Goals:

Integrate physical and graphical reasoning. Investigate existence and uniqueness of solutions. Employ graphing calculator to analyze a physical problem.

Mathematical Background:

Quadratic equations, inequalities, parabolas

Scientific Background:

Newton's law of gravity

Technology:

Graphing calculator

2.3.1 Introduction

This module is inspired by the ancient Germanic myth of the Niebelungen, a race of diminutive malefactors whose golden hoard was stashed at the bottom of the Rhine river. We treat the problem of locating and identifying a single isolated gravitational point source. Suppose a point mass m, say a nugget of gold, lies at the bottom of a calm river that is 1 unit deep. We make the simplifying assumption that all other gravitational sources are purely homogeneous so that the gravitational anomaly generated by the point source will be considered to be the only true gravitational effect. The determination of the gravitational force on a unit mass on the surface of the river engendered by the nugget at the bottom of the river is a very simple direct problem. Newton's law of gravitation holds that this force is equal to a known constant (the gravitational constant) times the product of the masses, divided by the square of the distance separating the masses. This is the famous "inverse square" law of gravitational attraction.

In this module we consider the inverse problem of determining the mass and location of a single nugget from measurements taken at the surface. The measurements consist of a distance x from a reference point on the surface and an estimate μ (obtained, say, by use of a delicate spring scale) of the vertical component of the gravitational force on the unit mass at position x on the surface engendered by the nugget below the surface. The situation is illustrated in Figure 2.5.

The square of the distance between the source nugget and the unit mass on the measuring device is given by the Pythagorean theorem, $1 + (x - s)^2$, and the product of the masses is $1m$, where m is the mass of the nugget. The vertical component, μ, of the gravitational effect at the position x on the surface

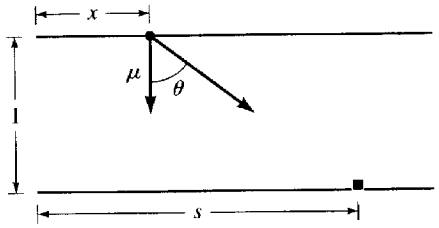

Figure 2.5: Vertical Force Engendered by a Point Source

is therefore, by Newton's law of gravitation,

$$\mu = \gamma \frac{m}{1 + (x - s)^2} \cos\theta,$$

where γ is the gravitational constant and θ is the angle pictured. Since the depth of the river is 1, we see from the figure that

$$\cos\theta = \frac{1}{\sqrt{1 + (x - s)^2}},$$

and substituting this above we get

$$\mu = \gamma m (1 + (x - s)^2)^{-3/2}.$$

The *direct* problem in this context is to determine the vertical force μ on the unit mass at position x on the surface that the mass m at position s at the bottom of the river engenders. This direct problem clearly has a unique solution given by the equation above. We will consider the *inverse* problem of determining the mass m and position s of the source from observations of the force μ at surface sites x. Before getting into this problem we will reformulate the equation to make our problem a bit simpler. Define new variables M and G, which we will call the *effective mass* and *effective vertical force*, respectively, by

$$M = m^{2/3} \quad G = \left(\frac{\mu}{\gamma}\right)^{2/3}$$

A measurement of μ then uniquely determines G and knowledge of M uniquely determines m. With these definitions, the equation above is easily seen to be equivalent to

$$M - G = G(x - s)^2.$$

The inverse problem now is equivalent to determining M, from which the mass m of the nugget can be obtained, and the location s of the nugget from knowledge of x and the effective force G at position x. We will call the pair (x, G) an *observation* because it consists of observing the effective force G (obtainable from μ) at the site x. The inverse problem therefore is equivalent to determining a pair (s, M), which we will call a *source*, from observations (x, G). If a unique source (s, M) is determined, then we have found the location s and the mass $m = M^{3/2}$ of the inaccessible nugget without getting wet—a feat that would surely arouse the envy of the Rhine Maidens!

2.3.2 Activities

1. Question Does a single observation (x, G) uniquely determine the source (s, M)?

2. Question Suppose s is plotted on a horizontal axis and M is plotted on a vertical axis. What is the shape of the *source curve* associated with a given observation (x, G)? (In other words, what is the graph of all possible single-point sources (s, M) that could account for the observation (x, G)?)

3. Question How does the shape and position of the source curve change with changes in the observation (x, G)?

4. Calculation Plot the source curve associated with the observation $(1, 2)$.

5. Problem Suppose an observation (x, G) is given. (a) Show that for every number $M > G$ there are two sources with effective mass M that can account for the observation. (b) Show that there is a unique source of effective mass $M = G$ that can account for a given observation (x, G). What is the location of this source? (c) Show that if $M < G$, then no source of effective mass M can account for the given observation.

6. Exercise Explain Problem 5 in intuitive physical terms rather than in mathematical terms.

7. Question Suppose observations (x_1, G) and (x_2, G) are recorded at distinct sites $x_1 \neq x_2$. What is the location of the source?

8. Exercise Find all sources (s, M) that can account for both of the observations $(0, 1)$ and $(1/\sqrt{2}, 2)$.

9. Question Is $\{(0, 1), (2, 6)\}$ a possible pair of observations? In other words, is there a single point source (s, M) that can engender both of the observations $(0, 1)$ and $(2, 6)$?

10. Exercise Find all sources that can account for the pair of observations $\{(0, 1), (1, 2)\}$.

11. Calculation Plot the source curve for the observation $(1.12, 2.7)$. On the same axes, plot the source curve for the observation $(3.1, 4.89)$. Estimate the sources (position and effective mass) that give rise to these observations.

12. Calculation Estimate the point on the source curve engendered by the observation $(2.1, 4)$ that is closest to the origin.

13. Calculation Estimate all sources that can generate the observations $(-1.1, 2.2)$ and $(0.9, 8.9)$.

14. Calculation Are the observations $(1.2, 3.4)$, $(-2.1, 1.1)$, and $(3.07, 2.7)$ consistent? (In other words, is there a source (s, M) that generates these observations?)

15. Question What is the largest number of possible sources that can account for a set of two or more observations at distinct sites?

16. Problem Find conditions on distinct observations (x_1, G_1) and (x_2, G_2) for which (a) the observations are inconsistent, (b) the observations determine a unique source, or (c) the observations may be accounted for by two distinct sources.

17. Problem Suppose two observations with $G_1 \neq G_2$ *uniquely* determine a source (s, M). Show the following:

(a) The distance between the observation sites is

$$|x_1 - x_2| = \frac{|G_1 - G_2|}{\sqrt{G_1 G_2}}.$$

(b) The source is located at

$$s = \frac{G_1 x_1 - G_2 x_2}{G_1 - G_2}.$$

(c) The effective mass is

$$M = G_1 + G_2.$$

18. Problem Show that at most one source can be located between distinct observation sites (i.e., given observations (x_1, G_1), (x_2, G_2) with $x_1 \neq x_2$, there can be at most one source (s, M), with s between x_1 and x_2, that gives rise to the observations).

sum is 1, while the other diagonal (SW to NE) sum is 3. Use 'art1' and 'displa', with the zero vector as an initial approximation, to get a picture of the worm.

10. Computation A 6×6 object has row sums (top to bottom) 0, 2, 0, 2, 6, 0 and column sums (left to right) 2, 2, 1, 1, 2, 2. Try to reconstruct the object using 'art1' and 'displa' using the zero vector as an initial approximation. Now add four more views and measurements consisting of a ray through pixels 4, 11, and 18 with sum 1; a ray through pixels 24, 29, and 34 with sum 2; a ray through pixels 19, 26, and 33 with sum 2; and a ray through pixels 3, 8, and 13 with sum 1. Try to reconstruct the picture using this additional information, and compare the result with the previous reconstruction.

11. Computation Repeat the previous computations using various (nonzero) initial approximations, and compare the results with the previous reconstructions.

12. Computation Repeat the previous computation, blending 5 percent uniform random noise into the measurements, and compare the results to the previous reconstructions.

5.2.3 Notes and Further Reading

The word "tomography" is based on the Greek root "tomos" meaning a cut or slice. What we have called "views" in this module could be called slices through the object. The shadowy images sometimes seen in the Activities above arise from the underdetermined nature of the linear systems involved. In the tomography community such spurious images are called "ghosts."

The ART algorithm is not new; it dates back to the work of Stefan Kaczmarz in the mid-1930s (Kaczmarz was murdered in a Nazi roundup of intellectuals following the invasion of Poland in 1939). A convergence proof of ART, in a more general context than that of this module, can be found, for example, in C. W. Groetsch, *Inverse Problems in the Mathematical Sciences*, Vieweg, Braunschweig, 1993. The ART algorithm has a number of advantages over direct methods for the tomography problem. Since all components of view vectors are either zero or one, the view vectors may be stored as bit strings, and individual projections may be computed very quickly. Also, new view vectors and corresponding measurements can be easily introduced during the course of the computation if new data becomes available. Furthermore, a priori information can be incorporated simply via the initial approximation vector.

"Image Reconstruction from Projections," by R. Gordon, G. Herman, and S. Johnson, *Scientific American*, Vol. 233 (October 1975), pp. 56–68, is an excellent popular article on computed tomography that gives more details on the medical technology involved. Ivars Peterson's article, "Inside Averages," *Science News* (May 1986), pp. 300–301, discusses an interesting application of tomography to literature.

5.3 Nonpolitical Pull

Course Level:

Linear Algebra

Goal:

Investigate the instability of a model problem in geophysics.

Mathematical Background:

Midpoint rule, matrix inverses, eigenvalues

Scientific Background:

Inverse-square law of gravity

Technology:

Graphics–symbolic calculator, MATLAB

5.3.1 Introduction

Let's renew our acquaintance with the Rhine maidens (see the module *das Rheingold*). But now, instead of a discrete nugget of gold, we wish to identify a nonhomogeneous mass density $w(s)$, $0 \leq s \leq 1$. Such a mass distribution engenders an inhomogeneity $\mu(x)$ in the vertical force of gravity at the surface. The relationship between w and μ is easily obtained as in the earlier module and is illustrated in Figure 5.5: The vertical component of force $\Delta\mu(x)$ at position

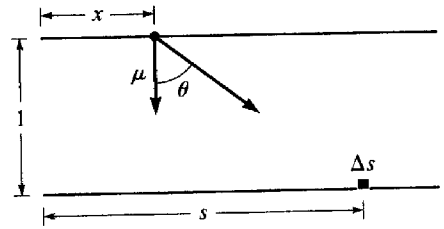

Figure 5.5: Gravitational Attraction of a Distributed Mass

x engendered by a mass segment of length Δs at position s on the subsurface is

$$\begin{aligned}\Delta\mu(x) &= \gamma w(s)((x-s)^2+1)^{-1}\cos\theta\Delta s \\ &= \gamma((x-s)^2+1)^{-3/2}w(s)\Delta s,\end{aligned}$$

where $w(s)$ is the mass density at position s. The usual summation and limit process leads to the model

$$\mu(x) = \int_0^1 ((x-s)^2+1)^{-3/2}w(s)ds$$

(here, and henceforth, we take $\gamma = 1$ for convenience). The problem of determining the gravitational inhomogeneity μ from the mass density w is a straightforward *direct* problem. On the other hand, the problem of determining the inaccessible mass density w from a known (i.e., measured) gravitational inhomogeneity μ is a classic *inverse* problem.

A notable feature of the model above is that μ is generally smoother than w. Even for quite "rough" (e.g., discontinuous) functions w, the function μ is infinitely differentiable because it inherits its smoothness in x from the kernel function $((x-s)^2+1)^{-3/2}$. In "filtering" the function w through the integral we can expect some of the fine detail in w to be "smoothed out" in the process. The essential point is that μ contains less information than w and we can therefore expect that the inverse problem of reconstructing w from knowledge of μ will be difficult.

We can form a concrete appreciation for the difficulties involved by studying an approximating discrete problem. One way of doing this comes about if we replace the integral by an approximate quadrature rule. For example, we might use the *midpoint rule*, that is,

$$\int_0^1 f(s)\,ds \approx h\sum_{j=1}^{n} f(s_j),$$

where $h = 1/n$ and $s_j = (j-\frac{1}{2})h$ for $j = 1,\ldots,n$. Applying this rule to the model above we have

$$\mu(x) \approx h\sum_{j=1}^{n}((x-s_j)^2+1)^{-3/2}w(s_j).$$

If we insist that this relationship hold at each of the midpoints $x = s_i$ for $i = 1,\ldots,n$, we obtain vectors $\mathbf{w}$ and $\boldsymbol{\mu}$, whose components approximate the values of the functions w and μ, respectively, at the midpoints of the subintervals

formed in the discretization process. The vectors $\mathbf{w}$ and $\boldsymbol{\mu}$ are related by the matrix equation

$$A\mathbf{w} = \boldsymbol{\mu},$$

where A is the $n \times n$ matrix with entries

$$a_{ij} = h((s_i - s_j)^2+1)^{-3/2} \quad i,j = 1,\ldots,n.$$

The *discrete* model

$$A\mathbf{w} = \boldsymbol{\mu}$$

may then be taken as an approximation to the *continuous* model

$$\int_0^1 ((x-s)^2+1)^{-3/2}w(s)\,ds = \mu(x).$$

The program 'geo' will produce, for a given positive integer n, the $n \times n$ matrix A that is the discrete model of the geophysical prospecting problem, along with the vector $\mathbf{s}$ of midpoint samples. It is then easy and entertaining to visualize the dramatic smoothing properties of the discrete model. For example, in Figure 5.6 a very rough mass density on 50 midpoints is plotted that consists of random noise in the interval $[0, 2]$. Applying the matrix A, obtained from

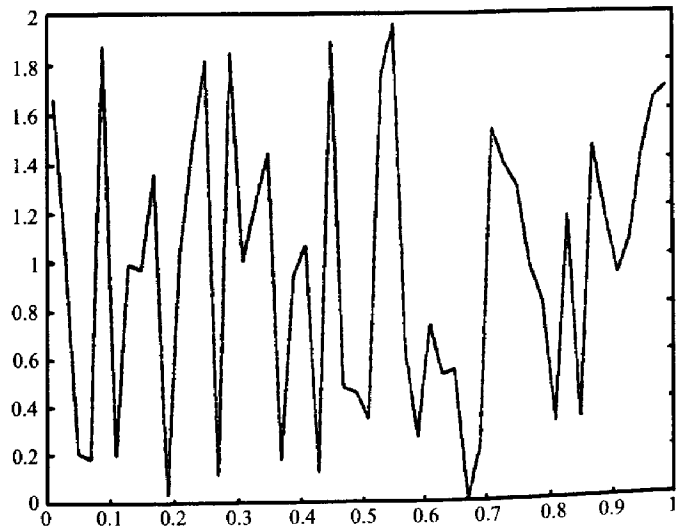

Figure 5.6: A Random Mass Density

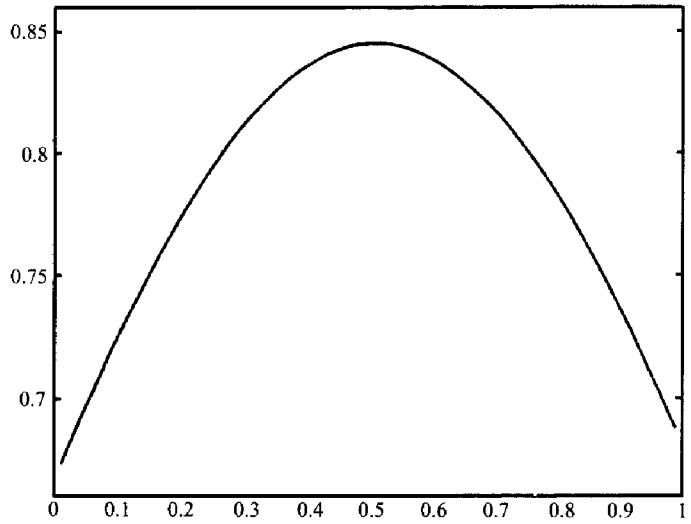

Figure 5.7: Resulting Gravitational Inhomogeneity

'geo' with $n = 50$, to this discrete density results in the very smooth curve plotted in Figure 5.7. The plot in the second figure represents the smooth, information-poor function μ, while that in the first figure illustrates the rough, information rich density w that engenders μ.

Since the direct model appears to reduce fluctuations, i.e., to smooth the input, it should come as no surprise that the inverse process will tend to magnify fluctuations in the data. To put it another way, the inverse problem is *unstable*. The activities below contain a number of computational exercises in which this phenomenon is explored. The instability of the inverse problem for the discrete model can be explained in terms of the *condition number* of the model matrix A. This matrix is symmetric and positive definite and its condition number, relative to the usual euclidean norm $\| \cdot \|$, is the ratio of its largest eigenvalue, say λ_n, to its smallest eigenvalue, say λ_1:

$$cond(A) = \lambda_n/\lambda_1.$$

General treatments of the role of the condition number in perturbation analysis can be found in most numerical analysis texts. We illustrate the possibilities with a particular example. Suppose that $0 < \lambda_1 < \cdots < \lambda_n$ are the eigenvalues of A and that $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are corresponding orthonormal eigenvectors. If $\boldsymbol{\mu} = \lambda_n \mathbf{u}_n$, then

$$A\mathbf{w} = \boldsymbol{\mu},$$

where $\mathbf{w} = \mathbf{u}_n$. If the right-hand side of this equation is perturbed by

$$\Delta\boldsymbol{\mu} = \frac{a}{\lambda_1}\mathbf{u}_1$$

for some positive scalar a, then the solution of the perturbed system is

$$\mathbf{w} + \Delta\mathbf{w} = \mathbf{u}_n + \frac{a}{\lambda_1}\mathbf{u}_1.$$

The relative sizes of the perturbations are then related in the following way:

$$\frac{\| \Delta\mathbf{w} \|}{\| \mathbf{w} \|} = \frac{\lambda_n}{\lambda_1}\frac{\| \frac{a}{\lambda_1}\mathbf{u}_1 \|}{\| \lambda_n\mathbf{u}_n \|}$$

$$= cond(A)\frac{\| \Delta\boldsymbol{\mu} \|}{\| \boldsymbol{\mu} \|}.$$

Therefore, the relative error in the right-hand side may be magnified by a factor of $cond(A)$. Linear systems with coefficient matrices having large condition numbers are called *ill-conditioned*. Unstable inverse problems give rise to discrete models with ill-conditioned matrices. The solution of such ill-conditioned systems is particularly challenging as the data of the problem, which is represented in the right-hand side of the discrete model, invariably contains errors. Even if the error is only that which results from representing the real numbers in a computer in floating-point form, a severely ill-conditioned matrix can have very unpleasant effects, as will be seen in some of the computations below. In general, data errors are greatly magnified in the solution process for ill-conditioned problems, and special measures must be taken to dampen this error magnification. One method for accomplishing this is hinted at in the activities.

5.3.2 Activities

1. Computation Generate the discrete models A for $n = 5$, 10, 20, and 40, and in each case use the MATLAB function 'cond' to find the condition number of A.

2. Computation Use the program 'geo' to produce the 100×100 discrete model matrix A. Generate various random 100-vectors $\mathbf{x}$, plot $\mathbf{x}$ and $A\mathbf{x}$, and note the qualitative features of both.

3. Computation Show that the matrix A of the discrete model has positive eigenvalues for $n = 5$, 10, and 20.

4. Problem Find the gravitational inhomogeneity μ engendered by the constant mass density $w(s) = 1$ (this is the density used in the program 'geo').

provided to allow the student to carry out numerical simulations of the inverse problems with the aim not only of constructing approximate solutions but also of investigating the inherent instability of the problems.

5.1 Cause and Identity

Course Level:

Linear Algebra

Goal:

Interpret some basic problems of linear algebra as inverse problems of causation and model identification.

Mathematical Background:

Matrices and linear equations, vector and matrix norms, Gauss–Jordan elimination method

Scientific Background:

Ohm's Law, Kirchhoff's Law

Technology:

MATLAB or other high-level numerical software

5.1.1 Introduction

Linear algebra is the one course in the undergraduate curriculum in which the issues of existence, uniqueness, and stability raised by inverse problems get serious, though often inadequate, attention. The *direct* problem of linear algebra consists of determining the action of a linear transformation represented (relative to given bases) by a matrix: Given an $m \times n$ matrix A and an n-vector $\mathbf{x}$, determine the m-vector $\mathbf{b} = A\mathbf{x}$. The inverse *causation* problem, that is, the problem of finding all solutions $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$, probably gets more attention than any other problem in elementary linear algebra. A less frequently treated inverse problem is the *identification* problem: Identify the matrix A, given an appropriate collection of "input–output" pairs $(\mathbf{x}, \mathbf{b})$ satisfying $A\mathbf{x} = \mathbf{b}$. This module is an elementary presentation of both of these inverse problems for $m \times n$ real matrices.

First we consider the inverse causation problem. A solution $\mathbf{x}$ of this problem, that is, a vector $\mathbf{x} \in R^n$ satisfying $A\mathbf{x} = \mathbf{b}$, where A is a given $m \times n$ real matrix and $\mathbf{b} \in R^m$ is a given vector, exists if and only if $\mathbf{b}$ lies in the *range* of A, that is, in the subspace

$$R(A) = \{A\mathbf{x} : \mathbf{x} \in R^n\}.$$

This subspace is, according to the definition of the action of the matrix A on a vector $\mathbf{x}$, just the subspace of R^m consisting of all linear combinations of the column vectors of A. Determining whether $\mathbf{b} \in R(A)$, that is, whether a solution exists, and finding all solutions, is accomplished by that excellent algorithm, the method of Gaussian elimination.

The uniqueness issue is addressed by a subspace of R^n associated with A, the *null-space*

$$N(A) = \{\mathbf{x} \in R^n : A\mathbf{x} = \mathbf{0}\}.$$

Again the Gaussian elimination algorithm is an effective means of characterizing the null-space and thereby settling the uniqueness question.

The stability, with respect to perturbations in the right-hand side $\mathbf{b}$ of the solution $\mathbf{x}$ of the problem $A\mathbf{x} = \mathbf{b}$, can be quantified in terms of the *condition number* of the matrix A. We assume that a unique solution exists for each $\mathbf{b}$, that is, that A is an invertible matrix. We would like to know to what extent relatively small errors in $\mathbf{b}$ can lead to relatively large changes in the solution $\mathbf{x}$. Suppose $\tilde{\mathbf{b}}$ is a perturbation of the right-hand side $\mathbf{b}$. The size of this perturbation relative to the size of $\mathbf{b}$, measured in terms of a given norm $\|\cdot\|$, is then $\|\mathbf{b} - \tilde{\mathbf{b}}\|/\|\mathbf{b}\|$. Let x be the unique solution of the system corresponding to the right-hand side $\mathbf{b}$, and let $\tilde{\mathbf{x}}$ be that corresponding to the right-hand side $\tilde{\mathbf{b}}$. Then

$$\|\mathbf{x} - \tilde{\mathbf{x}}\| = \|A^{-1}\mathbf{b} - A^{-1}\tilde{\mathbf{b}}\| \leq \|A^{-1}\|\|\mathbf{b} - \tilde{\mathbf{b}}\|;$$

hence the matrix norm $\|A^{-1}\|$ gives a bound for the change in the solution arising from a perturbation in the right-hand side. A relative measure of this change is obtained as follows:

$$\|\mathbf{x} - \tilde{\mathbf{x}}\| \leq \|A^{-1}\|\|\mathbf{b}\|\frac{\|\mathbf{b} - \tilde{\mathbf{b}}\|}{\|\mathbf{b}\|} \leq \|A^{-1}\|\|A\|\|\mathbf{x}\|\frac{\|\mathbf{b} - \tilde{\mathbf{b}}\|}{\|\mathbf{b}\|},$$

and hence

$$\frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} \leq cond(A)\frac{\|\mathbf{b} - \tilde{\mathbf{b}}\|}{\|\mathbf{b}\|},$$

where $cond(A) = \|A\|\|A^{-1}\|$ is called the *condition number* of the matrix A (with repect to the norm $\|\cdot\|$). The condition number therefore gives an upper bound for the relative error in the solution caused by a given relative error in the right-hand side. For matrices with large condition numbers, that is, *ill-conditioned* matrices, relatively small perturbations in the right-hand side can give rise to relatively large changes in the solution. It is in this sense that ill-conditioned systems are said to be unstable.

We now consider what can be accomplished with linear systems that have no solution, or too many solutions. In the case when $\mathbf{b} \in R^m$ is not in the range of the $m \times n$ matrix A, there is no solution to the problem $A\mathbf{x} = \mathbf{b}$, but all is not lost. A remarkable relationship between the null-space, range, and transpose allows the development of a type of generalized solution, and such generalized solutions *always* exist. The sort of generalized solution we have in mind is a *least-squares* solution, that is, a vector $\mathbf{u} \in R^n$ that minimizes the quantity $\|A\mathbf{x} - \mathbf{b}\|$ over all $\mathbf{x} \in R^n$, where the norm is the usual euclidean norm. Note that this minimum is zero if and only if the system has a solution. If $\mathbf{u}$ is a least-squares solution, then for any vector $\mathbf{v} \in R^n$, the function

$$g(t) = \|A(\mathbf{u} + t\mathbf{v}) - \mathbf{b}\|^2 = \|A\mathbf{u} - \mathbf{b}\|^2 + 2(A\mathbf{v}, A\mathbf{u} - \mathbf{b})t + \|A\mathbf{v}\|^2 t^2,$$

where $(\cdot, \cdot)$ is the familiar euclidean inner product, has a minimum at $t = 0$. The necessary condition $g'(0) = 0$ for a minimum then gives

$$(A\mathbf{v}, A\mathbf{u} - \mathbf{b}) = 0,$$

and hence $(\mathbf{v}, A^TA\mathbf{u} - A^T\mathbf{b}) = 0$ for all $\mathbf{v} \in R^n$. That is, if $\mathbf{u}$ is a least-squares solution, then

$$A^TA\mathbf{u} = A^T\mathbf{b},$$

where A^T is the transpose of A.

Conversely, if $A^TA\mathbf{u} = A^T\mathbf{b}$, then for any $\mathbf{x} \in R^n$,

$$\begin{aligned}\|A\mathbf{x} - \mathbf{b}\|^2 &= \|A(\mathbf{x} - \mathbf{u}) + A\mathbf{u} - \mathbf{b}\|^2 \\ &= \|A(\mathbf{x} - \mathbf{u})\|^2 + 2(A(\mathbf{x} - \mathbf{u}), A\mathbf{u} - \mathbf{b}) + \|A\mathbf{u} - \mathbf{b}\|^2 \\ &= \|A(\mathbf{x} - \mathbf{u})\|^2 + 2(\mathbf{x} - \mathbf{u}, A^TA\mathbf{u} - A^T\mathbf{b}) + \|A\mathbf{u} - \mathbf{b}\|^2 \\ &\geq \|A\mathbf{u} - \mathbf{b}\|^2,\end{aligned}$$

that is, $\mathbf{u}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$.

So, least-squares solutions of $A\mathbf{x} = \mathbf{b}$ coincide with ordinary solutions of the symmetric problem $A^TA\mathbf{x} = A^T\mathbf{b}$. Now this symmetric problem *always* has

a solution since $R(A^T) = R(A^TA)$ (see Problem 9) and hence $A^T\mathbf{b} \in R(A^TA)$ for *any* $\mathbf{b} \in R^m$. Therefore, any linear system $A\mathbf{x} = \mathbf{b}$ has a least-squares solution. However, least-squares solutions need not be unique. Indeed, if $\mathbf{u}$ is a least-squares solution, then so is $\mathbf{u} + \mathbf{v}$ for any $\mathbf{v} \in N(A)$, that is, the set of least-squares solutions forms a hyperplane parallel to the null-space. Therefore, if A has a nontrivial null-space, then $A\mathbf{x} = \mathbf{b}$ has infinitely many least-squares solutions. However, one least-squares solution can be distinguished from the others, namely, the one that is orthogonal to the null-space. There can be at most one such least-squares solution, because if $\mathbf{u}$ and $\mathbf{w}$ are both least-squares solutions that are orthogonal to $N(A)$, then $\mathbf{u} - \mathbf{w}$ is orthogonal to $N(A)$. Also, $A^TA(\mathbf{u} - \mathbf{w}) = A^T\mathbf{b} - A^T\mathbf{b} = \mathbf{0}$, and hence $\mathbf{u} - \mathbf{w} \in N(A^TA) = N(A)$ (see Problem 6). Therefore, $\mathbf{u} - \mathbf{w} \in N(A) \cap N(A)^\perp$, that is, $\mathbf{u} = \mathbf{w}$. On the other hand, there is always a least-squares solution that is orthogonal to the null-space (see Problem 10), and hence *any* linear system has a unique least-squares solution that is orthogonal to the null-space of the coefficient matrix. If we agree to accept this notion of generalized solution, then *every* linear system has a unique (generalized) solution.

Finally, we briefly consider the identification problem, that is, the inverse problem of determining an $m \times n$ matrix A, given pairs of vectors (x, b) related by $A\mathbf{x} = \mathbf{b}$. For each such pair we call $\mathbf{x}$ the input and $\mathbf{b}$ the corresponding output. Our job is to identify the "black box" A, by "interrogating" it with appropriate inputs $\mathbf{x}$ and observing the outputs $\mathbf{b}$. Because we control the inputs, we can arrange it so that they are linearly independent, and we will assume that this has been done. It is convenient to express things in matrix form by aggregating the independent inputs $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_p$ as the column vectors of an $n \times p$ matrix $\mathbf{x}$, and similarly thinking of the corresponding outputs $\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_p$ as the column vectors of an $m \times p$ matrix B. We say that A is *identifiable* from the matrix pair (X, B) if there is a unique $m \times n$ matrix A satisfying $AX = B$.

We consider three cases, each premised on different relative sizes for n and p. First note that $p > n$ is impossible, since $\{\mathbf{X}_1, \ldots, \mathbf{X}_p\}$ is a linearly independent set in R^n. If $p = n$, then X is invertible and A is identifiable, in fact, $A = BX^{-1}$. In this case a simple modification of the Gauss–Jordan elimination method provides an algorithm for identifying A. The method rests on the observation that if $AX = B$, then $X^TA^T = B^T$. One can therefore solve for the transpose of the model matrix A by augmenting the transpose of the input matrix with the transpose of the output matrix and performing the Gauss–Jordan algorithm in the usual way:

$$[X^T : B^T] \rightarrow \cdots \rightarrow [I : A^T].$$

In the last case, when $p < n$, one suspects that the input–output information is insufficient to identify A. This is indeed so since in this case there is a vector $\mathbf{q} \in R^n$ that is orthogonal to $\mathbf{X}_1, \ldots, \mathbf{X}_p$. Let C be the $m \times n$ matrix whose first row is $\mathbf{q}^T$ and whose other rows are zero vectors. Then CX is the $m \times p$ zero matrix. Therefore, if $AX = B$, then, likewise, $(A + C)X = B$, and hence A is not identifiable from the information (X, B).

5.1.2 Activities

1. Problem A company manufactures three types of circuit boards. Each board consists of three types of components, say, diodes, transistors, and resistors. Board 1 requires two diodes, seven transistors, and three resistors. Board 2 requires three diodes, five transistors, and two resistors. Board 3 requires one diode, nine transistors, and four resistors. Each of the components has a certain unit cost. Is it possible for the costs (in some monetary unit) of the constituents of the boards to be 24, 20, and 15 (for diodes, transistors, and resistors, respectively)?

2. Problem Referring to Problem 1, suppose the constituent costs of the three types of boards are 24, 20, and 28 dollars, respectively, and that the sum of the unit costs of the three types of components is a minimum. Estimate (to the nearest cent) the unit costs of the components.

3. Question Does small-determinant imply ill-conditioned?

4. Problem Given an arbitrarily small positive number ϵ, construct a matrix A with $\det A = \epsilon$ and $cond(A) = 1 + \epsilon$.

5. Computation Suppose the finite Laplace transform

$$F(s) = \int_0^1 e^{-su} f(u)du, \quad 0 \le s \le 1$$

is discretized to produce the $n \times n$ matrix A with $a_{ij} = e^{-ij/n^2}$, for $i, j = 1, \ldots, n$. Find $cond(A)$ for $n = 10, 20, 50$. For each such N, generate an n-vector with random components in $[-1, 1]$ and compute the vector $\mathbf{b} = A\mathbf{x}$. Plot the vector $\mathbf{x}$ and the vector $\mathbf{b}$. Explain the results.

6. Problem Show that $N(A) = N(A^TA)$ for any real matrix A.

7. Problem Show that for any real matrix A, $R(A^T)^\perp = N(A)$.

8. Problem Show that if W and V are subspaces of R^n and $W^\perp = V^\perp$, then $W = V$.

9. Problem Use Problem 8 to show that $R(A^T) = R(A^TA)$ for any real matrix A.

10. Problem Suppose $\mathbf{u}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$, and let $P\mathbf{u}$ be the orthogonal projection of $\mathbf{u}$ onto $N(A)$ (i.e., $P\mathbf{u} = \sum_{j=1}^{k} (\mathbf{u}, \mathbf{v}^{(j)})\mathbf{v}^{(j)}$, where $\{\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(k)}\}$ is an orthonormal basis for $N(A)$). Show that $\mathbf{u} - P\mathbf{u}$ is a least-squares solution that is orthogonal to $N(A)$.

11. Exercise Suppose $\mathbf{b} = [1, 0, 2]^T$ and

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 1 & 1 \end{bmatrix}.$$

Show that the system $A\mathbf{x} = \mathbf{b}$ has no ordinary solution, but that it does have a unique least-squares solution.

12. Exercise Suppose $\mathbf{b} = [1, 0, 1]^T$ and

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Show that the system $A\mathbf{x} = \mathbf{b}$ has no ordinary solution, but that it does have infinitely many least-squares solutions. Find the least-squares solution that is orthogonal to the null-space.

13. Exercise Suppose B is a real, symmetric matrix. A given nonzero $n \times 1$ vector $\mathbf{y}$ is an eigenvector of B if the n equations in the single unknown μ

$$\mathbf{y}\mu = B\mathbf{y}$$

have an ordinary solution. Show that for each nonzero $\mathbf{y}$ this equation has the unique least-squares solution

$$\mu = \frac{(B\mathbf{y}, \mathbf{y})}{\|\mathbf{y}\|^2}.$$

This quantity is called the *Rayleigh quotient* for $\mathbf{y}$, and it may be taken as an approximation to an eigenvalue of B.

14. Problem Show that there are infinitely many 2×3 matrices that produce the outputs $[1, 1]^T$ and $[0, 1]^T$, given the respective inputs $[1, 1, -1]^T$ and $[-1, 1, 0]^T$.