2.13 Exercises

Problem 1 *Using an eigenvector/eigenvalue analysis, solve (a)*

$$\left\{\begin{array}{ccc} 1 & 1 & -2 \\ 1 & 2 & -1 \\ -2 & -1 & 6 \end{array}\right\} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \tag{2.463}$$

and (b)

$$\left\{\begin{array}{ccc} 1 & 1 & -2 \\ 1 & 2 & -1 \\ 1.5 & 2 & -2.5 \end{array}\right\} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \tag{2.464}$$

Problem 2 *(a) Find the ranges and null spaces of*

$$\mathbf{A} = \left\{\begin{array}{ccc} 2 & -1 & 1 \\ 3 & 2 & 1 \end{array}\right\} \tag{2.465}$$

and calculate the solution and data resolution matrices. (b) Let there be a set of observations $\mathbf{y}$, *such that*

$$\mathbf{Ax} + \mathbf{n} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \tag{2.466}$$

This problem is clearly formally undetermined. Find the solution which minimizes

$$J = \mathbf{x}^T\mathbf{x} \tag{2.467}$$

and compare it to the SVD solution with null space set to zero. What is the uncertainty of this solution? (c) Now consider instead

$$\mathbf{A} = \left\{\begin{array}{cc} 2 & 3 \\ -1 & 2 \\ 1 & 1 \end{array}\right\}, \tag{2.468}$$

and the formally overdetermined problem

$$\mathbf{Ax}+\mathbf{n}=\begin{bmatrix}1\\2\\-1\end{bmatrix} \tag{2.469}$$

and find the least-squares solution which minimizes $\mathbf{n}^T\mathbf{n}$. *What is the uncertainty of this solution? How does the solution compare to the SVD solution? (d) For an arbitrary* $\mathbf{A}$, *solve the least-squares problem of minimizng*

$$J=\mathbf{x}^T\mathbf{x}+\boldsymbol{\alpha}^{-2}\mathbf{n}^T\mathbf{n} \tag{2.470}$$

and re-write the solution in terms of its SVD. Discuss what happens to the small singular value contributions.

Problem 3 *There is one observation*

$$x+n_1=1 \tag{2.471}$$

and a priori statistics $<n>=<x>=0, <n^2>=1/2, <x^2>=1/2$. *(a)What is the best estimate of* x, n? *(b) A second measurement becomes available,*

$$x+n_2=3 \tag{2.472}$$

with $<n_2>=0, <n_2^2>=4$. *What is the new best estimate of* x *and what is its estimated uncertainty. Are the various a priori statistics consistent with the final result?*

Problem 4 *Two observations of unknown* x *produce the apparent results*

$$x = 1 \tag{2.473}$$

$$x = 3 \tag{2.474}$$

Produce a reasonable value for x *under the assumption that (a) both observations are equally reliable, and (b) that the second observation is much more reliable (but not infinitely so) than the first (make some reasonable numerical assumption about what "reliable" means and state what you are doing). Can you re-write eqs. (2.473,2.474) in a more sensible form?*

Problem 5 *For the Neumann problem in the last example, let the right-hand boundary flux condition be unknown, but from the forward solution computed in the example, determine as best you can the values of the missing boundary fluxes, from knowledge of* $\mathbf{x}$ *on the interior grid points.*

Problem 6 *Two observations of 3 unknowns, x, y, z produce the apparent result,*

$$x - y - z = 1 \tag{2.475}$$

$$x - y - z = 3 \tag{2.476}$$

Discuss what if, anything, might be inferred from such a peculiar result. You can make some sensible assumptions about what is going on, but say what they are.

Problem 7 *The temperature along an oceanic transect is believed to satisfy a linear rule, $\theta = ar + b$, where r is the distance from a reference point, and a, b are constants. Measurements of θ at sea, called y, produce the following values, $r = 0, y = 10; r = 1, y = 9.5; r = 2, y = 11.1, r = 3, y = 12$. (a) Using ordinary least-squares, find an estimate of a, b and the noise in each measurement, and their standard errors. (b) Solve it again using the SVD and discuss, via the resolution matrices, which of the observations, if any proved most important. Is the solution fully resolved?*

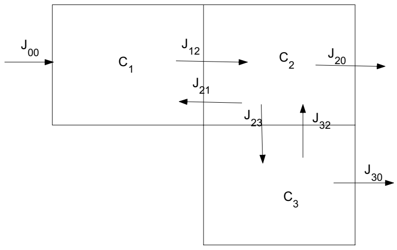
Problem 8 *Consider the system of equations*

$$\left\{\begin{matrix} 1 & 2 & 1 \\ \\ 1 & 2.1 & 1 \end{matrix}\right\} \mathbf{x} + \mathbf{n} = \begin{bmatrix} 1 \\ \\ 2 \end{bmatrix}. \tag{2.477}$$

Using the SVD, compare the solutions at ranks $1, 2$ for the two cases of

$$\mathbf{R} = \mathbf{I}_2, \mathbf{R} = \left\{\begin{matrix} 1 & 0.99999 \\ \\ 0.99999 & 1 \end{matrix}\right\}. \tag{2.478}$$

How do the rank 1 solutions differ in their treatment of the noise? What is the difference in the solutions at rank 2?



{boxprob.eps}

Figure 2.15: Three box model describing tracer movement as depicted.

Figure 2.15 depicts a simple "box-model". There are concentrations C_i in each of three boxes and the mass flux from box i to box j is $J_{ij} > 0$. Box "0" corresponds to externally imposed conditions. (a) Write the simultaneous equations for mass conservation in each box. (b) Let the concentration source or sink in box i be denoted q_i. Write the simultaneous equations for concentration steady-state in each box. (c) Initially all J_{ij} are thought to be about 8 (this is a not very sophisticated way of dealing with the positivity constraint on J_{ij}) and measurements show $C_0 = 5, C_1 = 3, C_3 = 1, q_1 = 20 \pm 2, q_2 = -2 \pm 2, q_3 = 8 \pm 10$. Assuming the measurements of C_i are perfect, make a better estimate of J_{ij},by finding the various corrections ΔJ_{ij}. (d) Assuming $< \Delta J_{ij} >= 0, < \Delta J_{ij}^2 >= 10$, find a solution using the truncated and tapered SVD and the Gauss-Markov Theorem. Find the uncertainty of the estimates. (e) Solve the problem by linear programming without using the *a priori* variances, but enforicing the positivity constraints on the J_{ij}.

Problem 9 *For the Laplace-Poisson equation $\nabla^2\phi = \rho$ with Dirichlet boundary conditions in a square domain, put it into discrete form and code it on a computer so that it can be written,*

$$\mathbf{Ax} = \mathbf{b}. \tag{2.479}$$ {probl3}

Choose any reasonable dimension for the number of grid points or finite elements or basis functions. Confirm that $\mathbf{A}$ is square. (a) For any reasonable boundary conditions ϕ_b and values of ρ, solve (2.479) as a forward problem (b) Add some random noise to ϕ_b and solve it again. (c) Omit any knowledge of ρ over some part of the domain and find at least one possible solution (you could use least-squares). (d) Omit any knowledge of ϕ_b over some part of the domain and find at least one possible solution. (e) Suppose ϕ from (a) is known over part of the domain, use that knowledge to help improve the solutions in (b-d).

Problem 10 *At rank 2, the SVD solution is $\tilde{\mathbf{x}} = [0.27, -1.3, -0.55, -1.55]^T$ which differs from the true solution by the nullspace vectors. How does one interpret this solution?*

Problem 11 *Describe and discuss the above solution when $k^2 < 0$.*

Problem 12 *By the same methods used in this last example, study the behavior of the solution to the modified Bessel equation*

$$r^2\frac{d^2x}{dr^2} + r\frac{dx}{dr} - r^2x = 0, a \le r \le b.$$

Problem 13 *Consider the simultaneous equations,* $\mathbf{Ax} = \mathbf{y}$,

$$\left\{\begin{matrix} 1 & 1 & -1 \\ 2 & 1 & 1 \\ -1 & 0 & -2 \end{matrix}\right\} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

(a) Using a numerical routine for symmetric matrix eigenvalue/eigenvector problems (i.e., do not *use a singular value decomposition program such as MATLAB's SVD), find the singular value decomposition for the matrix* $\mathbf{A}$. *(b) Find the null space and ranges of* $\mathbf{A}, \mathbf{A}^T$ *(c) Using the singular vectors and singular values, find the general solution to the equations and explain the behavior of this solution. Are there any residuals? (d) Find the resolution matrix for the solution and for the "data",* $\mathbf{y}$.

Problem 14 *You have five data points,* $y_t = 1, -2, -3, -2, -1$ *,*$t = 0, 1, 3, 4, 5$ *and you have reason to believe they are given by a reduced Fourier Series*

$$y_t = acos(2\pi t/6) + bsin\,(2\pi t/3) + n_t \tag{2.480}$$

where n_t *is noise. Solve this problem for estimates of* $a_,\ b, n_t$ *in three ways (a) As an ordinary least-squares problem. (You can use a matrix inversion routine if you wish.). (b) As an under-determined problem in 7 unknowns. (c) By the singular value decomposition (you may use an svd routine if you want). Explain the differences among the solutions. (d) The noise variance is believed to be* $< n_t^2 >= 1.5$. *Make an estimate of the uncertainty in your estimates of* a, b.

Problem 15 *Extend the discussion of determining a mean in a correlated time series (P. 135) to the problem of finding a trend, and calculate the dependence of the slope of the trend on the correlation.*

Problem 16 *(a) Set up the Neumann problem as in Problem 9 and show explicitly that there is a solution and "observation" null space. Interpret them. (b) Let the normal boundary condition be* $\partial\phi/\partial n = 3$ *everywhere. Is there any difficulty? What is its character, and how might it be dealt with?*

Problem 17 *For the Neumann problem, write the model equations with error terms, and solve the problem with additional information providing estimates of* ϕ_{ij} *at several grid points (rendering the problem formally overdetermined).*

Notes

[9]Noble & Daniel (1977); Strang (1988).

[10]Lawson & Hanson (1974)

[11]"Positive definite" will be defined below. Here it suffices to mean that $\mathbf{c}^T\mathbf{W}\mathbf{c}$ should never be negative, for any $\mathbf{c}$.

[12]Golub & Van Loan (1989)

[13]Haykin (1986, p. 61).

[14]Press *et al.* (1992); Lawson & Hanson (1974); Golub & van Loan (1989); etc.

[15]Determinants are used only rarely in this book. Their definition and properties are left to the references, as they are usually encountered in high school mathematics.

[16]Rogers (1980) is an entire volume of matrix derivative identities, and many other useful properties are discussed by Magnus and Neudecker (1988).

[17]Magnus and Neudecker (1988), P. 183

[18]Liebelt (1967, Section 1–19)

[19]The history of this not-very-obvious identity is discussed by Haykin (1986, p. 385).

[20]A good statistics text such as Cramér (1946), or one on regression such as Seber (1977), should be consulted.

[21]Feller (1957) and Jeffreys (1961) represent differing philosophies. Jaynes (2003) forcefully and colorfully argues the case for so-called Bayesian inference (following Jeffreys), and it seems likely that this approach to statistical inference will ultimately become the default method; Gauch (2003) has a particularly clear account of Bayesian methods. For most of the methods in this book, however, we use little more than the first moments of probability distributions, and hence can ignore the underlying philosophical debate.

[22]It follows from the Cauchy-Schwarz inequality: Consider $\left\langle (ax'+y')^2\right\rangle = a\left\langle x'^2\right\rangle + \left\langle y'^2\right\rangle + 2a\left\langle x'y'\right\rangle \geq 0$ for any constant a. Choose $a = -\left\langle x'y'\right\rangle / \left\langle x'^2\right\rangle$, and one has $-\left\langle x'y'\right\rangle^2 / \left\langle x'^2\right\rangle + \left\langle y'^2\right\rangle \geq 0$, or $1 \geq \left\langle x'y'\right\rangle^2 / \left(\left\langle x'^2\right\rangle \left\langle y'^2\right\rangle\right)$. Taking the squareroot of both sides, the required result follows.

[23]Draper & Smith (1998); Seber and Lee (2003).

[24]Numerical schemes for finding $\mathbf{C}_{\xi\xi}^{1/2}$ are described by Lawson and Hanson (1976) and Golub and Van Loan (1989)

[25]Cramér (1946) discusses what happens when the determinant of $\mathbf{C}_{\xi\xi}$ vanishes, that is, if $\mathbf{C}_{\xi\xi}$ is singular.

[26]Bracewell (1978).

[27]Cramér (1946).

[28]In problems involving time, one needs to be clear that "stationary" is not the same idea as "steady."

[29]If the means and variances are independent of i, j and the first cross-moment is dependent only upon $|i-j|$, the process x is said to be stationary in the "wide-sense." If all higher moments also depend only on $|i-j|$, the process is said to be stationary in the "strict-sense," or more simply, just stationary. A Gaussian process has the unusual property that wide-sense stationarity implies strict-sense stationarity.

[30]The terminology "least-squares" is reserved in this book, conventionally, for the minimization of discrete sums such as Eq. (2.90). This usage contrasts with that of Bennett (2002) who applies it to continuous integrals, such as, $\int_a^b (u(q) - r(q))^2\,dq$ leading to the calculus of variations and Euler-Lagrange equations.

[31]Seber (1977) or Box et al. (1994) or Draper and Smith (1981) are all good starting points.

[32]Draper and Smith (1981, Chapter 3) and the references given there.

[33]Gill, Murray and Wright (1981).

[34]Wunsch & Minster (1982).

[35]Morse & Feshbach (1953, p. 238); Strang (1988) .

[36]See Sewell (1987) for an interesting discussion.

[37]But the matrix transpose is not what the older literature calls the "adjoint matrix," and which is quite different. In the more recent literature the latter has been termed the "adjugate" matrix to avoid confusion.

[38]In the meteorological terminology of Sasaki (1970) and others, exact relationships are called "strong" constraints, and those imposed in the mean-square are "weak" ones.

[39]Claerbout (2001) displays more examples, and Lanczos (1960) gives a very general discussion of operators and their adjoints, Green functions, and their adjoints. See also the Appendix to this Chapter.

[40]Wiggins (1972).

[41]Brogan (1985) has a succinct discussion.

[42]Lanczos (1961), pages 117-118, sorts out the sign dependencies.

[43]Lawson and Hanson (1974).

[44]The singular value decomposition for arbitrary non-square matrices is apparently due to the physicist-turned-oceanographer Carl Eckart (Eckart & Young, 1939; see the discussion in Haykin, 1986; Klema & Laub, 1980; or Stewart, 1993). A particularly lucid account is given by Lanczos (1961) who however, fails to give the decomposition a name. Other references are Noble and Daniel (1977), Strang (1986) and many recent books on applied linear algebra. The crucial role it plays in inverse methods appears to have been first noticed by Wiggins (1972).

[45]Munk et al. (1996).

[46]In physical oceanography, the distance would be that steamed by a ship between stops for measurement, and the water depth is clearly determined by the local topography.

[47]Lawson & Hanson (1974), or Hansen (1992). Hansen's (1992) discussion is particularly interesting because he exploits the "generalized SVD," which is used to simultaneously diagonalize two matrices.

[48]Munk and Wunsch (1982).

[49]Seber (1977).

[50]Luenberger (1984).

[51]In oceanographic terms, the exact constraints describe the Stommel Gulf Stream solution. The eastward intensification of the adjoint solution corresponds to the change in sign of β in the adjoint model. See Schröter and Wunsch (1986) for details and an elaboration to a non-linear situation.

[52]Lanczos (1970) has a good discussion.

[53]See Lanczos (1961, Section 3.19)

[54]The derivation follows Liebelt (1967).

[55]Bretherton *et al.* (1976) obtain similar results.

[56]The time series was generated as $y_t = 0.999 y_{t-1} + \theta_t$, $\langle \theta_t \rangle = 0$, $\left\langle \theta_t^2 \right\rangle = 1$, a so-called autoregressive process of order 1 (AR(1)). The covariance $\langle y_i y_j \rangle$ can be determined analytically; see Priestley (1981), p.119. Many geophysical processes obey similar rules.

[57]Brogan (1985); Stengel (1986).

[58]Paige & Saunders (1982)

[59]See especially, van Trees (1968).

[60]Liebelt (1967, P.164)