Lecture 20

Continuous Problems

Linear Operators and Their Adjoints

Syllabus

Lecture 01	Describing Inverse Problems
Lecture 02	Probability and Measurement Error, Part 1
Lecture 03	Probability and Measurement Error, Part 2
Lecture 04	The L_2 Norm and Simple Least Squares
Lecture 05	A Priori Information and Weighted Least Squared
Lecture 06	Resolution and Generalized Inverses
Lecture 07	Backus-Gilbert Inverse and the Trade Off of Resolution and Variance
Lecture 08	The Principle of Maximum Likelihood
Lecture 09	Inexact Theories
Lecture 10	Nonuniqueness and Localized Averages
Lecture 11	Vector Spaces and Singular Value Decomposition
Lecture 12	Equality and Inequality Constraints
Lecture 13	L_1 , L_∞ Norm Problems and Linear Programming
Lecture 14	Nonlinear Problems: Grid and Monte Carlo Searches
Lecture 15	Nonlinear Problems: Newton's Method
Lecture 16	Nonlinear Problems: Simulated Annealing and Bootstrap Confidence Intervals
Lecture 17	Factor Analysis
Lecture 18	Varimax Factors, Empircal Orthogonal Functions
Lecture 19	Backus-Gilbert Theory for Continuous Problems; Radon's Problem
Lecture 20	**Linear Operators and Their Adjoints**
Lecture 21	Fréchet Derivatives
Lecture 22	Exemplary Inverse Problems, incl. Filter Design
Lecture 23	Exemplary Inverse Problems, incl. Earthquake Location
Lecture 24	Exemplary Inverse Problems, incl. Vibrational Problems

Purpose of the Lecture

Teach you a *tiny bit* of analysis

enough for you to understand

Linear Operators and their Adjoints

because they are the core technique
used in the so-called

adjoint method of computing data kernels everything we do today

is based on the idea of

generalizing
discrete problems to continuous problems

a function

$m(x)$

is the continuous analog of a vector

$\mathbf{m}$

a function

$m(x)$

simplification:
one spatial
dimension x

is the continuous analog of a vector

$\mathbf{m}$

comparison

$\mathbf{m}$ is of length M

$m(x)$ is infinite dimensional

What is the continuous analog of a matrix $\mathbf{L}$?

We'll give it a symbol, $\mathcal{L}$

and a name, a *linear operator*

Matrix times a vector is another vector

$$\mathbf{b} = \mathbf{L}\,\mathbf{a}$$

so we'll want
linear operator on a function
is another function

$$b(x) = \mathcal{L}\,a(x)$$

Matrix arithmetic is *not* communative

$$\mathbf{L}^{(1)}\,\mathbf{L}^{(2)}\,\mathbf{a} \neq \mathbf{L}^{(2)}\,\mathbf{L}^{(1)}\,\mathbf{a}$$

so we'll not expect that property
for linear operators, either

$$\mathcal{L}^{(1)}\,\mathcal{L}^{(2)}\,a(x) \neq \mathcal{L}^{(2)}\,\mathcal{L}^{(1)}\,a(x)$$

Matrix arithmetic is associative

$$(\mathbf{L}^{(1)}\,\mathbf{L}^{(2)})\,\mathbf{L}^{(3)}\,\mathbf{a} = \mathbf{L}^{(1)}\,(\mathbf{L}^{(2)}\,\mathbf{L}^{(3)}\,)\,\mathbf{a}$$

so well want
that property for linear operators, too

$$(\mathcal{L}^{(1)}\,\mathcal{L}^{(2)}\,)\mathcal{L}^{(3)}\,a(x) = \mathcal{L}^{(1)}\,(\mathcal{L}^{(2)}\,\mathcal{L}^{(3)})\;a(x)$$

Matrix arithmetic is distributive

$$\mathbf{L}\ [\mathbf{a}+\mathbf{b}] = \mathbf{La} + \mathbf{Lb}$$

so well want
that property for linear operators, too

$$\mathcal{L}\,[a(x)+b(x)] = \mathcal{L}\,a(x) + \mathcal{L}\,b(x)$$

Hint to the identity of $\mathcal{L}$

matrices can approximate
derivatives
and integrals

$$\mathbf{L}^{\mathsf{A}} = \frac{1}{\Delta x}\begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ & & \ddots & & & \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{L}^{\mathsf{B}} = \Delta x \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 \\ & & \ddots & & & \\ 1 & 1 & 1 & \cdots & 1 & 1 \end{bmatrix}$$

$$\mathbf{L}^{\mathrm{A}} = \frac{1}{\Delta x}\begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ & & \ddots & & & \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{L}^{\mathrm{B}} = \Delta x \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 \\ & & \ddots & & & \\ 1 & 1 & 1 & \cdots & 1 & 1 \end{bmatrix}$$

$$\mathrm{L}^{\mathrm{A}}\mathrm{a} \approx da/dx \qquad \mathrm{L}^{\mathrm{a}}\mathrm{a} \approx \int_0^x m(x')\,\mathrm{d}x'$$

Linear Operator $\mathcal{L}$

any combination of functions,
derivatives
and integrals

all perfectly good $\mathcal{L}$a(x)'s

$$\mathcal{L}\, a(x) = c(x)\, a(x)$$

$$\mathcal{L}\, a(x) = \mathrm{d}a/\mathrm{d}x$$

$$\mathcal{L}a(x) = b(x)\mathrm{d}a/\mathrm{d}x + c(x)\,\mathrm{d}^2a/\mathrm{d}x^2$$

$$\mathcal{L}\, a(x) = \int_0^x a(\xi)d\xi$$

$$\mathcal{L}\, a(x) = f(x)\int_0^\infty a(\xi)g(x,\xi\,)\, d\xi$$

What is the continuous analog of the inverse $\mathbf{L}^{-1}$ of a matrix $\mathbf{L}$?

call it $\mathcal{L}^{-1}$

Problem
$\mathbf{L}^{\mathrm{A}}$ not square, so has no inverse

$$\mathbf{L}^{\mathrm{A}} = \frac{1}{\Delta x}\begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ & & \ddots & & & \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}$$

derivative determines a function only up to an additive constant

Patch by adding top row that sets the constant

$$\mathbf{L}^{\mathrm{C}} = \frac{1}{\Delta x}\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ & & \ddots & & & \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}$$

Now $\mathbf{L}^{\mathrm{B}}\,\mathbf{L}^{\mathrm{C}} = \mathbf{I}$

lesson 1: $\mathcal{L}$ may need to include *boundary conditions*

lesson 2:

if $\mathcal{L} = \mathrm{d}/\mathrm{d}x$ then $\mathcal{L}^{-1} = \int_0^x \mathrm{d}x'$

since $m(x) = \frac{\mathrm{d}}{\mathrm{d}x}\int_0^x m(x')\,\mathrm{d}x'$

the analogy to the matrix equation

$$\mathbf{L}\,\mathbf{m} = \mathbf{f}$$

and its solution

$$\mathbf{m} = \mathbf{L}^{-1}\,\mathbf{f}$$

is the differential equation

$$\mathcal{L}\,m = f$$

and its Green function solution

$$m(x) = \int_{-\infty}^{+\infty} F(x,\xi) f(\xi) \mathrm{d}\xi = \mathcal{L}^{-1} f(x)$$

so the inverse to a differential operator $\mathcal{L}$

is the Green function integral

$$\mathcal{L}^{-1}\, a(x) = \int_{-\infty}^{+\infty} F(x,\xi)\, a(\xi)\, d\xi$$

where F solves

$$\mathcal{L}\, F(x,\xi) = \delta(x-\xi)$$

What is the continuous analogy to a dot product ?

$$a^T b = \Sigma_i \, a_i \, b_i$$

The continuous analogy to a dot product

$$s = a^T b = \Sigma_i\, a_i\, b_i$$

is the *inner product*

$$s = \int_{-\infty}^{+\infty} a(x) b(x) \mathrm{d}x = (a, b)$$

squared length of a vector

$$|\mathbf{a}|^2 = \mathbf{a}^{\mathrm{T}}\mathbf{a}$$

squared length of a function

$$|a|^2 = (a,a)$$

important property of a dot product

$$(\mathbf{La})^{\mathrm{T}}\,\mathbf{b} = \mathbf{a}^{\mathrm{T}}\,(\mathbf{L}^{\mathrm{T}}\mathbf{b})$$

important property of a dot product

$$(\mathbf{La})^{\mathrm{T}}\,\mathbf{b} = \mathbf{a}^{\mathrm{T}}\,(\mathbf{L}^{\mathrm{T}}\mathbf{b})$$

what is the continuous analogy ?

$$(\mathcal{L}a, b) = (a, ?b)$$

in other words ...

what is the continuous analogy of the transpose of a matrix?

$$(\mathcal{L}a, b) = (a, ?b)$$

by analogy, it must be another linear operator since transpose of a matrix is another matrix in other words ...

what is the continuous analogy of the transpose of a matrix?

$$(\mathcal{L}a, b) = (a, ?b)$$

give it a name "adjoint" and a symbol $\mathcal{L}^\dagger$

so …

$$(\mathcal{L}a, b) = (a, \mathcal{L}^{\dagger} b)$$

so, given $\mathcal{L}$, how do you determine $\mathcal{L}^\dagger$?

so, given $\mathcal{L}$, how do you determine $\mathcal{L}^\dagger$?

various ways ,,,

the adjoint of a function is itself

$$\int_{-\infty}^{+\infty} (ca)\, b \,\mathrm{d}x = \int_{-\infty}^{+\infty} a\,(cb)\,\mathrm{d}x$$

if $\mathcal{L}=c(x)$ then $\mathcal{L}^{\dagger}=c(x)$

the adjoint of a function is itself

$$\int_{-\infty}^{+\infty} (ca)\, b \, \mathrm{d}x = \int_{-\infty}^{+\infty} a \, (cb) \, \mathrm{d}x$$

if $\mathcal{L} = c(x)$ then $\mathcal{L}^{\dagger} = c(x)$

a function is *self-adjoint*

the adjoint of a function is itself

$$\int_{-\infty}^{+\infty} (ca)\, b\, \mathrm{d}x = \int_{-\infty}^{+\infty} a\, (cb)\, \mathrm{d}x$$

if $\mathcal{L} = c(x)$ then $\mathcal{L}^{\dagger} = c(x)$

self-adjoint operator anagous to a symmetric matrixx

the adjoint of d/dx
(with zero boundary consitions)
is −d/dx

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}a}{\mathrm{d}x}\, b\, \mathrm{d}x = ab\Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} a\, \frac{\mathrm{d}b}{\mathrm{d}x}\, \mathrm{d}x$$

if $\mathcal{L}=d/dx$ then $\mathcal{L}^{\dagger}=-d/dx$

the adjoint of d/dx
(with zero boundary consitions)
is −d/dx

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}a}{\mathrm{d}x}\, b\, \mathrm{d}x = ab\Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} a\, \frac{\mathrm{d}b}{\mathrm{d}x}\, \mathrm{d}x$$

integration by parts

if $\mathcal{L}=d/dx$ then $\mathcal{L}^{\dagger}=-d/dx$

the adjoint of $\mathrm{d}^2/\mathrm{d}x^2$ is itself

apply integration by parts twice

if $\mathcal{L} = \mathrm{d}^2/\mathrm{d}x^2$ then $\mathcal{L}^\dagger = \mathrm{d}^2/\mathrm{d}x^2$

a function is *self-adjoint*

the adjoint of $\int_{-\infty}^{x} \mathrm{d}x$ is $\int_{x}^{+\infty} \cdot\, \mathrm{d}x$

$$(\mathcal{L}a, b) = \int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{x} a(\xi)\, \mathrm{d}\xi \right\} b(x)\, \mathrm{d}x =$$

trick using Heaviside step function

$$\int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{+\infty} H(x - \xi)\, a(\xi)\, \mathrm{d}\xi \right\} b(x)\, \mathrm{d}x =$$

$$\int_{-\infty}^{+\infty} a(\xi) \left\{ \int_{-\infty}^{+\infty} H(x - \xi) b(x) \mathrm{d}x \right\} \mathrm{d}\xi =$$

$$\int_{-\infty}^{+\infty} a(\xi) \left\{ \int_{\xi}^{+\infty} b(x) \mathrm{d}x \right\} \mathrm{d}\xi = (a, \mathcal{L}^{\dagger} b)$$

properties of adjoints

$$(\mathcal{L}^{\dagger})^{\dagger} = \mathcal{L} \quad \text{and} \quad (\mathcal{L}^{-1})^{\dagger} = (\mathcal{L}^{\dagger})^{-1}$$

$$(\mathcal{L}^{A}+\mathcal{L}^{B})^{\dagger} = (\mathcal{L}^{B})^{\dagger}+(\mathcal{L}^{A})^{\dagger} \quad \text{and} \quad (\mathcal{L}^{A}\mathcal{L}^{B})^{\dagger} = (\mathcal{L}^{B})^{\dagger}(\mathcal{L}^{A})^{\dagger}$$

table of adjoints

$c(x)$	$c(x)$
$\mathrm{d}/\mathrm{d}x$	$-\mathrm{d}/\mathrm{d}x$
d^2/dx^2	d^2/dx^2
$\int_{-\infty}^{x} \mathrm{d}x$	$\int_{x}^{+\infty} \cdot \, \mathrm{d}x$

analogies

$\mathbf{m}$	$m(x)$
$\mathbf{L}$	$\mathcal{L}$
$\mathbf{Lm}=\mathbf{f}$	$\mathcal{L}m(x)=f(x)$
$\mathbf{L}^{-1}$	$\mathcal{L}^{-1}$
$\mathbf{f}=\mathbf{L}^{-1}\mathbf{m}$	$f(x)=\mathcal{L}^{-1}f(x)$
$\mathbf{s}=\mathbf{a}^{T}\mathbf{b}$	$s=(a(x), b(x))$
$(\mathbf{La})^{T}\mathbf{b}=\mathbf{a}^{T}(\mathbf{L}^{T}\mathbf{b})$	$(\mathcal{L}a, b)=(a, \mathcal{L}^{\dagger}b)$
$\mathbf{L}^{T}$	$\mathcal{L}^{\dagger}$

how is all this going to help us?

step 1

recognize that
standard equation of inverse theory

$$d_i = \int_a^b G_i(z)\, m(z)\, \mathrm{dz}$$

is an inner product

$$d_i = (G_i, m)$$

step 2

suppose that we can show that

$$d_i = (h_i, \mathcal{L}m)$$

then do this

$$d_i = (\mathcal{L}^\dagger h_i, m)$$

so

$$G_i = \mathcal{L}^\dagger h_i$$

step 2

suppose that we can show that

$$d_i = (h_i, \mathcal{L}m)$$

then do this

$$d_i = (\mathcal{L}^\dagger h_i, m)$$

so

$$G_i = \mathcal{L}^\dagger h_i$$

formula for the data kernel