Engsci 721 Inverse Problems and Learning From Data

Oliver Maclaren (oliver.maclaren@auckland.ac.nz)

Lecture 1: Overview

Problem 1.

Consider the polynomial regression model

$$y = a + bx + cx^2 + dx^3$$

for given data $(x_1, y_1), (x_2, y_2), ..., (x_m, y_m)$. Show that this can be expressed as a *linear* forward model. What is the model linear with respect to?

Problem 2.

Suppose you are working on a problem involving an intergral equation and real world data with your supervisor 'Joe'. He says to you that there exists a unique solution to the integral equation and hence there is no need for 'ugly regularisation methods and all that'. You go away and invert your integral equation to find the desired solution using the measured data but you get a terrible looking solution. You check your solution method using a known solution and noise-free data and everything seems fine. Briefly explain (e.g. to Joe!) what is likely going wrong and why regularisation might be necessary. Mention Hadamard to impress Joe.

Problem 3.

Write down matrices to represent discrete versions of the following vector-to-vector operations

- Cumulative integration (summation)
- Differentiation (finite differencing)
- Local (window) averaging

Write down an expression representing the operation of taking the expected value as a *scalar-valued* mapping acting on a signal vector.

Problem 4.

Give a simple constructive example illustrating the instability of differentiation. Hint: see lecture 1!

Lecture 2: Inverses 1

Problem 5.

Suppose you have a 'tall' linear system with more rows than columns and that the columns are linearly independent. Explain in what sense you can 'solve' this system of equations. What characteristics does the corresponding solution possess? Which of Hadamard's conditions does this help with? Which doesn't this approach help with? What type of 'inverse' have we constructed: a left or right inverse?

Problem 6.

Conversely, suppose you have a 'wide' linear system with more columns than rows and that the rows are linearly independent. Explain in what sense you can 'solve' this system of equations. What characteristics does the corresponding solution possess? Which of Hadamard's conditions does this help with? Which doesn't this approach help with? What type of 'inverse' have we constructed: a left or right inverse?

Lecture 3: Inverses 2

Problem 7.

Suppose that you have a linear problem of the form

$$Ax = y.$$

Consider the under- and over-determined cases respectively. You may suppose that right and/or left) inverses R and L exist as needed and as appropriate. Describe the effect of the operators

$$RA$$

and

$$AL$$

What are these operators called? What are their domains and ranges (i.e. what spaces do they map between)? State and prove a key property of each that they share with the identity matrix.

Problem 8.

What are the two most general algebraic properties that *generalised* inverses of satisfy? What properties do these ensure that AA^+ and A^+A possess for generalised inverse A^+ (hint: proj...)? State two other properties that *pseudo*-inverses satisfy. What is the interpretation of these properties (hint: orth...)?

Problem 9.

What optimisation problems does a pseudo-inverse solve?

Problem 10.

Suppose you have an underdetermined linear system with linearly independendent rows. Define a constrained optimisation problem to 'solve' this problem. How does this relate to the pseudo inverse?

Problem 11.

Give general expressions for the data and model resolution operators and state conditions for each to be the identity.

Lecture 4: Linearity

Problem 12.

Consider a mathematical model of the form $F(x) = y$ where F is the 'forward' mapping, $x \in \mathbb{R}^n$ is a vector of length n, and y is a vector of length m. Suppose that F is a *linear* forward model, i.e.

$$F(ax_1 + bx_2) = aF(x_1) + bF(x_2)$$

for arbitrary vectors x_1, x_2 and scalars a, b. Show by construction that $F(x)$ can hence be written in the form

$$F(x) = Ax$$

where A is an $m \times n$ matrix. Hint: first take the standard basis for $\mathbb{R}^n$, i.e. $\{e_1, e_2, ..., e_n\}$ where each e_i is a vector of length n with a one in the ith component and zeros everywhere else. Then consider an arbitrary x written in terms of this basis, $x = x_1e_1 + x_2e_2 + ...$ where x_i are scalars and e_i are vectors. Then consider the action of F on x. Finally, recall the definition of matrix-vector multiplication in terms of linear combinations of columns to help you construct A

Problem 13.

Carry out the matrix multiplication

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 4 & 1 \end{bmatrix}$$

in what we called:

- dot product form

- partial row/column dot product forms
- outer product form

and hence verify the answers are the same.

Problem 14.

Recall the key property of 'vec' for the product of three matrices:

$$\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B)$$

Use this to convert the matrix equation

$$Ax = b$$

into a matrix equation that can be solved for $\text{vec}(A)$, i.e that allows us to treat A (in the form $\text{vec}(A)$) as the unknown given x and b.

Illustrate with A a 2 by 3 matrix.

Problem 15.

Suppose we have an image stored as a matrix A. We can represent independent vertical and horizontal blurring with the multiplication:

$$B_c A B_r^T$$

for appropriate matrices B_c, B_r.

Show this can be represented in the form

$$M\text{vec}(A)$$

i.e. as a single matrix operating on A in vec form.

Problem 16.

Matrices can be thought of as representing linear functions from vectors to vectors. They can also be thought of as representing *bilinear* functions from vectors to numbers, by defining

$$B(x, y) = x^T Ay.$$

This is a linear function in each variable *while holding the other constant*, but is nonlinear overall when both can vary (i.e. is a 'quadratic form').

- Using 'vec', show how to write this as a (linear) dot product of two (large) vectors (i.e. a bilinear form is 'linear in a different space').

Lecture 5: Matrix calculus

Problem 17.

Starting from the least squares problem for an overdetermined linear system

$$\min_x ||y - Ax||^2$$

derive the normal equations

$$A^T Ax = A^T y$$

using matrix calculus. You may use the 'three key rules' from lecture 5 without proof.

Problem 18.

Use the basic properties of differentials listed in lecture 5 to derive the three key rules for derivatives of scalar/vector functions of scalar/vector variables.

Problem 19.

Define

$$\phi(x) = x^T Ax$$

- Find the first differential then take the differential of this to find $d^2\phi = d(d\phi)$.
- Write this in the form $d^2\phi = dx^T B dx$ for some matrix B and hence identify the so-called Hessian matrix $H\phi$ by using the rule $H\phi = (B + B^T)/2$ for any such B. This represents the quadratic term in an approximation of $\phi(x)$.

Problem 20.

For extra credit, derive the derivative matrix of the function $F(X) = X^{-1}$.