Extra Exercises

for

Modeling and Analysis Principles for Chemical and Biological Engineers

Michael D. Graham
Department of Chemical and Biological Engineering
University of Wisconsin-Madison
Madison, Wisconsin

James B. Rawlings
Department of Chemical Engineering
University of California, Santa Barbara
Santa Barbara, California

August 27, 2019

Nob Hill Publishing

Madison, Wisconsin

The extra exercises were set in Lucida using LaTeX.

Nob Hill Publishing, LLC
Cheryl M. Rawlings, publisher
Santa Barbara, CA 93117
orders@nobhillpublishing.com
http://www.nobhillpublishing.com

The extra exercises are intended for use by students and course instructors.

This document has been posted electronically on the website: www.chemengr.ucsb.edu/~jbraw/principles.

1

Linear Algebra

Exercise 1.79: Using the SVD to solve engineering problems

Consider the system depicted in Figure 1.12 in which we can manipulate an input $u \in \mathbb{R}^3$ to cancel the effect of a disturbance $d \in \mathbb{R}^2$ on an output $y \in \mathbb{R}^2$ of interest. The steady-state relationship between the variables is modeled as a linear relationship

$$y = Gu + d$$

and y, u, d are in deviation variables from the steady state at which the system was linearized. Experimental tests on the system have produced the following model parameters

$$G = \begin{bmatrix} U \end{bmatrix} \begin{bmatrix} \Sigma & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$$

$$U = \begin{bmatrix} -0.71 & -0.71 \\ -0.71 & 0.71 \end{bmatrix} \qquad \Sigma = \begin{bmatrix} 1.21 & 0.00 \\ 0.00 & 0.08 \end{bmatrix} \qquad V_1 = \begin{bmatrix} -0.98 & 0.11 \\ 0.20 & 0.84 \\ -0.098 & 0.54 \end{bmatrix}$$

If we have measurements of the disturbance d available, we would like to find the input u that exactly cancels d's effect on y, and we would like to know ahead of time what is the worst-case disturbance that can hit the system.

(a) Can you use u to exactly cancel the effect of d on y for *all* d? Why or why not?

(b) In terms of U, Σ, V_1, V_2, and d, what are all the inputs u that minimize the effect of d on y?

(c) What is the smallest input u that minimizes the effect of d on y?

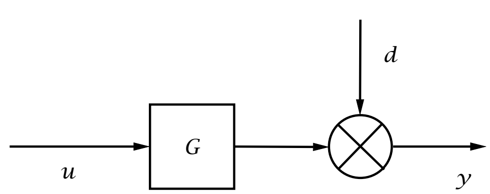

Figure 1.12: Manipulated input u and disturbance d combine to affect output y.

(d) What is the worst-case disturbance d for this process, i.e., what d of unit norm requires the *largest* response in u? What is the response u to this worst d?

(e) What is the best-case disturbance d for this process, i.e., what disturbance d of unit norm requires the *smallest* response u. What is the response u for this best d?

Exercise 1.80: Functions of matrices

I recall that two solutions to the linear, scalar differential equation $\ddot{x} = x$ are $\cosh t$ and $\sinh t$. Say the boundary conditions for this differential equation are the values of x and its first derivative at $t = 0$, $x(0) = x_0$ and $\dot{x}(0) = \dot{x}_0$.

(a) What is the full solution (in terms of sinh and cosh) to the scalar differential equation

$$\ddot{x} = mx$$

with these initial conditions and parameter $m > 0$?

(b) Therefore, what is the solution to the set of n coupled differential equations

$$\ddot{x} = Mx \qquad x(0) = x_0 \quad \dot{x}(0) = \dot{x}_0$$

in which $x, x_0, \dot{x}_0 \in \mathbb{R}^n$ and $M \in \mathbb{R}^{n \times n}$ and matrix M has positive eigenvalues. Hint: make sure that all matrix-vector multiplications are defined.

(c) Explain how you would evaluate this solution numerically. You may assume M has distinct eigenvalues.

Exercise 1.81: Concepts related to eigenvalues, SVD, and the fundamental theorem

Consider a general complex-valued matrix $A \in \mathbb{C}^{m \times n}$.

(a) Establish the relationship between the eigenvalues of A^*A and AA^*. For example: (i) how many eigenvalues does each matrix have? (ii) are these eigenvalues equal to each other? (iii) all of them? etc. Do not use the SVD factorization of A in your derivation because this result is used to establish the SVD factorization.

(b) Show that the following two null spaces are the same: $N(A^*A) = N(A)$. Hint: first show the simple result that $x \in N(A)$ implies $x \in N(A^*A)$ so $N(A)$ is contained in $N(A^*A)$. Next show that $x \in N(A^*A)$ implies that $x \in N(A)$ to complete the derivation. You may want to use the fundamental theorem of linear algebra for this second step.

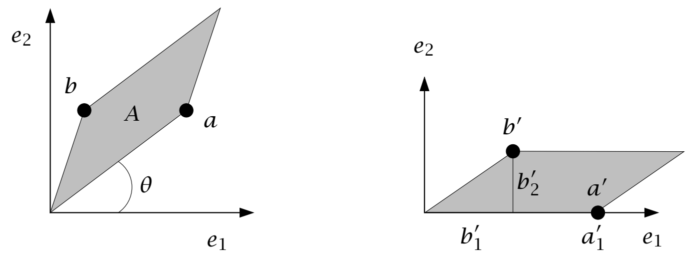

Figure 1.13: Parallelogram and rotation to align with x_1 axis.

Exercise 1.82: Determinant of a matrix, area, and volume

(a) Given a parallelogram with sides $a, b \in \mathbb{R}^2$ depicted in Figure 1.13, show that its area is given by the formula

$$A = |\det C| \qquad C = \begin{bmatrix} a & b \end{bmatrix}$$

In other words, form a partitioned matrix by placing the column vectors representing the two sides next to each other, and take the absolute value of the determinant of that matrix to calculate the area. Note that the absolute value is required so that area is positive.
Hint: to reduce the algebra, first define the rotation matrix that rotates the parallelogram so that the a side is parallel to the x_1-axis; then use the fact that the area of a parallelogram is the base times the height.

(b) Show that the result generalizes to finding the volume of a parallelepiped with sides given by $a, b, c \in \mathbb{R}^3$

$$V = |\det C| \qquad C = \begin{bmatrix} a & b & c \end{bmatrix}$$

(c) How does the result generalize to n dimensions.

Note that the determinant without absolute value is then called the *signed* area (volume). Signed area is positive when rotation of side a into side b is counter-clockwise as shown in Figure 1.13, and negative when the rotation is clockwise (right-hand rule, same sign as cross product).

Exercise 1.83: Pseudoinverse

We have already seen two different forms of the pseudoinverse arise when solving in the least-squares sense

$$Ax = b$$

with the singular value decomposition in Section 1.4.7. The different forms motivate a more abstract definition of pseudoinverse that is general enough to cover all these cases. Consider the following definition

Definition 1.1 (Pseudoinverse). Let $A \in \mathbb{C}^{m \times n}$. A matrix $A^\dagger \in \mathbb{C}^{n \times m}$ satisfying

$$(1)\ AA^\dagger A = A \qquad (2)\ A^\dagger A A^\dagger = A^\dagger \qquad (3)\ (AA^\dagger)^* = AA^\dagger \qquad (4)\ (A^\dagger A)^* = A^\dagger A$$

is called a pseudoinverse of A.

(a) Show that $(A^*)^\dagger = (A^\dagger)^*$ by direct substitution into the four properties of the definition.

(b) Show that (if it exists) $A^\dagger$ is unique for all A. Hint: let both B and C satisfy the properties of the pseudoinverse. First show that $AB = AC$ and that $BA = CA$. Given these, show that $B = C$.

Obtaining uniqueness of the pseudoinverse is the motivation for properties 3 and 4 in the definition.

(c) Let $A \neq 0$ and consider the SVD, $A = U\Sigma V^*$. Show that

$$A^\dagger = V_1 \Sigma_r^{-1} U_1^* \qquad A \neq 0 \tag{1.38}$$

in which $r \geq 1$ is the rank of A, satisfies the definition of the pseudoinverse.

(d) From the definition and inspection, what is a pseudoinverse for $A = 0$? Therefore we have shown that the pseudoinverse exists and is unique for all A, and we have a means to calculate it using the SVD.

(e) Show that the pseudoinverse reduces to $A^\dagger = (A^*A)^{-1}A^*$ when A has linearly independent columns, and that it reduces to $A^\dagger = A^*(AA^*)^{-1}$ when A has linearly independent rows.

Exercise 1.84: Minimum-norm, least-squares solution to $Ax = b$

(a) Show that the solution

$$x^0 = A^\dagger b$$

with the pseudoinverse defined in Exercise 1.83 is the unique minimum-norm least-squares solution to $Ax = b$ for arbitrary A, b. This result covers *all* versions of the least-squares problem.

(b) When A has neither linearly independent columns nor linearly independent rows, find the set of *all* least-squares solutions.

Exercise 1.85: Pseudoinverse and Kronecker product

Establish the formula

$$(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$$

which generalizes (1.24) in the text.

Exercise 1.86: Solution to the general matrix equation $AXB = C$

Show that the unique minimum-norm X that minimizes $\|AXB - C\|_F$ is given by

$$X^0 = A^\dagger C B^\dagger$$

Here we are using the square root of the sum of the squares of the elements in the matrix as the matrix norm. This norm is called the Frobenius norm (see also Exercise 1.90). This result covers *all* versions of the matrix least-squares problem.

Exercise 1.87: Some linear algebra questions and the singular value decomposition (SVD)

I have an $A \in \mathbb{R}^{12\times 12}$ and I call Octave's or MATLAB's `[U,S,V]=svd(A)` function and obtain the return arguments shown on the next page.

Answer the following questions.

(a) What is the rank of A? What is $\|A\|_2$?

(b) What are the smallest dimension matrices that you can multiply together to obtain A? Clearly mark and label these matrices as parts of the U, S, V, matrices printed on the next page.

(c) Is there a solution x to $Ax = b$ for every $b \in \mathbb{R}^{12}$? How do you know?

(d) Say I happen to choose a b such that $Ax = b$ has a solution. Is this solution unique? How do you know?

(e) Does the solution to the optimization problem $\min_x \|Ax - b\|_2^2$ exist for every right-hand side b? Is this solution unique? How do you know?

(f) What is $A^\dagger$ in terms of the three matrices you identified part (b)? What do you call this matrix?

(g) How can you calculate the smallest solution to $\min_x \|Ax - b\|_2^2$ given a b? Is this solution unique?

```
U =

  -0.4328  -0.0409  -0.3260   0.2898  -0.0375   0.2669  -0.0471   0.0107  -0.5901  -0.4301   0.0769   0.0815
  -0.2593   0.5397  -0.2572  -0.0451  -0.0347   0.0161   0.0883   0.0042   0.3733  -0.2694  -0.5661  -0.1779
  -0.3013   0.1062   0.1712   0.5850   0.3381  -0.0406   0.4231   0.0140   0.3349   0.0650   0.3365   0.0395
  -0.2345   0.0860  -0.1324  -0.2227   0.4816  -0.2773   0.0806   0.1276  -0.3850   0.3865  -0.0405  -0.4895
  -0.2267   0.0518   0.4199  -0.4300   0.4074   0.3383   0.1524  -0.2094  -0.1005  -0.0433  -0.1801   0.4391
  -0.3732  -0.4287  -0.4275  -0.3879  -0.0369  -0.3253   0.1659   0.0834   0.3058  -0.0743   0.1558   0.2805
  -0.2056   0.4478  -0.1227  -0.0921  -0.2727  -0.1315  -0.1419  -0.6237  -0.0594   0.3056   0.3442   0.1320
  -0.3213   0.0349   0.4643  -0.3101  -0.4018   0.0897   0.1180   0.1767   0.0393  -0.2405   0.3237  -0.4527
  -0.3199  -0.3116   0.1391   0.1287   0.2369   0.0435  -0.7090  -0.2495   0.2899  -0.0421  -0.0673  -0.2268
  -0.3137  -0.3351   0.1398   0.2182  -0.4261   0.0718   0.2340  -0.1007  -0.0910   0.4807  -0.4809  -0.0023
  -0.1626   0.2053  -0.1698  -0.0586  -0.0508   0.4846  -0.2808   0.5557   0.1496   0.4404   0.1626   0.1773
  -0.1905   0.2204   0.3511   0.1205  -0.0882  -0.6021  -0.2889   0.3622  -0.1693  -0.0583  -0.1288   0.3791

S =
```

```
  2.6e+01         0         0         0         0         0         0         0         0         0         0         0
        0   2.4e+00         0         0         0         0         0         0         0         0         0         0
        0         0   1.9e+00         0         0         0         0         0         0         0         0         0
        0         0         0   9.8e-01         0         0         0         0         0         0         0         0
        0         0         0         0   9.0e-01         0         0         0         0         0         0         0
        0         0         0         0         0   4.2e-01         0         0         0         0         0         0
        0         0         0         0         0         0   3.6e-01         0         0         0         0         0
        0         0         0         0         0         0         0   2.1e-01         0         0         0         0
        0         0         0         0         0         0         0         0   4.8e-02         0         0         0
        0         0         0         0         0         0         0         0         0   1.0e-15         0         0
        0         0         0         0         0         0         0         0         0         0   3.5e-16         0
        0         0         0         0         0         0         0         0         0         0         0   8.0e-17

V =

  -0.4328  -0.0409  -0.3260   0.2898  -0.0375   0.2669  -0.0471   0.0107  -0.5901   0.0736   0.1188   0.4220
  -0.2593   0.5397  -0.2572  -0.0451  -0.0347   0.0161   0.0883   0.0042   0.3733   0.6505  -0.0258   0.0309
  -0.3013   0.1062   0.1712   0.5850   0.3381  -0.0406   0.4231   0.0140   0.3349  -0.3280   0.1022   0.0311
  -0.2345   0.0860  -0.1324  -0.2227   0.4816  -0.2773   0.0806   0.1276  -0.3850   0.0460   0.2453  -0.5730
  -0.2267   0.0518   0.4199  -0.4300   0.4074   0.3383   0.1524  -0.2094  -0.1005   0.0276  -0.4263   0.2114
  -0.3732  -0.4287  -0.4275  -0.3879  -0.0369  -0.3253   0.1659   0.0834   0.3058  -0.1975  -0.1298   0.2293
  -0.2056   0.4478  -0.1227  -0.0921  -0.2727  -0.1315  -0.1419  -0.6237  -0.0594  -0.4583  -0.0529  -0.1284
  -0.3213   0.0349   0.4643  -0.3101  -0.4018   0.0897   0.1180   0.1767   0.0393  -0.0370   0.6028   0.0534
  -0.3199  -0.3116   0.1391   0.1287   0.2369   0.0435  -0.7090  -0.2495   0.2899   0.1488   0.1678  -0.0863
  -0.3137  -0.3351   0.1398   0.2182  -0.4261   0.0718   0.2340  -0.1007  -0.0910   0.2288  -0.3892  -0.5084
  -0.1626   0.2053  -0.1698  -0.0586  -0.0508   0.4846  -0.2808   0.5557   0.1496  -0.3689  -0.2213  -0.2584
  -0.1905   0.2204   0.3511   0.1205  -0.0882  -0.6021  -0.2889   0.3622  -0.1693   0.0090  -0.3483   0.2057
```

Exercise 1.88: Solutions of linear differential equations with repeated eigenvalues

We want to solve a set of linear differential equations

$$\frac{d}{dt}x = Ax \qquad x(0) = x_0$$

with $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times n}$

(a) What is the solution to this differential equation? Give the Taylor series for evaluating the exponential of a matrix e^C.

To help us express the solution, we performed an eigenvalue decomposition and found $Q, \Lambda \in \mathbb{C}^{n \times n}$ such that

$$A = Q\Lambda Q^{-1}$$

(b) Substitute this expression for A into the Taylor series of your solution and express the solution in terms of t, Λ, Q, and Q^{-1}.

Unfortunately, it appears that our A is not diagonalizable because Λ has the form

$$\Lambda = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$

(c) Perform the indicated multiplications on Λ and write out the following: Λ^2, Λ^3, Λ^4.

(d) How does this result generalize for Λ^k, $k \geq 1$?

(e) Substitute these results into the Taylor series applied to $e^{\Lambda t}$ and obtain a series expansion in terms of λ. What function of λ and t appear on the diagonal in $e^{\Lambda t}$?

(f) Simplify the sum of terms in your series expansion and show what functions of λ and t appear above the diagonal in $e^{\Lambda t}$?

(g) OK, based on your work so far, let's make a conjecture. What functions will appear in the matrix $e^{\Lambda t}$ when you have p repeated eigenvalues λ in $\Lambda_p \in \mathbb{C}^{p \times p}$

$$\Lambda_p = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{bmatrix} \qquad e^{\Lambda_p t} = ?$$

(Note that missing entries denote zeroes.)

Exercise 1.89: Changing basis for a linear space

Let V be a linear space of dimension n with a set of basis vectors $\{b_i\}_{i=1}^n$. Consider an arbitrary set of n-linearly independent vectors $\{a_i\}_{i=1}^n$. Show that the set $\{a_i\}_{i=1}^n$ is also a basis for V.

Exercise 1.90: The Frobenius norm of a matrix, the trace, and the vec operator

(a) Let $A, B \in \mathbb{C}^{m \times n}$. Show that

$$\mathrm{tr}(A^* B) = \mathrm{vec}(A)^* \mathrm{vec}(B)$$

(b) Let $X \in \mathbb{C}^{m \times n}$. Show that

$$\|X\|_F^2 = \mathrm{tr}(X^* X) = (\mathrm{vec} X)^* \mathrm{vec} X = \|\mathrm{vec} X\|_2^2$$

2

Ordinary Differential Equations

Exercise 2.80: Changing variables of integration

(a) For the two-dimensional integral, derive the formula

$$\int_{\mathbb{A}_x} g(x_1, x_2)dx_1 dx_2 = \int_{\mathbb{A}_y} \overline{g}(y_1, y_2) \left|\det(\partial x/\partial y)\right| dy_1 dy_2$$

in which the variable transformation is $y = f(x)$, which is assumed invertible, $\overline{g}(y) = \overline{g}(y(x)) = g(x)$, $\partial x/\partial y$ is the Jacobian matrix of the transformation, and the area of integration in the y-variables is given by

$$\mathbb{A}_y = \{(y_1, y_2) \mid (y_1, y_2) = f(x_1, x_2), (x_1, x_2) \in \mathbb{A}_x\}$$

Hint: consider the rectangular element of integration shown in Figure 2.35 with area $dx_1 dx_2$, and find its area under the transformation f; then use the properties of the determinant derived in Exercise 1.82.

(b) How does this result generalize to three-dimensional integrals? n-dimensional integrals?

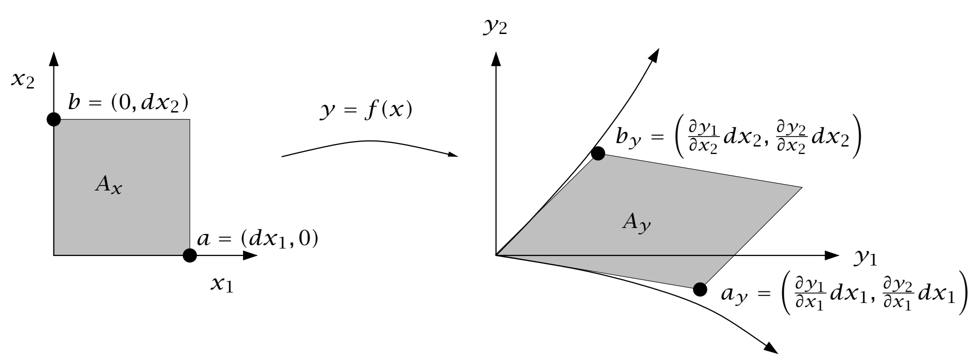

Figure 2.35: Change of area element due to coordinate transformation

Exercise 2.81: Existence and uniqueness of a nonhomogeneous boundary-value problem

Consider the second-order nonhomogeneous boundary-value problem

$$Lu = f \qquad B_1 u = 0 \qquad B_2 u = 0 \tag{2.101}$$

for unknown function $u(x)$, $x \in [0,1]$, in which $f(x)$ is given. The differential operator and two boundary functionals are defined by

$$\begin{aligned} Lu(x) &= \frac{d^2}{dx^2}u(x) + k^2 u(x) \\ B_1 u(x) &= u(0) \\ B_2 u(x) &= u(1) \end{aligned}$$

(a) Find the adjoint operator and boundary functionals L^*, B_1^*, B_2^* so that $(Lu, v) = (u, L^* v)$ for all $u(x), v(x)$ satisfying $B_1 u = 0$, $B_2 u = 0$, $B_1^* v = 0$, and $B_2^* v = 0$.

Is the boundary-value problem self adjoint?

(b) Find the null space of L, $N(L)$, for this boundary-value problem. Find also $N(L^*)$.

(c) Given your results, what does the alternative theorem say about the existence and uniqueness of the *nonhomogeneous* BVP 2.101? In particular, does the existence of the solution depend on the form of $f(x)$?

Exercise 2.82: More forcing on the boundary

We may wonder how general is the technique to move nonhomogeneity from the boundary conditions to the differential equation using impulses. To explore this issue, consider the fully nonhomogeneous second-order BVP for $u(x)$, with $x \in [0,1]$

$$Lu = f \qquad B_1 u = \gamma_1 \qquad B_2 u = \gamma_2$$

with the general form

$$Lu = u_{xx} + \rho_1 u_x + \rho_0 u$$

$$\begin{bmatrix} B_1 u \\ B_2 u \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & B_{13} & B_{14} \\ B_{21} & B_{22} & B_{23} & B_{24} \end{bmatrix} \begin{bmatrix} u(0) \\ u(1) \\ u_x(0) \\ u_x(1) \end{bmatrix}$$

and we assume that the rows of the B matrix are linearly independent so the boundary conditions are well posed for a second-order BVP.

(a) The first task is to define the adjoint problem. Find operator L^* and $J(u,v)$ so that

$$\langle Lu, v \rangle = \langle u, L^* v \rangle + J(u,v)\big|_0^1$$

(b) Define the following vectors to contain the boundary information

$$\mathbf{u} = \begin{bmatrix} u(0) & u(1) & u_x(0) & u_x(1) \end{bmatrix}^T \qquad \mathbf{v} = \begin{bmatrix} v(0) & v(1) & v_x(0) & v_x(1) \end{bmatrix}^T$$

and find matrix M such that

$$J(u,v)\,\big|_0^1 = \mathbf{v}^T M \mathbf{u}$$

What is the rank of matrix M?

(c) We define the adjoint boundary conditions so that $J(u,v)\ |_0^1 = 0$ for all $\mathbf{u}$ satisfying $B\mathbf{u} = 0$ and $\mathbf{v}$ satisfying $\widetilde{B}\mathbf{v} = 0$.[1]

Equivalently, we are saying that $\mathbf{v}^T M \mathbf{u} = 0$ for all $\mathbf{u} \in N(B)$ and $\mathbf{v} \in N(\widetilde{B})$. Let the columns of matrix N_B be a basis for $N(B)$, the null space of B, and let the columns of matrix $N_{\widetilde{B}}$ be a basis for $N(\widetilde{B})$, the null space of $\widetilde{B}$.

What is the rank and dimension of matrix N_B? Justify your answer.

(d) We then have that $\mathbf{u} \in N(B)$ and $\mathbf{v} \in N(\widetilde{B})$ are of the form $N_B\alpha$ and $N_{\widetilde{B}}\beta$, respectively, for arbitrary vectors α, β. So we have the condition

$$\beta^T (N_{\widetilde{B}}^T M N_B)\alpha = 0 \qquad \text{for all } \alpha, \beta$$

Show that this requirement implies

$$N_{\widetilde{B}}^T M N_B = 0 \tag{2.102}$$

(e) So we can compute the adjoint BC $N_{\widetilde{B}}$ matrix by defining its columns to be a basis for $N((MN_B)^T)$. In MATLAB or Octave,

```
N_Btilde = null((M*N_B)')
```

What is the rank and dimension of matrix $N_{\widetilde{B}}$ given the ranks of M and N_B. Justify your answer.

(f) With the linear algebra preliminaries out of the way, we are ready to move the nonhomogeneity into the differential equation. We consider adding a term to the ODE of the form

$$p(x) = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \end{bmatrix} \begin{bmatrix} \delta(x) \\ \delta(x-1) \\ -\delta_x(x) \\ -\delta_x(x-1) \end{bmatrix}$$

[1]We use matrix $\widetilde{B}$ to denote the adjoint's boundary functional matrix in place of B^* to avoid conflicting with the notation for the adjoint (conjugate, transpose) of matrix B.

where we consider adding singlets and doublets to both ends of the interval $[0, 1]$. Note the minus sign on the last two elements.

Show that this $p(x)$ produces the following integral

$$\langle p, v \rangle = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \end{bmatrix} \begin{bmatrix} v(0) \\ v(1) \\ v_x(0) \\ v_x(1) \end{bmatrix}$$

or

$$\langle p, v \rangle = \mathbf{p}^T \mathbf{v} = \mathbf{v}^T \mathbf{p}$$

(g) The solvability condition is $\langle f, v_k \rangle = J(u, v_k)\big|_0^1 = \mathbf{v}_k^T M \mathbf{u}$ where $v_k(x)$ is any solution to the fully homogeneous adjoint problem, and $u(x)$ is the solution to the nonhomogeneous problem. To make $\hat{f} = f + p$ orthogonal to v_k, show that we require that

$$\mathbf{v}_k^T (M\mathbf{u} + \mathbf{p}) = 0$$

(h) Since $u(x)$ satisfies the nonhomogeneous problem, $B\mathbf{u} = \gamma$. Show that all $\mathbf{u}$ satisfying this restriction are given by $\mathbf{u} = B^\dagger \gamma + N_B \alpha$ with α arbitrary. Since $v_k(x)$ satisfies the homogeneous adjoint problem, it satisfies $\mathbf{v}_k = N_{\tilde{B}} \beta$ with β arbitrary. Therefore, show that $\mathbf{p}$ satisfying

$$\mathbf{p} = -MB^\dagger \gamma$$

satisfies the orthogonality equation of the previous part. Is this perturbation unique? Justify your answer.

We have therefore found a perturbation $p(x)$ using impulses at the boundaries for any second-order nonhomogeneous problem. The same procedure works for any nth-order nonhomogeneous BVP with n linearly independent nonhomogeneous boundary conditions.

Exercise 2.83: Forcing on the boundary without tears

For problems arising in applications, we usually do not face the completely general nonhomogeneous BVP of the last exercise. For most of the common BVPs we can deduce the forcing on the boundary by inspection. To see this consider boundary conditions of the form

$$B\mathbf{u} = \gamma \qquad B = \begin{bmatrix} B_{11} & B_{12} & B_{13} & 0 \\ 0 & B_{22} & B_{23} & B_{24} \end{bmatrix}$$

If we ever see two columns with zeros in different rows, we can create the forcing term by inspection. First move the fourth column to the second column so that we have

$$B = \begin{bmatrix} B_{11} & 0 & B_{12} & B_{13} \\ 0 & B_{24} & B_{22} & B_{23} \end{bmatrix}$$

Note that we have now also switched the boundary data to

$$u = \begin{bmatrix} u(0) & u_x(1) & u(1) & u_x(0) \end{bmatrix}$$

in the boundary condition $B\mathbf{u} = \gamma$. From the first row of B, when we change to $Bu = 0$, we require $B_{11}u(0^+) = \gamma_1$, a jump in $u(0)$ of size γ_1/B_{11}. From the second row of B, we require a jump in $u_x(1^+)$ of size $-\gamma_2/B_{24}$. Since these jumps handle the nonhomogeneity in the BCs, we can set $u(1) = u_x(0) = 0$ for convenience. Next we examine what the selected jumps do to the differential operator L

$$\begin{aligned} u_{xx} &= (\gamma_1/B_{11})\delta_x(x) - (\gamma_2/B_{24})\delta(x-1) \\ u_x &= (\gamma_1/B_{11})\delta(x) \end{aligned}$$

So the total change to the operator $Lu = u_{xx} + \rho_1 u_x + \rho_0 u$ is

$$Lu = f + \frac{\gamma_1}{B_{11}}(\delta_x(x) + \rho_1\delta(x)) - \frac{\gamma_2}{B_{24}}\delta(x-1)$$

Check that the solvability condition for the original nonhomogeneous problem is equivalent to the orthogonality relation $\langle \hat{f}, v_k\rangle = 0$ with modified $\hat{f} = f + p$ with p defined above.

Exercise 2.84: Amaze your friends; homogenous BVPs by inspection

Given the general second-order nonhomogeneous, $u_{xx} + \rho_1 u_x + \rho_0 u = f$, $B_1 u = \gamma_1$, $B_2 u = \gamma_2$, consider the following special cases that commonly arise in applications. Give ODE perturbations $p(x)$ for $Lu = f + p$ that provide orthogonal solvability conditions $\langle f + p, v_k\rangle = 0$.

(a) $\rho_1 = 0, B_1 u = u(0), B_2 u = u(1)$.

(b) $\rho_1 = 0, B_1 u = u_x(0), B_2 u = u_x(1)$.

(c) $\rho_1 = 0, B_1 u = u(0) - u(1), B_2 u = u_x(0) + u_x(1)$.

3

Vector Calculus and Partial Differential Equations

Exercise 3.56: Brief transport discussion

(a) In polar coordinates (r, θ), calculate $\nabla \cdot \mathbf{x}$ in which $\mathbf{x}$ is the position vector, $\mathbf{x} = r\mathbf{e}_r$.

(b) Show that you obtain the same result for $\nabla \cdot \mathbf{x}$ if you operate in Cartesian coordinates (x, y) with $\mathbf{x} = x\mathbf{e}_x + y\mathbf{e}_y$. How do you suppose this result generalizes to an n-dimensional space?

(c) Starting from a statement of conservation of moles in a volume element, derive the continuity equation

$$\frac{\partial}{\partial t} c_A = -\nabla \cdot c_A \mathbf{v}_A + R_A$$

in which c_A is the molar concentration of species A, $c_A\mathbf{v}_A$ is the molar flux vector of species A, and R_A is the molar rate/volume of generation of species A. What is a common model for the flux of species A in a multicomponent fluid with average velocity $\mathbf{v}$?

Exercise 3.57: A sphere with prescribed heat flux at its surface

A sphere of radius R, initially at uniform temperature T_0, is placed in a uniform radiation field that delivers constant heat flux F_0 normal to the surface of the sphere. We wish to find the transient temperature response of the sphere.

(a) Starting with the general energy balance

$$\rho \hat{C}_p \frac{\partial}{\partial t} T = \nabla \cdot k \nabla T$$

reduce the model as far as possible in spherical coordinates for the conditions given in the problem statement. State appropriate boundary and initial conditions.

(b) Using the following dimensionless variables

$$\tau = \frac{k}{\rho \hat{C}_p R^2} t \qquad \xi = \frac{r}{R} \qquad \Theta = \frac{T - T_0}{T_0}$$

nondimensionalize the PDE and IC and BCs. How many dimensionless parameters are there in this problem?

(c) Take the Laplace transform of your PDE, IC, and BCs and show that the transform satisfies the second-order ODE

$$\frac{1}{\xi^2}\frac{d}{d\xi}(\xi^2\frac{d\overline{\Theta}}{d\xi}) - s\overline{\Theta} = 0$$

Choose two linearly independent solutions to this equation that allow you to eliminate one of the solutions easily using the BCs.

(d) Solve the ODE, apply the two BCs, and show that

$$\overline{\Theta}(\xi, s) = \frac{\beta}{\xi}\frac{\sinh\sqrt{s}\xi}{s(\sqrt{s}\cosh\sqrt{s} - \sinh\sqrt{s})}$$

in which β is a dimensionless heat flux.

(e) To invert the transform, we need to find and classify the order of the singularities of $\overline{\Theta}(s)$. Obviously there is a singularity at $s = 0$ because of the s term in the denominator.

Now for the other term. Using the substitution $\sqrt{s} = i\alpha$, for what α is $\sqrt{s}\cosh\sqrt{s} - \sinh\sqrt{s} = 0$? Draw a sketch of these α_n, $n = 1, 2, 3, \ldots$ roots. You should find an infinite number of them. These are all simple zeros.

But notice that $\alpha = 0$, which implies $s = 0$, is a zero of this term as well, so the zero at $s = 0$ is *second order.*

(f) So the structure of the inverse of the transform is

$$\Theta(\xi, \tau) = a_{00}\tau + a_{01} + \sum_{n=1}^{\infty} a_n e^{-\alpha_n^2\tau}$$

where the first two terms come from the double zero at $s = 0$. Find the a_n, $n \geq 1$ using the Heaviside theorem since these are simple zeros.

(g) Find a_{00} and a_{01} using the Heaviside theorem for repeated roots for the root at $s = 0$.

(h) Given the appearance of this linear τ term, what is the final value of the temperature? Discuss the physical significance of this result, and critique the model we are using.

Exercise 3.58: Steady-state heat conduction in two dimensions.

Consider a square thin plate of side length ℓ. We are interested in finding the steady-state temperature profile $T(x, y)$ in response to an arbitrary heat generation rate $f(x, y)$ within the body. The four sides of the plate are maintained at steady temperature T_0.

(a) Write down the steady-state energy balance considering heat conduction and heat generation in the solid. Note that this will be a partial differential equation. State the boundary conditions for this PDE.

(b) Nondimensionalize your problem and choose variables so that the boundary conditions are homogeneous. Denote the nondimensional T, x, y variables as Θ, ξ, η, respectively.

State the nondimensional PDE and BCs for this problem. How many nondimensional parameters appear in the problem?

(c) Take the Laplace transform in the ξ variable. Show that the ordinary differential equation that arises for the transform is

$$\begin{gathered} \overline{T}_{\eta\eta} + s^2\overline{T} = \overline{g}(s,\eta) \\ \overline{T}(s,0) = 0 \qquad \overline{T}(s,1) = 0 \end{gathered} \tag{3.105}$$

What is $\overline{g}(s,\eta)$ for this problem?

(d) Solve the ODE for the transform, and show that the solution is

$$\overline{T}(s,\eta) = \frac{1}{s}\int_0^1 \overline{G}(s,\eta,\eta')\overline{g}(s,\eta')d\eta'$$

with

$$\overline{G}(s,\eta,\eta') = \begin{cases} -\frac{\sin(s\eta')\sin s(1-\eta)}{\sin s}, & \eta' < \eta \\ -\frac{\sin(s\eta)\sin s(1-\eta')}{\sin s}, & \eta < \eta' \end{cases}$$

Note that you may find it helpful to review Example 2.15 where it is shown that (in our variable names)

$$\sinh s(\eta-\eta') - \frac{\sinh(s\eta)\sinh s(1-\eta')}{\sinh s} = -\frac{\sinh(s\eta')\sinh s(1-\eta)}{\sinh s}$$

Without rederiving anything, how do you know that the result above holds also for sin replacing sinh?

(e) Using the inverse transform of Exercise 3.59, show that the inverse of the Green's function transform is

$$G(\xi,\eta,\eta') = 2\sum_{n=1}^{\infty} \sin(n\pi\eta')\sin(n\pi\eta)\cosh(n\pi\xi)$$

(f) Invert $\overline{T}(s,\eta)$ and show that

$$\Theta(\xi,\eta) = \int_0^1\int_0^{\xi} G(\xi',\eta,\eta')\Theta_{\xi}(0,\eta')d\xi'd\eta' + \int_0^1\int_0^{\xi}\int_0^{\xi'} G(\xi'-\xi'',\eta,\eta')F(\xi'',\eta')d\xi''d\xi'd\eta'$$

Substitute the above expression for G, switch the order of ξ' and ξ'' integrals in the second term, and then perform the ξ' integrals in both terms to obtain

$$\Theta(\xi,\eta) = 2\sum_{n=1}^{\infty}\frac{1}{n\pi}\Big[a_n\sinh(n\pi\xi)+ \\ \int_0^1\int_0^{\xi}\sin(n\pi\eta')\sinh(n\pi(\xi-\xi'))F(\xi',\eta')d\xi'd\eta'\Big]\sin(n\pi\eta) \quad (3.106)$$

with the Fourier coefficients of the flux defined as

$$a_n = \int_0^1 \Theta_\xi(0,\eta')\sin(n\pi\eta')d\eta'$$

(g) Evaluate at $\xi = 1$ and use the boundary condition to show that

$$0 = \sum_{n=1}^{\infty}\frac{1}{n\pi} \\ \left[\sinh(n\pi)a_n + \int_0^1\int_0^1\sin(n\pi\eta')\sinh(n\pi(1-\xi'))F(\xi',\eta')d\eta'd\xi'\right]\sin(n\pi\eta)$$

Since the right-hand side of this equation is the Fourier series of the zero function for $\eta \in [0,1]$, its coefficients must all be zero giving

$$a_n = \frac{-1}{\sinh(n\pi)}\int_0^1\int_0^1\sin(n\pi\eta')\sinh(n\pi(1-\xi'))F(\xi',\eta')d\eta'd\xi'$$

(h) Substitute this result for a_n into (3.106) and show that the solution can be expressed as

$$\Theta(\xi,\eta) = 2\sum_{n=1}^{\infty}\frac{1}{n\pi}\int_0^1\int_0^1 G_n^{\xi}(\xi,\xi')G_n^{\eta}(\eta,\eta')F(\xi',\eta')d\eta'd\xi'$$

with

$$G_n^{\eta}(\eta,\eta') = \sin(n\pi\eta)\sin(n\pi\eta')$$

$$G_n^{\xi}(\xi,\xi') = \begin{cases} -\frac{\sinh(n\pi\xi')\sinh(n\pi(1-\xi))}{\sinh n\pi}, & \xi' < \xi \\ -\frac{\sinh(n\pi\xi)\sinh(n\pi(1-\xi'))}{\sinh n\pi}, & \xi < \xi' \end{cases}$$

Notice that both G_n^{η} and G_n^{ξ} are symmetric functions.

Exercise 3.59: Laplace transform inverse.

Invert the following transform

$$\overline{f}(s) = \frac{\sin(as)\sin(bs)}{\sin s}$$

Hint: use the Heaviside expansion theorem. Note that the Heaviside expansion theorem does not strictly apply to this $\overline{f}(s)$. Why not? See Exercise 3.60 for a way around this technical difficulty.

Exercise 3.60: The fine print.

We are stretching the Laplace transform pretty hard to cover Exercise 3.58, so let's check that the approach is providing a valid solution.

(a) First of all, an infinite sum of increasing exponentials does not appear to converge, so let's check that the $G_n(\xi, \xi')$ is well defined in the limit $n \to \infty$.

$$G_n^{\xi}(\xi, \xi') = \begin{cases} -\frac{\sinh(n\pi\xi')\sinh(n\pi(1-\xi))}{\sinh n\pi}, & \xi' < \xi \\ -\frac{\sinh(n\pi\xi)\sinh(n\pi(1-\xi'))}{\sinh n\pi}, & \xi < \xi' \end{cases}$$

What is $\lim_{n\to\infty} G_n^{\xi}(\xi, \xi')$?

(b) Since $G_n^{\xi}(\xi, \xi')$ *is* bounded as $n \to \infty$, we can probably justify the use of the Laplace transform. The first serious problem is that we blithely applied the Heaviside expansion theorem to

$$\overline{f}(s) = \frac{\sin(as)\sin(bs)}{\sin s}$$

even though $1/\sin(s)$ has poles at $s = \pm n\pi, n = 0, 1, 2, \ldots$, and we cannot choose a positive constant c such that all the poles of $1/\sin(s)$ are to the left of the line $\mathrm{Re}(s) = c$. And we see the problem when we try to invert anyway and obtain

$$f(t) = 2\sum_{n=1}^{\infty}(-1)^n \sin(n\pi a)\sin(n\pi b)\cosh(n\pi t)$$

which is not (absolutely) convergent for any t, even *finite* t, because of the increasing orders of the cosh terms!

So we have to fix this problem. First consider the Mittag-Leffler expansion of $1/\sin(s)$

$$\frac{1}{\sin s} = \sum_{n=-\infty}^{\infty}\frac{(-1)^n}{s - n\pi}$$

and truncate this series for $|n| > M$ and define

$$\frac{1}{\sin_M(s)} = \sum_{n=-M}^{M}\frac{(-1)^n}{s - n\pi}$$

and we have that $\lim_{M\to\infty} 1/\sin_M(s) = 1/\sin(s)$ at all $s \neq n\pi, n = 0, \pm1, \pm2, \ldots$. Next replace $1/\sin(s)$ with $1/\sin_M(s)$, resolve Exercise 3.59, and show that

$$\overline{f}_M(s) = \frac{\sin(as)\sin(bs)}{\sin_M(s)} \qquad f_M(t) = 2\sum_{n=1}^{M}(-1)^n \sin(n\pi a)\sin(n\pi b)\cosh(n\pi t)$$

which is now well defined for all $t \in [0, \infty)$ and finite integer $M > 0$.

(c) Notice what changes when you resolve the rest of Exercise 3.58 using $1/\sin_M(s)$ in place of $1/\sin(s)$. Show that

$$\Theta_M(\xi, \eta) = 2\sum_{n=1}^{M}\frac{1}{n\pi}\int_0^1\int_0^1 G_n^{\eta}(\eta, \eta')G_n^{\xi}(\xi, \xi')F(\xi', \eta')d\xi' d\eta'$$

Finally take the limit as $M \to \infty$ to solve the original problem and obtain the result stated in Exercise 3.58.

Exercise 3.61: Two forms of the steady-state heat conduction solution

Review the eigenfunction expansion approach to the steady-state heat conduction problem presented in Example 3.7.

(a) Show that another form of the solution to Exercise 3.58 is

$$\Theta(\xi, \eta) = -\frac{4}{\pi^2}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{1}{n^2+m^2} \left[\int_0^1\int_0^1 F(\xi', \eta')\sin(m\pi\xi')\sin(n\pi\eta')d\xi' d\eta'\right]\sin(m\pi\xi)\sin(n\pi\eta)$$

Notice that this solution has a pleasing symmetry in ξ and η that is missing from the previous solution. On the other hand, this solution requires a double summation, which is computationally more expensive than the previous solution.

Let's establish that these two solutions are equivalent.

(b) Equate the two forms of the solution and show that they are the same if and only if

$$\frac{2}{n\pi}G_n^{\xi}(\xi, \xi') = -\frac{4}{\pi^2}\sum_{m=1}^{\infty}\frac{1}{n^2+m^2}\sin(m\pi\xi')\sin(m\pi\xi)$$

which must hold for all $n \geq 1$ and $\xi, \xi' \in [0, 1]$.

(c) Consider ξ' a constant parameter and ξ a variable. Show that for the right-hand side to be a Fourier series representation of the left-hand side, the coefficients must satisfy

$$-\frac{n}{\pi(n^2+m^2)}\sin(m\pi\xi') = \int_0^1 G_n^{\xi}(\xi, \xi')\sin(m\pi\xi)d\xi \tag{3.107}$$

which must hold for all $n, m \geq 1$ and all $\xi' \in [0, 1]$.

(d) Given $G_n(\xi, \xi')$ from the previous exercise

$$G_n^{\xi}(\xi, \xi') = \begin{cases} -\frac{\sinh(n\pi\xi')\sinh(n\pi(1-\xi))}{\sinh n\pi}, & \xi' < \xi \\ -\frac{\sinh(n\pi\xi)\sinh(n\pi(1-\xi'))}{\sinh n\pi}, & \xi < \xi' \end{cases}$$

perform the indicated integral and verify that the coefficients satisfy 3.107.

Hints: Before jumping in, it pays to have a few integrals available. Verify the following pair of integrals

$$\int_0^a \sinh(n\pi x)\sin(m\pi x)dx = \frac{1}{\pi(n^2+m^2)}\left[-m\cos(m\pi a)\sinh(n\pi a) + n\sin(m\pi a)\cosh(n\pi a)\right]$$

$$\int_a^1 \sinh(n\pi(1-x))\sin(m\pi x)dx = \frac{(-1)^m}{\pi(n^2+m^2)}\left[m\cos(m\pi(1-a))\sinh(n\pi(1-a)) - n\sin(m\pi(1-a))\cosh(n\pi(1-a))\right]$$

The first can be derived by noting that

$$\sinh(n\pi x)\sin(m\pi x) = -\mathrm{Im}(\cos((n+im)\pi x))$$

and the second by noting that

$$\sinh(n\pi x)\sin(m\pi(1-x) = -\mathrm{Im}((-1)^n\cos((-n+im)\pi(1-x)))$$

(e) Choose a few values of n and ξ' and plot $G_n^{\xi}(\xi, \xi')$ versus ξ and its Fourier series. How many terms in the series are required for an accurate reconstruction of G_n^{ξ}?

Which of the two solutions do you prefer and why?

Exercise 3.62: Transient heat conduction

In Exercise 2.81, you showed that the following steady-state heat conduction problem with forcing term (heat addition rate $-f(x)$) has a unique solution provided $k \neq n\pi, n = 1, 2, \ldots$

$$\frac{d}{dx^2}u(x) + k^2u(x) = f(x) \qquad u(0) = u(1) = 0$$

One of your colleagues is wondering how we can show analytically whether the steady-state solution is stable?

So consider the transient problem with no external heating, $f(x) = 0$, for which we know that a steady state exists for *all* k,

$$\frac{\partial}{\partial t}u(x,t) = \frac{\partial^2}{\partial x^2}u(x,t) + k^2 u(x,t)$$

The initial and boundary conditions are

$$u(x,t) = \begin{cases} u_0(x), & t = 0, x \in (0,1) \\ 0, & t > 0, x = 0 \\ 0, & t > 0, x = 1 \end{cases}$$

with arbitrary initial condition function $u_0(x)$.

(a) Take the Laplace transform of the PDE and BCs and show that the transform satisfies

$$\frac{d^2}{dx^2}\overline{u}(x,s) + (k^2 - s)\overline{u}(x,s) = -u_0(x) \qquad (3.108)$$
$$\overline{u}(0,s) = \overline{u}(1,s) = 0$$

(b) Given the arbitrary $u_0(x)$ initial condition, let's solve this differential equation using Fourier series. Given the form of the boundary conditions, we choose the complete, orthogonal set $\{\sin n\pi x\}, n = 1, 2, \ldots$ since these satisfy the boundary conditions at $x = 0, 1$. Denote the Fourier coefficients of the solution as $a_n(s)$ and the initial condition as t_n, so that

$$\overline{u}(x,s) = \sum_{n=1}^{\infty} a_n(s) \sin n\pi x \qquad u_0(x) = \sum_{n=1}^{\infty} t_n \sin n\pi x$$

Provide a formula for calculating the $t_n, n = 1, 2, \ldots$.

(c) Double check that $\overline{u}(x,s)$ satisfies the boundary conditions for all choices of $a_n(s)$. Substitute these expansions for $\overline{u}(x,s)$ and $u_0(x)$ into (3.108) and solve for $a_n(s)$.

(d) With $\overline{u}(x,s)$ determined, where are its singularities? Invert the transform to obtain $u(x,t)$.

(e) For what k values is the steady state stable? What is the steady-state solution?

4

Probability, Random Variables, and Estimation

Exercise 4.60: PLS regression versus minimum norm regression

The text lists the PLS regression formula as

$$B_{\text{PLS}} = RT^T Y$$

and mentions that this is a (nonunique) least-squares solution to the regression problem after replacing X with the low rank (rank ℓ) approximation $X_\ell = TP^T$

$$Y = X_\ell B_{\text{PLS}} \tag{4.97}$$

Consider the SVD of X_ℓ written as

$$X_\ell = \widetilde{U}_\ell \widetilde{\Sigma} \widetilde{V}_\ell^T$$

(a) Work out the dimensions of matrices $\widetilde{U}_\ell$, $\widetilde{\Sigma}_\ell$, and $\widetilde{V}_\ell$.

(b) What is the minimum-norm least-squares solution B_{MLS} of (4.97) in terms of the SVD matrices? Is this solution unique?

(c) For the data in Example 4.23, compare the estimates B_{PLS} and B_{MLS} for all $\ell = 1, 2, \dots, q$. Check that both are least-squares solutions. Check that indeed the norm of B_{PLS} is larger than B_{MLS} for all ℓ. Is it much larger?

Exercise 4.61: Estimating mean from samples, vector case

Consider random variable $\xi \in \mathbb{R}^p$ with mean m and variance matrix P. Let $\hat{x}_n$ be an estimator of the mean of ξ based on n samples of ξ, $x_1, x_2, \dots, x_n$. Say we have worked out the mean and variance of the particular estimator, and they are

$$\mathcal{E}(\hat{x}_n) = m' \qquad \text{var}(\hat{x}_n) = P'$$

(a) What is the bias of this estimator?

(b) Calculate the expectation of the square of the 2-norm of the estimate error $\mathcal{E}(\|\hat{x}_n - m\|^2)$. State this result in terms of the bias and variance of the estimator. This metric of estimate error allows us, for example, to quantify the tradeoff between bias and variance in estimation.

Hint: start by expressing $\|\hat{x}_n - m\|^2 = (\hat{x}_n - m)^T(\hat{x}_n - m) = ((\hat{x}_n - m') - (m - m'))^T((\hat{x}_n - m') - (m - m'))$. Expand the quadratic and then take expectation. Also recall that $x^T x = \mathrm{tr}(xx^T)$.

(c) For the sample-average estimator, what are m' and P'? What is the expectation of the square of the norm of the estimate error using the sample-average estimator?

(d) Consider taking just one of the samples, x_i, as the estimator of the mean. What is the expectation of the square of the norm of the estimate error using the single-sample estimator?

Exercise 4.62: The limiting value of $x^T(xx^T)^{-1}x$

Consider a nonzero vector $x \in \mathbb{R}^p$. We would like to find the limiting value of $x^T(xx^T)^{-1}x$.

(a) Show that $(xx^T)^{-1}$ does not exist (for $p \geq 2$) by showing that $\mathrm{rank}(xx^T) = 1$.

(b) Therefore consider the full rank matrix $xx^T + \rho I$ and take the limit as $\rho \to 0$ to show that

$$\lim_{\rho \to 0} x^T(xx^T + \rho I)^{-1}x = 1$$

Hint: Consider the SVDs of x and xx^T.

(c) Next consider a collection of $n \geq 1$ x_i vectors, not all zero, and show that

$$\mathrm{rank}\Big(\sum_{i=1}^{n} x_i x_i^T\Big) = r$$

where r is the number of linearly independent x_i vectors. Then show that

$$\lim_{\rho \to 0} \sum_{j=1}^{n} x_j^T \left(\Big(\sum_{i=1}^{n} x_i x_i^T \Big) + \rho I \right)^{-1} x_j = r$$

Hint: Place the vectors in matrix $X = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$ and consider the SVDs of X and the product XX^T.

Exercise 4.63: Maximum likelihood estimate of mean and variance of a normal

Let random variable $x \in \mathbb{R}^p$ be distributed normally $x \sim N(m, P)$. Let $x_i, i = 1, 2, \ldots, n$ denote $n \geq 1$ samples of x.

(a) Compute the maximum likelihood estimate of m and P, denoted $\hat{m}$, $\hat{P}$. Compare your result to the one stated in Theorem 4.22.

(b) How large does n need to be for the maximum likelihood problem to have a solution? What happens, for example, for a single sample, $n = 1$?

Hint: consider the result derived in Exercise 4.62.

(c) Assume that the mean is known to be zero, i.e., $x \sim N(0, P)$. How large does n need to be for the maximum likelihood problem for estimating variance (only) to have a solution?

(d) Take the expectations of the estimates $\hat{m}$ and $\hat{P}$. Are these unbiased? Explain why or why not.

Exercise 4.64: A maximum likelihood estimation problem

Consider a model

$$y = X\theta + e$$

with $y \in \mathbb{R}^p$ a vector of measured responses that are linear in parameter vector $\theta \in \mathbb{R}^{n_p}$ with measurement error e modeled as a random variable distributed as $e \sim N(0, R)$ with *known* positive definite variance R.

(a) Given this density of e, what is the density of measurement y as a function of model parameter (not RV) θ.

(b) One of your colleagues recalls from a class he took as a graduate student at Wisconsin, that if you maximize the probability of the measurements, $p_y(y; \theta)$, over parameter θ, you are also solving a weighted least-squares problem. What is the equivalent weighted least-squares problem that you are solving here?

(c) Solve this least-squares problem and find the maximum likelihood estimate of parameter θ, denoted $\hat{\theta}$. What have you assumed about matrix X when solving this problem.

(d) Given your result above, find the residual of the model fit $r = y - X\hat{\theta}$, assuming that the measurements come from the linear model with some true model parameter value θ_0, i.e., $y = X\theta_0 + e$.

(e) Show that the following matrix B is a projection operator

$$B = I - X(X^T R^{-1} X)^{-1} X^T R^{-1}$$

(f) Finally compute the expectation of the square of the norm of the residual for the maximum likelihood estimate of θ.

$$\mathcal{E}(r^T r)$$

Given e's distribution, what is $\mathcal{E}(e^T e)$? Which is smaller, $\mathcal{E}(r^T r)$ or $\mathcal{E}(e^T e)$? What does your result reduce to if $R = \sigma^2 I_p$, which is the first estimation problem solved in the text in Section 4.7.1 with the number of measurements p replacing the number of samples n of the scalar measurement.

Exercise 4.65: Least squares with deficient model matrix

Complications arise when solving least-squares problems with rank deficient matrices. Consider the following problem

$$\min_{\theta} \, (y - X\theta)^T H (y - X\theta) \qquad H > 0 \quad \text{any } X$$

Here we keep a positive definitive Hessian, $H \in R^{p\times p} > 0$, but allow any $X \in \mathbb{R}^{p\times n_p}$, and do not assume that the columns (or even rows) are linearly independent as in the standard least-squares problem.

(a) Show that the set of all solutions to this least-squares problem is

$$\hat{\theta} = \left(V_1 \Sigma_r^{-1} (U_1^T H U_1)^{-1} U_1^T H \right) y + V_2 \alpha_2$$

with free parameter $\alpha_2 \in \mathbb{R}^{n_p - r}$, $r = \operatorname{rank}(X)$, and U and V are defined by the partitioned SVD of X as follows

$$X = USV^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_r & \\ & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$$

(b) What does this solution reduce to when X has linearly independent columns, $r = n_p$?

(c) What is the minimum norm solution, $\hat{\theta}_{\mathrm{mn}}$?

Exercise 4.66: Least squares with deficient variance matrix

Complications also arise in least-squares estimation for semidefinite variance of the measurement error. Consider the following problem

$$\min_{\theta} \, (y - X\theta)^T R^{-1} (y - X\theta) \qquad R \geq 0 \quad \text{any } X$$

with semidefinite matrix, $R \in R^{p\times p} \geq 0$, and arbitrary $X \in \mathbb{R}^{p\times n_p}$.

(a) Show that the solution exists if and only if the data y satisfy the condition

$$\widetilde{y}_2 \in \text{ range of } \widetilde{X}_2$$

with $\widetilde{y}_2 = U_2^T y$, $\widetilde{X}_2 = U_2^T X$, and the U matrix is defined in the partitioned (symmetric) SVD of matrix R

$$R = USU^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_r & \\ & 0 \end{bmatrix} \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix}$$

and r is the rank of R.

How would you numerically check that the data satisfy this existence condition?

(b) Given that y satisfies the existence condition, show that the original problem can be converted into the following equality-constrained least-squares problem with a positive definite Hessian

$$\min_{\theta} (\widetilde{y}_1 - \widetilde{X}_1\theta)^T \Sigma_r^{-1} (\widetilde{y}_1 - \widetilde{X}_1\theta) \quad \text{subject to } \widetilde{X}_2\theta = \widetilde{y}_2$$

and $\widetilde{y}_1 = U_1^T y$, $\widetilde{X}_1 = U_1^T X$.

Exercise 4.67: Least squares with equality constraint

Consider the equality-constrained least-squares problem that appears in the previous exercise

$$\min_{\theta} (y - X\theta)^T H (y - X\theta) \quad \text{subject to } A\theta = b$$

with $b \in$ range of A and $H > 0$.

(a) Define the SVD of A as $A = USV$ and change the coordinate system so that $\theta = V\alpha$. Show that the constraint $A\theta = b$ can be solved in the transformed coordinates $\alpha = (\alpha_1, \alpha_2)$ to obtain

$$\alpha_1 = \Sigma_r^{-1} U_1^T b \qquad \alpha_2 \text{ arbitrary}$$

with the solvability condition $b \in$ the range of U_1. Show that this condition is equivalent to the original solvability condition $b \in$ the range of A.

(b) With α_1 determined by the constraint, show that the remaining unconstrained optimization problem is

$$\min_{\alpha_2} (\widetilde{y} - M_2\alpha_2)^T H (\widetilde{y} - M_2\alpha_2)$$

with $M_1 = XV_1$, $M_2 = XV_2$, and $\widetilde{y} = y - M_1\Sigma_r^{-1}U_1^T b$. Since $H > 0$, this problem can be solved as in Exercise 4.65.

Exercise 4.68: The whole enchilada—Generalized maximum likelihood estimation for the linear model

With the preliminaries of the last three exercises, we are ready to solve the problem of interest. Consider again the linear model with normally distributed measurement error

$$y = X\theta + e \qquad e \sim N(0, R)$$

with $y, e \in \mathbb{R}^p$, $\theta \in \mathbb{R}^{n_p}$, and $X \in \mathbb{R}^{p \times n_p}$. We are interested in the case where X may not have independent columns, and $R \geq 0$ may not be positive definite.

(a) Solve this problem by (i) first applying the result of Exercise 4.66 to handle the semidefinite variance matrix; (ii) next applying the result of Exercise 4.67 to handle the imposed equality constraint; and (iii) next applying the result of Exercise 4.65 to solve the optimization problem.

(b) What solvability condition do the data y have to satisfy for the problem to have a solution?

(c) Is the solution unique? If not, what is the minimum norm solution?

(d) Finally, is the resulting estimator *linear*? Explain why or why not.

5

Stochastic Models and Processes

Exercise 5.27: Optimizing a constrained, quadratic matrix function

We are familiar with the vector version of the problem

$$\min_x (1/2) x^T Q x \qquad \text{s.t. } Ax = b$$

with $x \in \mathbb{R}^n$, and $Q > 0$, and $A \in \mathbb{R}^{p \times n}$ having linearly independent rows. The solution is

$$x^0 = Q^{-1} A^T (A Q^{-1} A^T)^{-1} b \tag{5.87}$$

which is readily derived using the method of Lagrange multipliers. The solution exists for all $b \in \mathbb{R}^p$.

We would like to extend this result to the matrix version of the problem

$$\min_X (1/2) \operatorname{tr}(X^T Q X) \qquad \text{s.t. } AX = B$$

with $X \in \mathbb{R}^{n \times n}$, $Q > 0$, and $A \in \mathbb{R}^{p \times n}$ having linearly independent rows. See also Humpherys, Redd, and West (2012).

To solve this constrained problem we consider a matrix of Lagrange multipliers, $\Lambda \in \mathbb{R}^{p \times n}$, and express the equality constraints as the vector equation $\operatorname{vec}(AX - B) = 0$. Next we augment the objective function in the usual way to form the Lagrangian

$$L(X, \Lambda) = (1/2) \operatorname{tr}(X^T Q X) - (\operatorname{vec}\Lambda)^T \operatorname{vec}(AX - B)$$

The constrained problem is then equivalent to the following unconstrained minmax problem

$$\min_X \max_\Lambda L(X, \Lambda)$$

(see Section 1.5.2 of Chapter 1 for a brief review of minmax games.) The necessary and sufficient conditions for a solution to the minmax problem are the matrix equations

$$\frac{dL}{dX} = 0 \qquad \frac{dL}{d\Lambda} = 0$$

(a) Show that for any matrices A and B with identical dimensions

$$(\operatorname{vec}A)^T (\operatorname{vec}B) = \operatorname{tr}(A^T B)$$

Use this result to convert the Lagrangian to

$$L(X, \Lambda) = (1/2) \operatorname{tr}\Big(X^T Q X - 2\Lambda^T (AX - B)\Big)$$

(b) Compute the required derivatives (Table A.3 may be useful for this purpose) and show that the optimal solution satisfies the linear equations

$$\begin{bmatrix} Q & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} X^0 \\ \Lambda^0 \end{bmatrix} = \begin{bmatrix} 0 \\ B \end{bmatrix}$$

(c) Solve these equations to obtain

$$X^0 = Q^{-1}A^T(AQ^{-1}A^T)^{-1}B \tag{5.88}$$

(and $\Lambda^0 = (AQ^{-1}A^T)^{-1}B$). Compare (5.88) to (5.87).

Exercise 5.28: Optimal linear controller

Consider the linear discrete time system

$$x^+ = Ax + Bu$$

with infinite horizon control objective function

$$V(x, \mathbf{u}) = \sum_{k=0}^{\infty} L(x(k), u(k))$$

with input sequence defined as $\mathbf{u} = \{u(0), u(1), \ldots\}$, and stage cost given as

$$L(x, u) = (1/2)\left(x^T Q x + u^T R u\right)$$

From dynamic programming we have already found that the optimal controller is a linear feedback control $u^0(x) = Kx$, and the optimal cost is $V^0(x) = (1/2)x^T \Pi x$. The matrix Π was shown to satisfy the following discrete algebraic Riccati equation with the optimal gain K given by

$$\Pi = Q + A^T \Pi A - A^T \Pi B (B^T \Pi B + R)^{-1} B^T \Pi A \tag{5.89}$$

$$K = -(B^T \Pi B + R)^{-1} B^T \Pi A \tag{5.90}$$

Now we would like to find a shortcut method to derive these two results. Postulate that the optimal control is a linear control law $u = Kx$, and substitute that control law into the model, assume $A + BK$ is stable, compute $V(x)$ for this control law, and verify that $V(x) = (1/2)x^T \Pi x$ with Π satisfying

$$\Pi = Q_K + A_K^T Q_K A_K + (A_K^T)^2 Q_K A_K^2 + \cdots$$

with $Q_K = Q + K^T R K$ and $A_K = A + BK$. Multiply both sides from the left by A_K^T and right by A_k and subtract from Π to show that Π satisfies

$$\Pi - (A + BK)^T \Pi (A + BK) = Q + K^T R K \tag{5.91}$$

Let $\Pi(K)$ denote the solution to (5.91), which is a Lyapunov equation for Π given a fixed value of K. For a controller gain K_0 to be the unique optimal gain, we require that for all $x \in \mathbb{R}^n$ and all $K \neq K_0$

$$x^T \Pi(K_0) x < x^T \Pi(K) x$$

or $\Pi(K_0) < \Pi(K)$. Note that matrix inequality $A < B$ denotes that $A - B$ is a negative definite matrix.

(a) Assume K_0 is the unique optimal gain and denote the solution to (5.91) as $\Pi_0 = \Pi(K_0)$. Let P_K be an arbitrary perturbation and consider $K = K_0 + P_K$. Substitute this K into (5.91), and denoting the solution as $\Pi(K) = \Pi(K_0) + P_\Pi$, derive an equation for the perturbation P_Π. It should be a Lyapunov equation.

(b) Using the result of Exercise 2.63, find a condition for K_0 such that $P_\Pi > 0$ for all $P_K \neq 0$. Show that this condition is (5.90). Thus K_0 is optimal if it satisfies (5.90). Since the optimal solution is assumed unique, this K_0 is the only optimal solution.

(c) Substitute the optimal gain into (5.91) and simplify to show that the optimal Π is given by the solution to (5.89).

Dynamic programming goes further and establishes that *linear* feedback is indeed optimal over all feedback policies, including nonlinear policies, and that the optimal gain is unique.

Exercise 5.29: Optimal linear state estimator

Given the linear stochastic system

$$x^+ = Ax + w \qquad y = Cx + v \qquad w \sim N(0, Q) \quad v \sim N(0, R)$$

we showed that the optimal steady-state estimator is given by

$$\hat{x}^+ = A\hat{x} + L(y - C\hat{x}) \tag{5.92}$$

in which the steady-state filter variance and gain are given by

$$P = Q + APA^T - APC^T(CPC^T + R)^{-1}CPA^T \tag{5.93}$$

$$L = APC^T(CPC^T + R)^{-1} \tag{5.94}$$

Note that these are the optimal state estimate and variance *before* measurement, i.e., $\hat{x} = \hat{x}^-$, $P = P^-$.

As in the previous exercise with regulation, we would like to derive the optimal estimator formulas with a shortcut method. First consider a linear estimator of the form given in (5.92), but with an arbitrary L.

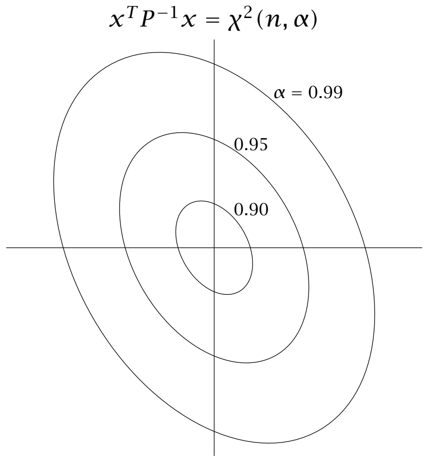

Figure 5.20: The α-level confidence intervals for a normal distribution. Maximizing the value of $x^T P^{-1} x$ over L gives the *smallest* confidence intervals for the estimate error.

(a) Show that for this estimator, the estimate error $\tilde{x} = x - \hat{x}$ satisfies

$$\tilde{x}^+ = (A - LC)\tilde{x} + w - Lv$$

Assuming that $A - LC$ is stable, show that, at steady state, the estimate error is zero mean with variance that satisfies

$$P = (A - LC)P(A - LC)^T + Q + LRL^T \tag{5.95}$$

Therefore, at steady state $\tilde{x} \sim N(0, P)$.

(b) As derived in Chapter 4, the α-level confidence interval for a normally distributed random variable is the ellipsoid

$$x^T P^{-1} x \le F_{\chi^2}^{-1}(\alpha; n)$$

in which $F_{\chi^2}^{-1}(\alpha; n)$ is the inverse cumulative distribution function for the chi-squared distribution. We would like to derive the optimal L so that the confidence interval for estimate error is as small as possible. As shown in Figure 5.20, this means we would like to solve the optimization

$$\max_L x^T P^{-1} x$$

and we would like this L to be optimal for all $x \in \mathbb{R}^n$. Show that, as in the previous exercise, the necessary condition for an optimum is

$$\frac{dP^{-1}}{dL_{ij}} = 0, \qquad i, j = 1, 2, \ldots, n$$

(c) Show that for a nonsingular matrix D, $dD^{-1}(x)/dx = 0$ if and only if $dD(x)/dx = 0$. Therefore the necessary condition can be stated equivalently as

$$\frac{dP_{ab}}{dL_{ij}} = 0, \qquad a, b, i, = 1, 2, \ldots, n, \quad j = 1, \ldots, p \tag{5.96}$$

(d) Differentiate the trace of (5.95) with respect to L, use the necessary condition (5.96), and show that the optimal L agrees with (5.94).

(e) Substitute the optimal gain into (5.95) and simplify to show that the optimal P is given by (5.93).

Deriving the statistically optimal filter in Section 5.4 established that: (i) the necessary condition for optimality (5.96) derived here is also sufficient, and (ii) the *linear* estimator is the optimal estimator.

Exercise 5.30: Diffusion and random walk with n species

One of your experimental colleagues has measured the multicomponent diffusivity *matrix* D corresponding to the flux model

$$N_i = -D_{ij}\nabla c_j \qquad i, j = 1, 2, \ldots, n$$

i.e., a gradient in *any* concentration results in a nonzero flux of *all* species. The matrix D is positive semi-definite (symmetric, with nonnegative eigenvalues).

You are considering the following random process to model this multicomponent diffusion in a single spatial dimension

$$x(k+1) = x(k) + Gw(k)$$

with $x \in \mathbb{R}^n$, $k \in \mathbb{I}_{\geq 0}$, and matrix $G \in \mathbb{R}^{n\times n}$, and $w(k) \sim N(0, I_n)$ distributed as a vector of independent, zero-mean, unit-variance normals.

(a) Solve this difference equation for $x(k)$, $k = 0, 1, \ldots$, as a function of the random term $w(k)$. What is the mean and variance of $x(k)$?

(b) If we let $t = k\Delta t$ and $GG^T = 2\Delta t D$, what is the mean and variance of the resulting continuous time process $x(t)$. Write down the probability density $p(x,t)$ for $x(t)$.

(c) If the simulation for $x(t)$ is considered to be a sample of some random process, provide the evolution equation of its probability density $p(x,t)$? In other words, what is the corresponding diffusion equation for this n-species system. You may assume that the diffusivity matrix D does not depend on x.

(d) How do you calculate a matrix G so that it corresponds to the measured diffusivity matrix D. Hint: think about the SVD (or, equivalently in this case, the eigenvalue decomposition) of D.

(e) We now want to make *sure* that we have the correct random process for the stated diffusion equation. To verify the diffusion equation, we need to take one t derivative and two x derivatives of $p(x,t)$. First establish the following helpful fact

$$D\colon \nabla\nabla \exp(a\, x^T D^{-1} x) = 2a \exp(a\, x^T D^{-1} x)(n + 2a\, x^T D^{-1} x)$$

or in component form

$$D_{ij} \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} \exp(a\, x_k D_{kl}^{-1} x_l) = 2a \exp(a\, x_k D_{kl}^{-1} x_l)(n + 2a\, x_k D_{kl}^{-1} x_l)$$

for a an arbitrary constant scalar.

Even if this part proves elusive, you can still do the next part.

(f) Using the result above, show that the $p(x,t)$ from your random walk satisfies your stated diffusion equation.

Bibliography

J. Humpherys, P. Redd, and J. West. A fresh look at the Kalman filter. *SIAM Rev.*, 54(4):801–823, 2012.