

$$(\hat{L}\bar{g})_n = (L^*\bar{g}, \phi_n)_{\underline{X}} = \sum_{n=1}^{N'} \hat{L}_{nn'} g_{n'}$$

where the matrix $\hat{L}_{nn'} = \sum_{n''=1}^{N'} G_{n''n} W_{n''n'}$

<u>Proof:</u> i) $(\hat{L}\bar{g}, \bar{g})_Y = (LL^*\bar{g}, \bar{g})_Y =$

$$= (L^*\bar{g}, L^*\bar{g})_{\underline{X}} \geq 0 \quad , \quad \forall \bar{g} \in Y$$

$\Rightarrow$ The operator $\hat{L}$ is non-negative

ii) For all $\bar{g}_1, \bar{g}_2 \in Y$

$$(\hat{L}\bar{g}_1, \bar{g}_2)_Y = (LL^*\bar{g}_1, \bar{g}_2)_Y =$$

$$= (L^*\bar{g}_1, L^*\bar{g}_2)_{\underline{X}} = (L^*\bar{g}_2, L^*\bar{g}_1)^*_{\underline{X}} =$$

$$= (LL^*\bar{g}_2, \bar{g}_1)^*_Y = (\bar{g}_1, \hat{L}\bar{g}_2)_Y$$

$\Rightarrow$ $\hat{L}$ self-adjont on Y

iii) The components of $\hat{L}\bar{g}$ is

$$(\hat{L}\bar{g})_n = (LL^*\bar{g})_n = (L^*\bar{g}, \phi_n)_{\underline{X}}$$

$$= \sum_{n',n''=1}^{N'} W_{n''n'} g_{n'} (\phi_{n''}, \phi_n)_{\underline{X}} =$$

$$= \sum_{n',n''=1}^{N'} G_{n''n} W_{n''n'} g_{n'} = \sum_{n'=1}^{N'} \hat{L}_{nn'} g_{n'}$$

QED.

The rank of $\hat{L}$ is equal to the rank of $G_{nn'}$.

This implies that $\mathrm{Rank}(\hat{L}) = N' = \#$ of linearly independent $\{\phi_n\}$.

Thus the matrix $\hat{L}_{nn'}$ has N' positive eigenvalues

$$\alpha_1^2 \geq \alpha_2^2 \geq \dots \geq \alpha_{N'}^2 > 0$$

and eigenvectors $\{\bar{v}_n\}_{n=1}^{N'}$, ON basis in $Y_{N'}$

Lemma: The operator $\tilde{L}$ has the same eigenvalues as $\hat{L}$. The eigenvectors $\{u_n\}_{n=1}^{N'}$; the ON basis in $\underline{X}_{N'}$ and

$$\boxed{\begin{aligned} L u_n &= \alpha_n \bar{v}_n \\ L^* \bar{v}_n &= \alpha_n u_n \end{aligned}}$$

Proof: Suppose $\bar{v}_n$ is an eigenvector of $\hat{L}$ with eigenvalue α_n^2.

$\hat{L}\,\bar{v}_n = \alpha_n^2\,\bar{v}_n \quad \Rightarrow$

$L^* L L^* \bar{v}_n = \alpha_n^2 L^* \bar{v}_n \quad \Rightarrow$

$L^* \bar{v}_n$ eigenvector with eigenvalue α_n^2 of $\tilde{L} = L^* L$

Define $u_n = \frac{1}{\alpha_n} L^* \bar{v}_n$

$\| u_n \|_{\underline{X}} = \alpha_n^{-2} (L^* \bar{v}_n, L^* \bar{v}_n)_{\underline{X}} = \alpha_n^{-2} (L L^* \bar{v}_n, \bar{v}_n)_Y = 1$

Thus $\{u_n\}_{n=1}^{N'}$, ON basis of $\underline{X}_{N'}$

On the other hand, assume u_n is an eigenvector $(\|u_n\|^2=1)$ of $\tilde{L}$ with eigenvalue $\lambda_n>0$.

$$\tilde{L}u_n=\lambda_n u_n \Rightarrow$$

$$LL^*Lu_n=\lambda_n Lu_n$$

$\Rightarrow$ Lu_n eigenvector of $\hat{L}=LL^*$ with eigenvector λ_n

$\Rightarrow$ $\lambda_n=\alpha_n^2$ and $\bar{v}_n=c_n L u_n$

$$1=\|\bar{v}_n\|^2=c_n^2(Lu_n,Lu_n)_Y=c_n^2(u_n,\tilde{L}u_n)_X=c_n^2\lambda_n$$

$$\Rightarrow \quad \bar{v}_n=\frac{1}{\alpha_n}L\bar{u}_n$$

QED.

<u>Definition:</u> <u>The singular system</u> of the operator L is the set of triples $\{\alpha_n;u_n,\bar{v}_n\}_{n=1}^{N'}$.

α_n are called the <u>singular values</u>, and u_n <u>the singular functions</u> and $\bar{v}_n$ <u>the singular vectors</u>.

<u>Theorem:</u> The unique generalized solution f^+ is

$$\boxed{f^+ = \sum_{n=1}^{N'} \alpha_n^{-1} (\bar{g}, \bar{v}_n)_Y u_n}$$

<u>Proof:</u> Expand $\bar{g}$ in $\{\bar{v}_n\}_{n=1}^{N'}$, ON basis in $Y_{N'}$

$$\bar{g} = \sum_{n=1}^{N'} (\bar{g}, \bar{v}_n)_Y \bar{v}_n + \bar{g}^\perp$$

where $\bar{g}^\perp \in Y_{N'}^\perp$

Since $f^+ \in X_{N'}$ expand in $\{u_n\}_{n=1}^{N'}$,

ON basis in $X_{N'}$

$$f^+ = \sum_{n=1}^{N'} a_n u_n$$

$$Lf^+ = \sum_{n=1}^{N'} a_n L u_n = \sum_{n=1}^{N'} a_n \alpha_n \bar{v}_n$$

$$\| Lf^+ - \bar{g} \|_Y = \sum_{n=1}^{N'} (a_n \alpha_n - (\bar{g}, \bar{v}_n)_Y)^2 + \|\bar{g}^\perp\|_Y \geq \|\bar{g}^\perp\|_Y$$

with minimum when

$$a_n \alpha_n = (\bar{g}, \bar{v}_n)_Y$$

$$a_n = \alpha_n^{-1} (\bar{g}, \bar{v}_n)_Y$$

QED

<u>Lemma:</u> $\|L\| = \alpha_1$ and $\|L^+\| = \alpha_{N'}^{-1}$

<u>Proof:</u> Expand $f = \sum_{n=1}^{N'} a_n u_n + f^\perp$

where $f^\perp \in \overline{X}_{N'}^\perp$

$$Lf = \sum_{n=1}^{N'} a_n \alpha_n \overline{v}_n$$

$$\|L\| = \sup_{f \in \underline{X}} \frac{\|Lf\|_Y}{\|f\|_{\underline{X}}} = \sup_{f \in \underline{X}} \frac{\sqrt{\sum_{n=1}^{N'} |a_n \alpha_n|^2}}{\left(\sum_{n=1}^{N'} |a_n|^2 + \|f^\perp\|^2\right)^{1/2}}$$

$$\leq \sup_{f \in \underline{X}_{N'}} \frac{\alpha_1 \sqrt{\sum_{n=1}^{N'} |a_n|^2}}{\sqrt{\sum_{n=1}^{N'} |a_n|^2}} = \alpha_1$$

and the supremum is obtained by $f = u_1$

$\Rightarrow \|L\| = \alpha_1$

The definition of $f^+ = L^+ \overline{g}$ implies

$$\|L^+\| = \sup_{\overline{g} \in Y} \frac{\|L^+ \overline{g}\|_{\underline{X}}}{\|\overline{g}\|_Y} = \sup_{\overline{g} \in Y} \frac{\sqrt{\sum_{n=1}^{N'} |\alpha_n^{-1} (\overline{g}, \overline{v}_n)_Y|^2}}{\|\overline{g}\|_Y}$$

$$\leq \alpha_{N'}^{-1} \sup_{\overline{g} \in Y} \sqrt{\sum_{n=1}^{N'} |(g, \overline{v}_n)_Y|^2} \Big/ \|g\|_Y \leq \alpha_{N'}^{-1}$$

and the supremum is obtained by $\overline{g} = \overline{v}_{N'}$

$\Rightarrow \|L^+\| = 1/\alpha_{N'}$

Q.E.D.

The condition number $C(L) = \|L\| \, \|L^+\|$ for L thus is

$$\boxed{C(L) = \alpha_1 / \alpha_N}$$

This number reflects the linear dependence of $\{\phi_n\}_{n=1}^{N}$ or of the Gram matrix.

If the scalar product in Y is the usual Euclidean norm $W_{nn'} = \delta_{nn'}$ then by the lemma on p. 33

$\hat{L}_{nn'} = G_{nn'}$ and the condition number $C(L)$ is just the quotient between the largest and the smallest eigenvalue of $G_{nn'}$.

4.5 Examples

4.5.1 Interpolation

Given $\{g_n\}_{n=1}^{N}$ on N points $a \le x_1 < x_2 < \dots < x_N \le b$

Determine a function f on $[a,b]$:

$$f(x_n) = g_n$$

(A) Search for a function f that is bandlimited with bandwith Ω, i.e.

$$\hat{f}(\omega) = 0 \qquad |\omega| > \Omega \qquad (\hat{f} \text{ Fourier transform})$$

Any such function $f(x)$ satisfies

$$f(x) = \int_{-\infty}^{\infty} \frac{\sin \Omega (x-x')}{\pi (x-x')} f(x')\,dx'$$

This interpolation problem can be put into the LIP by choosing $(\underline{X} = L^2)$ (c.f. $(f, \phi_n) = g_n$)

$$\boxed{\phi_n(x) = \frac{\sin \Omega (x - x_n)}{\pi (x - x_n)}}$$

The corresponding Gram matrix is

$$G_{nn'} = (\phi_n, \phi_{n'})_{\overline{X}} = \int_{-\infty}^{\infty} \phi_n(x)\, \phi_{n'}(x)\, dx$$

$$= \int_{-\infty}^{\infty} \frac{\sin \Omega(x-x_n)}{\pi(x-x_n)} \frac{\sin \Omega(x-x_{n'})}{\pi(x-x_{n'})}\, dx = \frac{\sin \Omega(x_n - x_{n'})}{\pi(x_n - x_{n'})}$$

since ϕ_n itself is a bandlimited function.

Special case: Sample points at Nyquist rate

$$\boxed{x_{n+1} - x_n = \pi/\Omega}$$

$$\Rightarrow \quad G_{nn'} = \frac{\Omega}{\pi}\delta_{nn'} \qquad \text{(diagonal)}$$

and we get the normal solution (Whittaker - Shannon)

(cf. p. 19)

$$\boxed{\begin{aligned} f^{+}(x) &= \sum_{n,n'=1}^{N} G^{-1}_{nn'}\, g_{n'}\, \phi_n(x) = \\ &= \sum_{n=1}^{N} g_n \frac{\sin \Omega(x-x_n)}{\Omega(x-x_n)} \end{aligned}}$$

If the points are sampled at a rate less than the Nyquist rate π/Ω

$$x_n = x_1 + (n-1)\sigma \quad , \; n = 1, \dots, N$$

and

$$W := \frac{\sigma \Omega}{2\pi} < 1/2$$

$$G_{nn'} = \frac{\sin 2\pi W (n-n')}{\pi (n-n')}$$

Slepian has studied this problem (Bell Syst. Tech. J. 57, 1371–1430 (1978))

and finds that

$$c(L) \approx \frac{\delta}{N^{1/4}} e^{\gamma N/2}$$

where
$$\beta = 1 - \cos 2\pi W$$
$$\gamma = \ln \left[1 + \frac{2\sqrt{\beta}}{2-\sqrt{\beta}} \right]$$

and the number of singular values significantly different from zero is approximately

$$N_0 = 2WN = \frac{\Omega}{\pi} x_N \Rightarrow \text{Nyquist rate!!}$$

(B) Search for a smooth function $f \in C^k([0,1])$

such that $f(x_n) = g_n$, $x_n \in [0,1]$

Introduce the scalar product

$$(f,h)_{X} = \sum_{j=0}^{k-1} f^{(j)}(0)\, h^{(j)}(0) + \int_0^1 f^{(k)}(x)\, h^{(k)}(x)\, dx$$

Taylors formula gives

$$f(x) = \sum_{j=0}^{k-1} \frac{f^{(j)}(0)}{j!} x^j + \frac{1}{(k-1)!} \int_0^x (x-x')^{k-1} f^{(k)}(x')\, dx'$$

This interpolation problem can be put into the LIP

by choosing

$$\phi_n(x) = \sum_{j=0}^{k-1} \frac{x^j x_n^j}{(j!)^2} + \sum_{j=0}^{k-1} (-1)^{k+j+1} \frac{x^{2k-j-1} \; x_n^j}{(2k-j-1)!\, j!}$$

$$+ \frac{(-1)^k}{(2k-1)!} H(x-x_n)\,(x-x_n)^{2k-1}$$

where $H(x)$ is the Heaviside function $H(x) = \begin{cases} 1 & , x>0 \\ 0 & , x<0 \end{cases}$

<u>Check:</u> $\phi_n^{(j)}(0) = \dfrac{x_n^j}{j!}$ for $j = 0, 1, \dots, k-1$

$$\phi_n^{(k)}(x) = \sum_{\nu=k}^{2k-1} (-1)^{k-\nu} \frac{x^{\nu-k}\, x_n^{2k-\nu-1}}{(2k-\nu-1)!\,(\nu-k)!}$$

$$+ \frac{(-1)^k}{(2k-1)!} H(x-x_n) \frac{(2k-1)!}{(k-1)!} (x-x_n)^{k-1}$$

$$= \sum_{j=0}^{k-1} (-1)^j \frac{x^j\, x_n^{k-j-1}}{(k-j-1)!\, j!}$$

$$+ \frac{(-1)^k}{(k-1)!} H(x-x_n)(x-x_n)^{k-1}$$

$$= \frac{(x_n - x)^{k-1}}{(k-1)!} \underbrace{\left[1 - H(x-x_n)\right]}_{H(x_n - x)}$$

— o —

The Gram matrix $G_{nn'} = (\phi_n, \phi_{n'})_{\underline{X}}$

$$G_{nn'} = \sum_{j=0}^{k-1} \frac{x_n^j\, x_{n'}^j}{(j!)^2} + \left[(k-1)!\right]^{-2} \int_0^{\min(x_n, x_{n'})} (x_n - x)^{k-1} (x_{n'} - x)^{k-1}\, dx$$

The normal solution $\overset{+}{f}(x) = \sum\limits_{n,n'=1}^{N} G^{-1}_{nn'}\, g_{n'}\, \phi_n(x)$

defines a <u><u>spline of degree $m = 2k-1$</u></u>

4.5.2 Moment problem

Find a function $f(x) \in C(0,1)$ given its first N moments, i.e.

$$\int_0^1 x^{n-1} f(x)\, dx = g_n \qquad n = 1, \ldots, N$$

This problem is unique if $N = \infty$ (N. & S. p. 329, Exercise 4)

Take $\underline{X} = L^2(0,1)$, $W_{nn'} = \delta_{nn'}$ and choose

$$\phi_n(x) = x^{n-1} \quad , \; n = 1, \ldots, N$$

The adjoint operator is then (p. 23)

$$(L^* \bar{g})(x) = \sum_{nn'=1}^{N} W_{nn'} g_{n'} \phi_n(x) = \sum_{n=1}^{N} g_n x^{n-1}$$

and $R(L^*) = \underline{X}_N = \{\text{polynomials of degree } N-1\}$

The normal solution is also a polynomial of degree $N-1$.

Represent the normal solution in (shifted) Legendre polynomials

$$f^{+}(x) = \sum_{n=1}^{N} c_n P_{n-1}(2x-1)$$

Determine the coefficients c_n from $(f, \phi_n)_{\mathbb{X}} = g_n$

$$\int_0^1 x^{n-1} f^{+}(x)\, dx = g_n \quad , \quad n = 1, \dots, N$$

or $$g_n = \sum_{n'=1}^{n} \beta_{nn'} c_{n'} \quad , \quad n = 1, \dots, N$$

where $$\beta_{nn'} = \int_0^1 x^{n-1} P_{n'}(2x-1)\, dx \quad , \quad n > n'$$

<u>Method:</u> i/ Compute $\beta_{nn'}$, $n' = 1, \dots, n$

ii/ Recursively determine c_n, $n = 1, \dots, N$

This procedure masks the severe ill-conditioning of the problem.

To see the ill-conditioning of the method find the operator $\tilde{L} = L^* L$

From p. 31

$$(\tilde{L} f)(x) = \sum_{n,n'=1}^{N} W_{nn'} \, (f, \phi_{n'})_{\underline{X}} \, \phi_n(x)$$

$$= \sum_{n=1}^{N} \int_0^1 x'^{\,n-1} f(x') \, dx' \, x^{n-1}$$

$$= \int_0^1 \frac{1-(xx')^N}{1-xx'} f(x') \, dx'$$

and from p. 33

$$\hat{L}_{nn'} = \sum_{n''=1}^{N} G_{n''n} W_{n''n'} = G_{n'n} = (\phi_{n'}, \phi_n)_{\underline{X}}$$

$$= \int_0^1 x^{n+n'-2} \, dx = 1/(n+n'-1)$$

<u>This is a Hilbert matrix</u>

The condition number $C(L)$ for large N is

$$C(L) \sim e^{1.75 N}$$

To find a normal solution stable with respect to errors of 0.1% on the moments, only five moments of f can be given $\left(C(L)_{N=5} = 690.6 \right)$.
However, using SVD, information on the function f increases even if rather slowly with N

Example: $N = 100$

Nine components can be used, since

$$\alpha_1 / \alpha_9 = 505.8$$

while

$$\alpha_1 / \alpha_{10} = 1319.$$

4.5.3. FT inversion

Assume $g(t)$ is a bandlimited function

$$g(t) = \int_{-\Omega}^{\Omega} f(\omega)\, e^{-it\omega}\, d\omega$$

Given $g(t)$ at $t_n = (n-1)\tau$, $n = 1, \dots, N$

Determine the function $f(\omega)$

Introduce $W := \frac{\tau\Omega}{2\pi}$ (see p. 42)

and change variable $x = \left(\frac{\tau}{2\pi}\right)\omega$

$$g(t) = \int_{-W}^{W} f\left(\frac{2\pi x}{\tau}\right) e^{-it\frac{2\pi x}{\tau}}\, \frac{2\pi}{\tau}\, dx$$

Our problem is a LIP : $g_n = (f, \phi_n)_{\underline{X}}$

$$\begin{cases} g_n = \frac{\tau}{2\pi}\, g((n-1)\tau) \\ \phi_n(x) = \exp 2\pi i (n-1)x \end{cases} , \quad n = 1, \dots, N$$

(A.) $W = \frac{1}{2}$ (Nyquist rate)

Then $\{\phi_n\}_{n=1}^{\infty}$ is an ON basis of $L^2(-W,W)$

The normal solotion f^{+} is (cf. p. 17)

$$f^{+}(x) = \sum_{n=1}^{N} g_n \phi_n(x)$$

$$= \sum_{n=1}^{N} g_n \exp[2\pi i (n-1) x]$$

This is nothing but the truncated Fourier series expansion of f with Fourier coefficients g_n

(B.) $W < \frac{1}{2}$ (oversampled data)

This problem becomes ill-conditioned and the analysis resembles very much that of example 4.5.1.A (interpolation of a bandlimited function)

4.6 Regularization algorithms

For a LIP with discrete data $\{g_n\}_{n=1}^{N}$ the generalized inverse L^{+} is always continuous (see theorem p. 27).

However, if $C(L) = \|L\| \|L^{+}\| >> 1$ the LIP is ill-conditioned and there is a need to regularize the problem.

(Note that the SUD decomposition itself is a kind of regularization of the original LIP).

Instead of considering the generalized inverse L^{+} we consider a family of of operators R_μ.

The precise definition is

<u>Definition</u>: A one-parameter family of linear operators $\{R_\mu\}_{\mu>0} : Y \to \underline{X}$ is a <u>regularization algorithm</u> for the approximate computation of the generalized inverse L^+ if

i) for any $\mu > 0$, $\mathcal{R}(R_\mu) \subseteq \underline{X}_N$,
(the range of $R_\mu \subseteq$ the span of $\{\phi_n\}_{n=1}^N$)

ii) for any $\mu > 0$, $\| R_\mu \| \le \| L^+ \| \underset{\uparrow}{=} \alpha_N^{-1}$,
p. 38

iii) $\lim_{\mu \to 0} R_\mu = L^+$

The parameter μ is called <u>the regularization parameter.</u> How to choose it will be a later problem. This choice is related to the statistics of the experimental errors in $\{g_n\}_{n=1}^N$.

Denote noisy data by $\bar{g}_\epsilon$ represented as

$$\boxed{\bar{g}_\epsilon = Lf^{\dagger} + \bar{h}_\epsilon}$$

$Lf^{\dagger}$ represents the noise-free signal

$\bar{h}_\epsilon$ represents the effect of noise, such that

$$\boxed{\| Lf^{\dagger} - \bar{g}_\epsilon \|_Y \leq \epsilon}$$

Notice that we do not assume the noise to be additive! $\bar{h}_\epsilon$ can be signal dependent.

Then, for given regularization algorithm and given value of the regularization parameter μ the approximation of the generalized solution $f^{\dagger}$ provided by R_μ is

$$\boxed{f_{\mu,\epsilon} = R_\mu g_\epsilon}$$

Compare with f^{+}

$$\boxed{f_{\mu,\epsilon} - f^{+} = \left(R_\mu L f^{+} - f^{+}\right) + R_\mu \bar{h}_\epsilon}$$

The <u>first term</u> represents the approximation error due to the algorithm R_μ instead of L^{+}

This term $\to 0$ as $\mu \to 0$ (condition iii) p. 52)

The <u>second term</u> is the error induced by the noise.

This term grows to unacceptable values as $\mu \to 0$.

Therefore, one looks for <u>a compromise</u> between approximation error and the error propagation from the data to the solution.

For any given $\mu > 0$ define $A_\mu : \underline{X} \to \underline{X}$

$$\boxed{A_\mu = R_\mu L}$$

The function $A_\mu f^+$ is the approximation of f^+ which can be obtained in the absence of noise.

In the limit $\mu \to 0$ this operator A_μ converges to the mapping

$$\boxed{A = L^+ L}$$

(cf. p. 19)

A is a projection operator, namely, the orthogonal projection onto the space $\underline{X}_N' = \mathrm{span}\{\phi_n\}_{n=1}^N$.

<u>Example</u> (cf. p. 19) $\underline{X} = L^2$

$$(f, h)_{\underline{X}} = \int f(x)\, h^*(x) \underbrace{w(x)}_{\text{weights} > 0} dx$$

Assume $N' = N$ (the normal solution f^+ exists)

From condition i) on p. 52 $\left(R(R_\mu) \subseteq \underline{X}_N \right)$

$$(R_\mu \bar{g})(x) = \sum_{n,n'=1}^{N} \rho_\mu^{n'n} g_{n'} \phi_n(x)$$

Condition iii) p. 52 is satisfied if and only if

$$\lim_{\mu \to 0} \rho_\mu^{n'n} = G^{-1}_{n'n} \qquad G_{n'n} = (\phi_{n'}, \phi_n)_{\underline{X}}$$

(cf. p. 17)

Therefore,

$$(A_\mu f)(x) = (R_\mu L f)(x) = \sum_{n,n'=1}^{N} \rho_\mu^{n'n} (f, \phi_{n'})_{\underline{X}} \phi_n(x)$$

$$= \int f(x') \sum_{n,n'=1}^{N} \rho_\mu^{n'n} \phi^*_{n'}(x') \phi_n(x)\, w(x')\, dx' = \int A_\mu(x, x') f(x')\, dx'$$

where the kernel $\quad A_\mu(x, x') = \sum_{n,n'=1}^{N} \rho_\mu^{n'n} \phi^*_{n'}(x') \phi_n(x)\, w(x')$

4.7 Spectral windows and numerical filtering

Of special interest is the following regularization algorithm

$$\boxed{R_\mu \bar{g} = \sum_{k=1}^{N'} W_{\mu,k}\, \alpha_k^{-1}\, (\bar{g}, \bar{v}_k)_Y\, u_k}$$

where $\{\alpha_k, \bar{v}_k, u_k\}_{k=1}^{N'}$ is the singular system of L

and where the set $\{W_{\mu,k}\}_{k=1}^{N'}$ satisfies

a/ $0 \le W_{\mu,k} \le 1$ for all $\mu > 0$ and $k = 1, 2, \dots, N'$

b/ $\lim_{\mu \to 0} W_{\mu,k} = 1$, $k = 1, 2, \dots, N'$

With these assumptions R_μ is a regularization algorithm

Check: i) $R(R_\mu) \subseteq X_{N'}$

since $\{u_k\}$ ON-basis in $X_{N'}$

ii) $$\|R_\mu\| = \sup_{\bar{g}\in Y} \frac{\|R_\mu \bar{g}\|_X}{\|\bar{g}\|_Y} = \sup_{\bar{g}\in Y} \frac{\sqrt{\sum_{k=1}^{N'} |W_{\mu,k}\,\alpha_k^{-1}\,(\bar{g},\bar{v}_k)_Y|^2}}{\|\bar{g}\|_Y}$$

$$\leq \max_k \left|\frac{W_{\mu,k}}{\alpha_k}\right| \sup_{\bar{g}\in Y} \sqrt{\sum_{k=1}^{N'} |(\bar{g},\bar{v}_k)_Y|^2} \Big/ \|\bar{g}\|_Y$$

$$\leq \max_k \left(\frac{W_{\mu,k}}{\alpha_k}\right)$$ with equalsign obtained for $\bar{g} = \bar{v}_{k'}$

(k'=the k value that gives max

Thus $$\|R_\mu\| = \max_k (W_{\mu,k}/\alpha_k) \leq \max_k (1/\alpha_k)$$

$$= 1/\alpha_{N'} = \|L^+\| \qquad \text{(p. 38)}$$

since $0 \leq W_{\mu,k} \leq 1$ by assumption a)

iii) $$\lim_{\mu\to 0} R_\mu \bar{g} = \sum_{k=1}^{N'} 1\cdot \alpha_k^{-1}\,(\bar{g},\bar{v}_k)_Y\, u_k = L^+\bar{g}$$

Thus $$\lim_{\mu\to 0} R_\mu = L^+$$

And all three conditions i) – iii) for R_μ to be a regularization algorithm are satisfied.

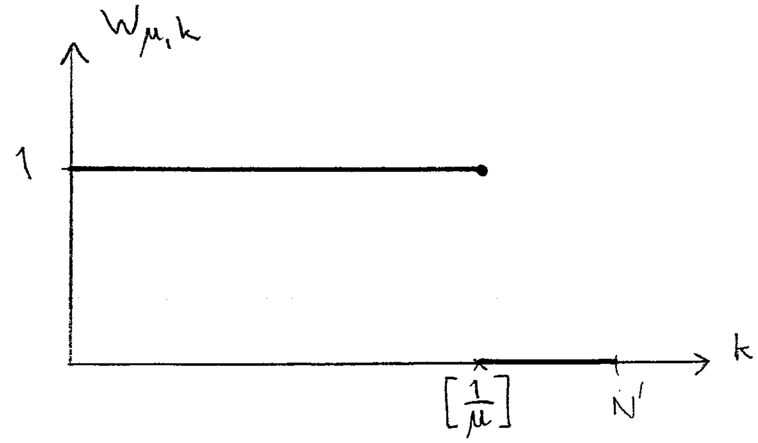
<u>Example 1</u> The rectangular window provides the simplest regularization algorithm

$$W_{\mu,k} = \begin{cases} 1 & k = 1, 2, \dots, [1/\mu] \\ 0 & k = [1/\mu]+1, \dots, N' \end{cases}$$

Check that conditions a) and b) are satisfied!

$$R_\mu \bar{g} = \sum_{k=1}^{[1/\mu]} \alpha_k^{-1} (\bar{g}, \bar{v}_k)_Y \, u_k$$

and $R_\mu = L^+$ if $\frac{1}{\mu} > N'$

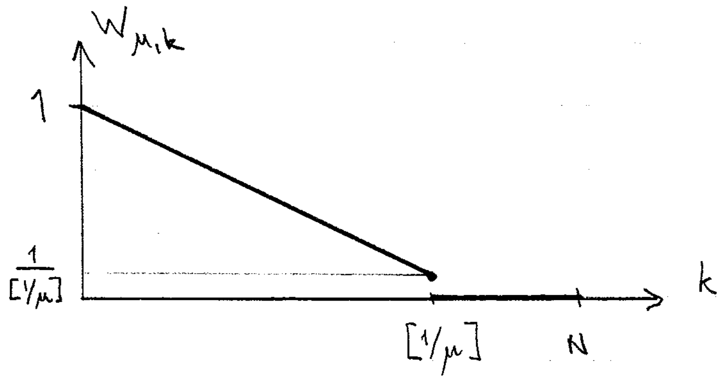


This method is also called <u>numerical filtering</u>

Example 2 The triangular window

$$W_{\mu,k} = \begin{cases} 1 - (k-1)/[1/\mu] & k = 1, 2, \dots, [1/\mu] \\ 0 & k = [1/\mu]+1, \dots, N' \end{cases}$$

Check the conditions a/ and b/ !



4.8 The Tikhonov regularization algorithm

This is a special kind of regularization algorithm. It was first introduced as a minimization problem. We will show that this minimization problem can be formulated as a spectral window.

Minimizing problem

For each $\mu > 0$, find $f = f_\mu$ that minimizes

$$\Phi_\mu[f] = \| Lf - \bar{g} \|_Y^2 + \mu \|f\|_{\underline{X}}^2$$

μ is the regularization parameter

Example of relevant norm in $\underline{X}$ (Sobolev)

$$\|f\|_{\underline{X}}^2 = \sum_{i=0}^{M} \int \underbrace{p_i(x)}_{>0} \left| \frac{d^i f(x)}{dx^i} \right|^2 dx$$

Since $\{u_k\}_{k=1}^{N'}$ is an ON-basis in $\underline{X}_{N'}$

and $\{\bar{v}_k\}_{k=1}^{N'}$ is an ON-basis in $R(L) = Y_{N'}$

Expand
$$\begin{cases} f = \sum_{k=1}^{N'} (f, u_k)_{\underline{X}}\, u_k + f_0 & , f_0 \in \underline{X}_{N'}^{\perp} \\ \bar{g} = \sum_{k=1}^{N'} (\bar{g}, \bar{v}_k)_Y\, \bar{v}_k + \bar{g}_0 & , \bar{g}_0 \in Y_{N'}^{\perp} \end{cases}$$

$$Lf = \sum_{k=1}^{N'} (f, u_k)_{\underline{X}} \underbrace{L u_k}_{=\alpha_k \bar{v}_k} = \sum_{k=1}^{N'} \alpha_k (f, u_k)_{\underline{X}}\, \bar{v}_k$$

since $\quad L f_0 = \bar{0} \qquad$ (cf. p. 22 $\quad N(L) = \underline{X}_{N'}^{\perp}$)

$$\Phi_\mu[f] = \| Lf - \bar{g} \|_Y^2 + \mu \| f \|_{\underline{X}}^2 =$$

$$= \sum_{k=1}^{N'} \Big[\big| \alpha_k (f, u_k)_{\underline{X}} - (\bar{g}, \bar{v}_k)_Y \big|^2 + \mu \big| (f, u_k)_{\underline{X}} \big|^2 \Big]$$

$$+ \| \bar{g}_0 \|_Y^2 + \mu \| f_0 \|_{\underline{X}}^2$$

$$= \sum_{k=1}^{N'} \Big[(\alpha_k^2 + \mu) \Big| (f, u_k)_{\underline{X}} - \frac{\alpha_k (\bar{g}, \bar{v}_k)_Y}{\alpha_k^2 + \mu} \Big|^2$$

$$+ |(\bar{g}, \bar{v}_k)_Y|^2 \left(1 - \alpha_k^2 (\alpha_k^2 + \mu)^{-1}\right) \Big]$$

$$+ \| \bar{g}_0 \|_Y^2 + \mu \| f_0 \|_{\underline{X}}^2$$

The functonal $\Phi_\mu[f]$ has a minimum

if one chooses $(f, u_k)_X = (\bar{g}, \bar{v}_k)_Y \, \alpha_k \, (\alpha_k^2 + \mu)^{-1}$

and $f_0 = 0$

The minimizing function f_μ is

$$\boxed{f_\mu = R_\mu \bar{g} = \sum_{k=1}^{N'} \alpha_k (\alpha_k^2 + \mu)^{-1} (\bar{g}, \bar{v}_k)_Y \, u_k}$$

Compare with the defition of spectral windows $W_{\mu,k}$ on p. 57

The Tikhonov regularization algorithm is equivalent to a spectral window

$$W_{\mu,k} = \alpha_k^2 / (\alpha_k^2 + \mu) \qquad \underline{\text{Tikhonov window}}$$

Note the condition a/ and b) on p. 57 are satisfied.

The Tikhonov regularizer for discrete data can always be reduced to a matrix inversion.

Make a variation of $\Phi_\mu[f]$ around f_μ. Define δf by $f = f_\mu + \delta f$

For a minimum $\delta \Phi_\mu[f_\mu + \delta f] = 0$ for all δf

$$\Phi_\mu[f] = (Lf - \bar{g}, Lf - \bar{g})_Y + \mu \|f\|_X =$$

$$= (Lf, Lf)_Y + \|\bar{g}\|_Y^2 - (Lf, \bar{g})_Y - (\bar{g}, Lf)_Y + \mu (f, f)_X$$

$$\Phi_\mu[f_\mu + \delta f] - \Phi_\mu[f_\mu] = (Lf_\mu, L\delta f)_Y - (\bar{g}, L\delta f)_Y + \mu (f_\mu, \delta f)_X$$

$$+ \text{ complex conjugates } + (L\delta f, L\delta f)_Y + \mu (\delta f, \delta f)_X$$

$$= ((L^*L + \mu I) f_\mu - L^* \bar{g}, \delta f)_X + \text{c.c} + \text{2nd order}$$

Minimum if

$$(L^*L + \mu I) f_\mu = L^* \bar{g}$$

$$f_\mu = R_\mu \bar{g}$$

Note that the operator $L^*L+\mu I$ has a bounded inverse on its range $(= \underline{X}_N')$ for every $\mu > 0$.

This follows from N. & S. Th. 5.7.1, p. 244 and the fact that $\tilde{L} = L^*L$ is a non-negative operator, see Lemma p. 31.

Thus $L^*L+\mu I$ is a <u>strictly</u> positive operator, i.e. (N. & S. p. 370)

$$(f, L^*Lf+\mu f)_{\underline{X}} \geq \mu \|f\|^2_{\underline{X}} \qquad \forall f \in \underline{X}$$

and a strictly positive is always bounded away from the origin, i.e.

$$\|(L^*L+\mu I)f\|_{\underline{X}} \|f\|_{\underline{X}} \overset{\text{Schwarz ineq.}}{\geq} |(f,(L^*L+\mu I)f)_{\underline{X}}| \geq \mu \|f\|^2_{\underline{X}}$$

and Th. 5.7.1 in N. & S. applies.

Thus $\quad R_\mu = (L^*L + \mu I)^{-1} L^*$

Operator identity

$$(L^*L + \mu I) L^* = L^*(LL^* + \mu I)$$

$LL^* + \mu I$ is always invertable (see lemma on p. 33 where $\hat{L} = LL^*$ is proven to be a non-negative operator) and we get

$$\boxed{R_\mu = L^*(LL^* + \mu I)^{-1}}$$

Note! Since $LL^* + \mu I$ is an operator on Y (i.e. finite dimensional space), the operator R_μ is a matrix inversion followed by an application of the operator L^*

4.9 Choice of regularization parameter

This section is devoted to the question of choosing the appropriate value of the regularization parameter μ, such that optimal resolution is obtained. Four different methods are reviewed, which apply to the Tikhonov regularizer (the first two can be generalized to other regularization algorithms).

(A.) Constrained least-squares solutions

Again we formulate the problem as a minimization problem ($\bar{g}$ could be with or without noise).

$\| Lf - \bar{g} \|_Y = \text{minimum}$ (solution set: $\tilde{S}$, see p. 25)

However, $\bar{g}$ is usually corrupted with noise. If the underlying physical problem is ill-posed the set of solutions, $\tilde{S}$, contains unphysical solutions, e.g. wildly oscillating functions. To remedy this introduce constraints.

$$\boxed{\begin{array}{c} \| Lf - \bar{g} \|_Y = \text{minimum} \\ \text{subject to } \| f \|_{\underline{X}} \leq E \end{array}}$$

The constraint could be e.g. bounds on the energy f^2 on the solution.

If the generalized solution f^+ satisfies $\| f^+ \|_{\underline{X}} \leq E$ this minimization problem is not uniquely soluble (p. 24).

This is however highly unlikely if $\bar{g}$ is corrupted with noise. (either wrong const. E or no need for regularization)

We may thus assume $\| f^+ \|_{\underline{X}} > E$.

Then the intersection of the set of solutions $\tilde{S}$ and the ball $\| f \|_{\underline{X}} \leq E$ is empty.

The solution to the constraint problem must then satisfy $\| f \|_{\underline{X}} = E$.

With this type of constraints Lagrange multipliers apply, which is equivalent to study

$$\| Lf - \bar{g} \|_Y^2 + \mu \| f \|_X^2 = \text{minimum}$$

with the condition $\| f_\mu \|_X = E$ on the minimizing solution f_μ.

There is a unique μ satisfying $\| f_\mu \|_X = E$ as seen from (see p. 63)

$$\nu^2(\mu) = \| f_\mu \|_X^2 = \sum_{k=1}^{N} \alpha_k^2 (\alpha_k^2 + \mu)^{-2} \, |(\bar{g}, \bar{v}_k)_Y|^2$$

$\nu(\mu)$ is a decreasing function of μ and its value at $\mu = 0$ is $\nu(0) = \| f^+ \|_X > E$, while $\nu(\infty) = 0$

The constrained least-squares solution satisfying $\| f_\mu \|_X = E$ is denoted f_0

So far we made no distinction between clean and noisy data. Now we want to study the effect of noisy data $\bar{g}_\epsilon$. Denote the generalized solution f_ϵ^+ (i.e. no regularizing algorithm applied) and the constrained least-squares solution $f_{0,\epsilon}$ (i.e. Tikhonov regularization algorithm applied and μ chosen so that $\|f_{0,\epsilon}\|_{\underline{X}} = E$)

Furthermore, let f^+ denote the generalized solution for noise free data $\bar{g}$.

Two cases occur as $\bar{g}_\epsilon \to \bar{g}$, $\epsilon \to 0$

1) $\|f^+\| \leq E$ and $\|f_\epsilon^+\| > E$

$f_\epsilon^+ \to f^+$ as $\epsilon \to 0$ (Theorem p. 27)

and $f_{0,\epsilon}^+ \to f^+$ (we do not prove this)

2) $\|f^+\| > E$ (unphysical constraint)

$f_\epsilon^+ \to f^+$ as before, but

$f_{0,\epsilon}^+ \not\to f^+$

B. Minimum norm solution and the discrepancy principle

One drawback with the former method is that it requires the knowledge of an upper bound E on the solution of the problem.

Suppose instead we have bounds on the error of the data. Then define a minimization problem

$$\boxed{\begin{aligned} &\|f\|_X = \text{minimum} \\ &\text{subject to } \|Lf - \bar{g}\|_Y \leq \epsilon \end{aligned}}$$

This is the dual of the former method, and has always a unique solution. This solution is not the zero solution provided $\|g\|_Y > \epsilon$ (more data than noise). Denote the solution corresponding to noisy data $\bar{g}_\epsilon$ as $f_{1,\epsilon}$. One can prove

$$f_{1,\epsilon} \to f^{+} \quad \text{as} \quad \epsilon \to 0.$$

(C.) Solution satisfying prescribed bounds

A combination of both methods above would be

$$\boxed{\begin{aligned} \|Lf - \bar{g}\|_Y &\leq \epsilon \\ \|f\|_X &\leq E \end{aligned}}$$

This solution is not unique.

One solution is

$$f = \sum_{k=1}^{N} \alpha_k \left(\alpha_k^2 + (\epsilon/E)^2 \right)^{-1} (\bar{g}, \bar{v}_k)_Y \, u_k$$

D. Cross-validation

This method can be used if no upper bounds on the error or on the solution are known.

It is based upon the idea of letting the data themselves choose the value of the regularization parameter.

A good value of the regularization parameter should allow the prediction of missing data value.

Let $f_{\mu,k}$ be the minimizer of (Y with euclidean norm)

$$\Phi_{\mu,k}[f] = N^{-1} \sum_{n \neq k} |(Lf)_n - g_n|^2 + \mu \|f\|^2_{\underline{X}}$$

Note that the k:th data point is missing.

Define <u>the generalized cross-validation function $V(\mu)$</u> as

$$\boxed{\begin{aligned} V(\mu) &= \left\{ N^{-1} \operatorname{Tr}[I - A(\mu)] \right\}^{-2} \left(N^{-1} \|[I - A(\mu)]\bar{g}\|^2 \right) \\ A(\mu) &= LL^* [LL^* + \mu I]^{-1} \qquad \text{(a matrix, c.f. p. 66)} \end{aligned}}$$

$V(\mu)$ is invariant with respect to permutations of data.

The value μ that minimizes $V(\mu)$ is the taken as the value of the regularization parameter.

Efficient and fast methods to find this minimum have been developed.

4.10 Iterative methods

Several iterative methods define regularizing algorithms in the sense of Tikhonov.
These iterative methods are methods computing the minimum of the quadratic form

$$\| Lf - \bar{g} \|_Y = \text{minimum}$$

From p. 24 & 26 the solution to this minimum problem is equivalent to the solotions of the normal equations (Euler equation)

$$L^* L f = L^* \bar{g}$$

Note that L^*L is a finite rank operator in $\underline{X}$ (see Lemma p. 31)

<u>Notations:</u> (Assume $N' = N$, so f^+ = normal solution)

- Let f_n denote the nth approximation of the normal solution f^+
- Let r_n denote the residual

$$r_n = L^* L f_n - L^* \bar{g}$$

- Introduce vectors $\bar{f}_n, \bar{r}_n \in Y$

$$\boxed{\begin{aligned} f_n &= L^* \bar{f}_n \\ r_n &= L^* \bar{r}_n \end{aligned}}$$

Then $r_n = L^* L f_n - L^* \bar{g} = L^* \underbrace{(L L^*)}_{= \hat{L}} \bar{f}_n - L^* \bar{g}$

↑ p. 31

So $\boxed{\bar{r}_n = \hat{L} \bar{f}_n - \bar{g}}$

Now define the iteration algorithms

1. Successive approximations

$$\bar{f}_0 = 0 \quad , \quad \bar{f}_{n+1} = \bar{f}_n - \tau \bar{r}_n$$

where τ is a parameter $0 < \tau < 2/\alpha_1^2$

2. Steepest descent

$$\bar{f}_0 = 0 \quad , \quad \bar{f}_{n+1} = f_n - \tau_n \bar{r}_n$$

where $\tau_n = (\hat{L}\bar{r}_n, \bar{r}_n)_Y \,/\, (\hat{L}^2\bar{r}_n, \bar{r}_n)_Y$

3. Conjugate gradient

$$\bar{f}_0 = 0 \quad , \quad \bar{f}_{n+1} = \bar{f}_n - \tau_n \bar{p}_n$$

where $\bar{p}_0 = \bar{r}_0 = -\bar{g}$, $\bar{p}_n = \bar{r}_n + \sigma_{n-1} p_{n-1}$

and $\tau_n = (\hat{L}\bar{r}_n, \bar{p}_n)_Y \,/\, (\hat{L}^2\bar{p}_n, \bar{p}_n)_Y$

$$\sigma_{n-1} = -(\hat{L}^2\bar{r}_n, \bar{p}_{n-1})_Y \,/\, (\hat{L}^2\bar{p}_{n-1}, \bar{p}_{n-1})_Y$$

All three methods have the following properties

1) $f_n \in \overline{X}_N$ (by definition of L^*, p. 23)

$f_n \to f^+$, as $n \to \infty$

2) $\|r_{n+1}\|_{\overline{X}} \leq \|r_n\|_{\overline{X}}$, for all n

$\|f_{n+1} - f^+\|_{\overline{X}} \leq \|f_n - f^+\|_{\overline{X}}$, for all n

3) $\|f_n\|_{\overline{X}} \leq \|f_{n+1}\|_{\overline{X}}$, for all n

The last inequality implies $\|f_n\|_{\overline{X}} \leq \|f^+\|_{\overline{X}}$, $\forall n$

Define a family of operators $\{R_n\}_{n=0}^{\infty}$ by

$$\boxed{f_n = R_n \bar{g}}$$

This family of operators satisfies assumptions i) – iii) on p. 52 and is therefore a regularization algorithm

(take e.g. $\mu = 1/n$, $n > 0$ as regularization parameter)

Check! i) O.K. since $\mathcal{R}(R_n) = \mathcal{R}(L^*) = \bar{X}_N$ (p. 23)

ii) $\|R_n\| = \sup\limits_{\|g\|_Y = 1} \|R_n \bar{g}\|_{\bar{X}} = \sup\limits_{\|g\|_Y = 1} \|f_n\|_{\bar{X}} \leq$

$\leq \sup\limits_{\|g\|_Y = 1} \|f^+\| = \|L^+\|$ (p. 28)

iii) O.K. since $\|f_n\|_{\bar{X}} \to \|f^+\|_{\bar{X}}$

The method of successive approximation

$$\bar{f}_0 = 0 \qquad \bar{f}_{n+1} = \bar{f}_n - \tau \bar{r}_n$$

is equivalent to a spectral window regularizer

To see this, prove by induction that

$$\bar{f}_n = \tau \sum_{k=0}^{n-1} (I - \tau \hat{L})^k \bar{g} \quad , \quad n \geq 1$$

$n = 1$ $\qquad \bar{f}_1 = \tau \bar{g} = \underbrace{\bar{f}_0}_{=0} - \tau \underbrace{\bar{r}_n}_{= -\bar{g}}$ $\qquad$ O.K.

Assume the sum holds for n, then

$$\begin{aligned}
\bar{r}_n = \hat{L} \bar{f}_n - \bar{g} &= \sum_{k=0}^{n-1} (I - \tau \hat{L})^k \tau \hat{L} \bar{g} - \bar{g} \\
&= \sum_{k=1}^{n} (I - \tau \hat{L})^{k-1} \tau \hat{L} \bar{g} - \bar{g} \\
&= -\sum_{k=1}^{n} (I - \tau \hat{L})^{k-1} (I - \tau \hat{L}) \bar{g} \\
&\quad + \sum_{k=0}^{n-1} (I - \tau \hat{L})^k \bar{g} - \bar{g} \\
&= -\sum_{k=0}^{n} (I - \tau \hat{L})^k \bar{g} + \sum_{k=0}^{n-1} (I - \tau \hat{L})^k \bar{g}
\end{aligned}$$

$$\bar{f}_{n+1} = f_n - \tau r_n = \tau \sum_{k=0}^{n} (I - \tau \hat{L})^k \bar{g} \qquad \text{i.e. holds for } n+1$$

Thus, we have

$$f_n = R_n \bar{g} = L^* \bar{f}_n = \tau L^* \sum_{k=0}^{n-1} (I - \tau L L^*)^k \bar{g}$$

$$= \tau \sum_{k=0}^{n-1} (I - \tau L^* L)^k L^* \bar{g}$$

Introduce the singular system for L

$$\bar{g} = \sum_{k=1}^{N} (\bar{g}, \bar{v}_k)_Y \bar{v}_k \quad \text{and} \quad L^* \bar{v}_k = \alpha_k u_k \,, \; L u_k = \alpha_k \bar{v}_k$$

$$f_n = \tau \sum_{m=0}^{n-1} (I - \tau L^* L)^m \sum_{k=1}^{N} \alpha_k (\bar{g}, \bar{v}_k)_Y u_k$$

$$= \tau \sum_{k=1}^{N} \underbrace{\sum_{m=0}^{n-1} (1 - \tau \alpha_k^2)^m}_{[1 - (1-\tau\alpha_k^2)^n] / \tau \alpha_k^2} \alpha_k (\bar{g}, \bar{v}_k)_Y u_k$$

We get

$$\boxed{\begin{array}{l} f_n = \sum_{k=1}^{N} \alpha_k^{-1} [1 - (1 - \tau \alpha_k^2)^n] (\bar{g}, \bar{v}_k)_Y u_k \\ \\ W_{n,k} = [1 - (1 - \tau \alpha_k^2)^n] \end{array}}$$

Note that $W_{n,k} \in [0,1]$ if $0 \leq \tau \leq \alpha_1^{-2}$

4.11 Backus – Gilbert method

On p. 56 we showed that when $\underline{X} = L^2$ any regularized solution can be written as an average over the true solution f

$$f_{reg}(x) = \int A(x,x')\, f(x')\, dx'$$

In general $A(x,x')$ is

$$A(x,x') = \sum_{n,n'=1}^{N} \rho_{\mu}^{nn'}\, \phi_{n'}^{*}(x')\, \phi_n(x)\, w(x)$$

The Backus – Gilbert method assumes

$$\boxed{A(x,x') = \sum_{n=1}^{N} a_n(x)\, \phi_n^{*}(x')}$$

where $a_n(x)$ is determined by a minimizing problem

$$\boxed{\begin{array}{l} \delta^2(x) = \int J(x,x')\, |A(x,x')|^2\, dx' = \text{minimum} \\ \int A(x,x')\, dx' = 1 \end{array}}$$

The function $J(x,x')$ should vanish when $x=x'$ and should increase monotonically ax x goes away from x'.

Ex. $J(x,x') = (x-x')^2$

Let $b_n = \int \phi_n(x)\,dx \qquad S_{nm}(x) = \int J(x,x')\phi_n(x')\phi_m^*(x')\,dx$

Then the minimizing problem becomes

$$\boxed{\begin{aligned} &\delta^2(x) = \sum_{n,m=1}^{N} S_{nm}(x)\, a_n^*(x)\, a_m(x) = \text{minimum} \\ &\sum_{n=1}^{N} b_n^* a_n(x) = 1 \end{aligned}}$$

This is solved by means of the method of Lagrange multipliers.

$$\boxed{a_n(x) = \left[\sum_{n,m=1}^{N} S_{nm}^{-1} b_m b_n^*\right]^{-1} \sum_{n=1}^{N} S_{nm}^{-1} b_m}$$

where S_{nm}^{-1} is the inverse matrix of S_{nm}

The Backus - Gilbert solution is

$$f_{BG}(x) = \int A(x,x')\, f(x')\, dx' = \sum_{n=1}^{N} a_n(x) \int f(x')\, \phi_n^*(x')\, dx'$$

$$= \sum_{n=1}^{N} g_n\, a_n(x)$$

This solution is a type of generalized solution even if it is different from f^+, since

$$f_{BG}(x) \in \operatorname{span}\{a_n(x)\}_{n=1}^{N}$$

$$f^+(x) \in \underline{X}_N = \operatorname{span}\{\phi_n(x)\}_{n=1}^{N}$$