mean. The median is insensitive to a data point, which, by some
r infinity. In fact, median and mean are equal when the prob-
is symmetrical. For a sample of n numbers $(x_i, i = 1, 2, \ldots, n)$,
nay be estimated by reordering the numbers from smallest to largest
ing the number in the middle as the estimate of the median $\hat{m}_1$.
the reordered x_i be denoted by x_i' where $x_i' \leq x_{i+1}'$. Then we have
w it turns out that without knowledge of the probability density
e random variables x_i we will still be able to compute the proba-
rue median m_1 is contained in the interval

$$x'_{n/2-\alpha\sqrt{n}} < m_1 < x'_{n/2+\alpha\sqrt{n}} \qquad (4\text{-}6\text{-}1)$$

et $\alpha = 1$ and $N = 100$, the assertion is that we can now calculate
hat the true median m_1 lies between the 40th and the 60th percentile
ie trick is this: Define a new random variable

$$y = \text{step}\,(x - m_1) \qquad (4\text{-}6\text{-}2)$$

on equals $+1$ if $x > m_1$ and equals 0 if $x < m_1$. The new random
on only values of zero and one with equal probability; thus we
bility function even though we may not know the probability
random variable x. Now define a third random variable s as

$$s = \sum_{i=1}^{n} y_i$$

zero or one, then s must be an integer between zero and n. Further-
bility that s takes the value j is given by the coefficient of Z^j of
v the probability that s lies in the interval $n/2 - \alpha\sqrt{n} < s < n/2 +$
determined by adding the required coefficients of Z^j, and this
definition equal to the probability that the median m_1 lies in the
For $\alpha = 1$ and large n this probability works out to about 95

5

MATRICES AND MULTICHANNEL TIME SERIES

Familiarity with matrices is essential to computer modeling in both physical and social sciences. As this is a big subject covered by many excellent texts at all levels, our review will be a quick one. We focus on those properties required in the succeeding chapters. We avoid proofs, and although constructions given should be useful in most situations, there will be occasional matrices (which we will dismiss as pathological cases) in which our constructions will fail. In practice, the user should always check computed results. Unfortunately, the so-called pathological cases arise in practice far more often than might be expected. When matrix difficulties arise, the first tendency of the scientist is to use a higher-precision arithmetic. In the author's experience, physically meaningful calculations rarely require high precision. When higher precision seems to be needed, it is often because something is happening physically which shows that the problem being solved is a poorly posed problem. If a slight change in the problem should not make a drastic change in the answer, then it may happen that a different organization of the calculations will obviate the need for high precision. Anyway, our discussion here will focus on the nonpathological cases, but the reader is warned that pathological cases will certainly be encountered in practice and when they are they will be a stern test of the reader's mathematical knowledge and physical insight.

OF MATRICES

eous equations may be written as

$$\mathbf{Ax}=\mathbf{b} \qquad (5\text{-}1\text{-}1a)$$

are matrix (nonsquare matrices are taken up in Chap. 6 on least
d **b** are column vectors. In a 2×2 case, (5-1-1*a*) becomes

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}=x_1\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}+x_2\begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} \qquad (5\text{-}1\text{-}1b)$$

) has a simple geometric interpretation. The two columns of the
lumn **b** are regarded as vectors in a plane. Equation (5-1-1*b*) says
column vector plus x_2 times the second column vector equals the
Difficulty arises when the two column vectors of the matrix
direction. Unless **b** just happens to be that direction, no solution
The same thing may be said about the general case. A solution
1-1*a*) exists if **b** lies in the space spanned by the columns of **A**.
situations the matrix **A** and the column **b** arise from independent
that it is often reasonable to require the columns of **A** to span a
contain an arbitrary **b** vector. If **A** is an $n\times n$ matrix, then its
ired to span an n-dimensional space. In particular, the n-dimen-
ped with edges given by the columns of **A** should not have a zero
olume is given by the determinant of **A**.
of simultaneous equations which arises frequently in practice is
ogeneous equations

$$\mathbf{Ax}=\mathbf{0} \qquad (5\text{-}1\text{-}2)$$

as the solution $\mathbf{x}=\mathbf{0}$ which is often called the trivial solution.
ve nontrivial solution values for **x** the determinant of A should
hat the columns of A do not span an n-dimensional space. We will
subject of actually solving sets of simultaneous equations.
ul feature of matrices is that their elements may be not only
they may be other matrices. Viewed differently, a big matrix may
o smaller submatrices. A surprising thing is that the product of
e same whether there are partitions or not. Study the identity

$$\left[\begin{array}{cc|c} a & b & c \\ d & e & f \end{array}\right]\left[\begin{array}{cc} g & h \\ i & j \\ \hline k & l \end{array}\right]=\begin{bmatrix} a & b \\ d & e \end{bmatrix}\begin{bmatrix} g & h \\ i & j \end{bmatrix}+\begin{bmatrix} c \\ f \end{bmatrix}[k \quad l] \qquad (5\text{-}1\text{-}3)$$

nation notation, the left-hand side of (5-1-3) means

$$C_{ik}=\sum_{j=1}^{3} A_{ij}B_{jk} \qquad (5\text{-}1\text{-}4)$$

-hand side means

$$C_{ik}=\sum_{j=1}^{2} A_{ij}B_{jk}+\sum_{j=3}^{3} A_{ij}B_{jk} \qquad (5\text{-}1\text{-}5)$$

Equations (5-1-4) and (5-1-5) are obviously the same; this shows that this partitioning of a matrix product is merely rearranging the terms. Partitioning does not really do anything at all from a mathematical point of view, but it is extremely important from the point of view of computation or discussion.

We now utilize matrix partitioning to develop the bordering method of matrix inversion. The bordering method is not the fastest or the most accurate method but it is quite simple, even for nonsymmetric complex-valued matrices, and it also gives the determinant and works for homogeneous equations. The bordering method proceeds by recursion. Given the inverse to a $k\times k$ matrix, the method shows how to find the inverse of a $(k+1)\times(k+1)$ matrix, which is the same old $k\times k$ matrix with an additional row and column attached to its borders. Specifically, **A**, **e**, **f**, g, and $\mathbf{A}^{-1}$ are taken to be known in (5-1-6). The task is to find **W**, **x**, **y**, and z.

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{f} \\ \hline \mathbf{e} & g \end{array}\right]\left[\begin{array}{c|c} \mathbf{W} & \mathbf{y} \\ \hline \mathbf{x} & z \end{array}\right]=\left[\begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} & 1 \end{array}\right] \qquad (5\text{-}1\text{-}6)$$

The first thing to do is multiply the partitions in (5-1-6) together. For the first column of the product we obtain

$$\mathbf{AW}+\mathbf{fx}=\mathbf{I} \qquad (5\text{-}1\text{-}7)$$

$$\mathbf{eW}+g\mathbf{x}=\mathbf{0} \qquad (5\text{-}1\text{-}8)$$

A choice of **W** of

$$\mathbf{W}=\mathbf{A}^{-1}(\mathbf{I}-\mathbf{fx}) \qquad (5\text{-}1\text{-}9)$$

leads to (5-1-7) being satisfied identically. This leaves **x** still unknown, but we may find it by substituting (5-1-9) into (5-1-8)

$$\mathbf{x}=\frac{\mathbf{eA}^{-1}}{\mathbf{eA}^{-1}\mathbf{f}-g} \qquad (5\text{-}1\text{-}10)$$

Now, to get the column unknowns **y** and z, we compute the second column of the product (5-1-6)

$$\mathbf{Ay}+\mathbf{f}z=\mathbf{0} \qquad (5\text{-}1\text{-}11)$$

$$\mathbf{ey}+gz=1 \qquad (5\text{-}1\text{-}12)$$

Multiply (5-1-11) by $\mathbf{A}^{-1}$

$$\mathbf{y}=-\mathbf{A}^{-1}\mathbf{f}z \qquad (5\text{-}1\text{-}13)$$

This gives the column vector **y** within a scale factor z. To get the scale factor, we insert (5-1-13) into (5-1-12)

$$-\mathbf{eA}^{-1}\mathbf{f}z+gz=1$$

$$z=\frac{1}{g-\mathbf{eA}^{-1}\mathbf{f}} \qquad (5\text{-}1\text{-}14)$$

```
      SUBROUTINE CMAINE(N,B,A)
C     A=MATRIX INVERSE OF B
      COMPLEX B,A,C,R,DEL
      DIMENSION  A(N,N),B(N,N),R(100),C(100)
      DO 10 I=1,N
      DO 10 J=1,N
10    A(I,J)=0.
      DO 40 L=1,N
      DEL=B(L,L)
      DO 30 I=1,L
      C(I)=0.
      R(I)=0.
      DO 20 J=1,L
      C(I)=C(I)+A(I,J)*B(J,L)
20    R(I)=R(I)+B(L,J)*A(J,I)
30    DEL=DEL-B(L,I)*C(I)
      C(L)=-1.
      R(L)=-1.
      DO 40 I=1,L
      C(I)=C(I)/DEL
      DO 40 J=1,L
40    A(I,J)=A(I,J)+C(I)*R(J)
      RETURN
      END
```

er program for matrix
on the bordering

shown that the determinant of the matrix being inverted is given
er all the bordering steps of the denominator of (5-1-14). Thus,
ing the recursion the denominator of (5-1-14) goes to zero, the
and the calculation cannot proceed.

rize the recursion: One begins with the upper left-hand corner of
rner is a scalar and its inverse is trivial. Then it is considered
a row and a column as shown in (5-1-6). Next, we find the
2 matrix. The process is continued as long as one likes. A typical
te z by (5-1-14) and then compute $\mathbf{A}^{-1}$ of one larger size by

$$\mathbf{A}^{-1} \leftarrow \begin{bmatrix} \mathbf{A}^{-1} \\ \hdashline \text{zeros} \end{bmatrix} + z \begin{bmatrix} \mathbf{A}^{-1}\mathbf{f} \\ \hdashline -1 \end{bmatrix} [\mathbf{e}\mathbf{A}^{-1} \vdots -1] \qquad (5\text{-}1\text{-}15)$$

s made up from (5-1-9), (5-1-10), and (5-1-13). A Fortran com-
achieve this is shown in Fig. 5-1.

ve to see what becomes of $\mathbf{A}^{-1}$ if $\mathbf{A}$ is perturbed steadily in such
terminant of $\mathbf{A}$ becomes singular. If the element g in the matrix
d closer and closer to $\mathbf{e}\mathbf{A}^{-1}\mathbf{f}$, then we see from (5-1-14) that z
What is interesting is that the second term in (5-1-15) comes to
, and the inverse tends to infinity times the product of a column

pressions $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ or $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ in the limit of small z^{-1} tend to

$$\mathbf{A}\mathbf{c}\mathbf{r} = z^{-1}\mathbf{I} \qquad (5\text{-}1\text{-}16)$$

$$\mathbf{c}\mathbf{r}\mathbf{A} = z^{-1}\mathbf{I} \qquad (5\text{-}1\text{-}17)$$

In the usual case (rank $\mathbf{A} = n-1$, not rank $\mathbf{A} < n-1$) where neither $\mathbf{c}$ nor $\mathbf{r}$ vanish identically, (5-1-16) and (5-1-17) in the limit $z^{-1} = 0$ become

$$\mathbf{A}\mathbf{c} = \mathbf{0} \qquad (5\text{-}1\text{-}18)$$

$$\mathbf{r}\mathbf{A} = \mathbf{0} \qquad (5\text{-}1\text{-}19)$$

In summary, then, to solve an ordinary set of simultaneous equations like (5-1-1), one may compute the matrix inverse of $\mathbf{A}$ by the bordering method and then multiply (5-1-1) by $\mathbf{A}^{-1}$ obtaining

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \qquad (5\text{-}1\text{-}20)$$

In the event $\mathbf{b}$ vanishes, we are seeking the solution to homogeneous equations and we expect that z will explode in the last step of the bordering process. (If it happens earlier, one should be able to rearrange things.) The solution is then given by the column $\mathbf{c}$ in (5-1-18).

The row homogeneous equations of (5-1-19) was introduced because such a set arises naturally for the solution to the row eigenvectors of a nonsymmetric matrix. In the next section, we will go into some detailed properties of eigenvectors. A column eigenvector $\mathbf{c}$ of a matrix $\mathbf{A}$ is defined by the solution to

$$\mathbf{A}\mathbf{c} = \lambda\mathbf{c} \qquad (5\text{-}1\text{-}21)$$

where λ is the so-called eigenvalue. At the same time, one also considers a row eigenvector equation

$$\mathbf{r}\mathbf{A} = \lambda\mathbf{r} \qquad (5\text{-}1\text{-}22)$$

To have a solution for (5-1-21) or (5-1-22), one must have $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. After finding the roots λ_j of the polynomial $\det(\mathbf{A} - \lambda\mathbf{I})$, one may form a new matrix $\mathbf{A}'$ for each λ_j by

$$\mathbf{A}' = \mathbf{A} - \lambda_j\mathbf{I} \qquad (5\text{-}1\text{-}23)$$

then the solution to

$$\mathbf{A}'\mathbf{x} = \mathbf{0} \qquad (5\text{-}1\text{-}24)$$

arises from the column $\mathbf{c}$ at the last step of the bordering. It is the column eigenvector. Likewise, the row eigenvector is the row in the last step of the bordering algorithm.

EXERCISES

1 Indicate the sizes of all the matrices in equations (5-1-7) to (5-1-14)
2 Show how (5-1-15) follows from (5-1-9), (5-1-10), (5-1-13), and (5-1-14).

TER'S MATRIX THEOREM

rem provides a rapid way to calculate functions of a matrix. nctions of a matrix of frequent occurrence are $\mathbf{A}^{-1}$ and $\mathbf{A}^N$(for N re matrix functions which are very important in wave propagation [2]. Before going into the somewhat abstract proof of Sylvester's ll take up a mumerical example. Consider the matrix

$$\mathbf{A}=\begin{bmatrix}3 & -2\\ 1 & 0\end{bmatrix} \tag{5-2-1}$$

sary to have the column eigenvectors and the eigenvalues of this e given by

$$\begin{bmatrix}3 & -2\\ 1 & 0\end{bmatrix}\begin{bmatrix}1\\ 1\end{bmatrix}=1\begin{bmatrix}1\\ 1\end{bmatrix} \tag{5-2-2}$$

$$\begin{bmatrix}3 & -2\\ 1 & 0\end{bmatrix}\begin{bmatrix}2\\ 1\end{bmatrix}=2\begin{bmatrix}2\\ 1\end{bmatrix} \tag{5-2-3}$$

x **A** is not symmetric, it has row eigenvectors which differ from the These are

$$[-1 \quad 2]\begin{bmatrix}3 & -2\\ 1 & 0\end{bmatrix}=1[-1 \quad 2] \tag{5-2-4}$$

$$[1 \quad -1]\begin{bmatrix}3 & -2\\ 1 & 0\end{bmatrix}=2[1 \quad -1] \tag{5-2-5}$$

iate equations (5-2-2) through (5-2-5) by

$$\begin{aligned}\mathbf{A}\ \mathbf{c}_1&=\lambda_1\ \mathbf{c}_1\\ \mathbf{A}\ \mathbf{c}_2&=\lambda_2\ \mathbf{c}_2\\ \mathbf{r}_1\ \mathbf{A}&=\lambda_1\ \mathbf{r}_1\\ \mathbf{r}_2\ \mathbf{A}&=\lambda_2\ \mathbf{r}_2\end{aligned} \tag{5-2-6}$$

observe that **r** or **c** could be multiplied by an arbitrary scale factor ld still be valid. The eigenvectors are said to be normalized if scale n chosen so that $\mathbf{r}_1\cdot\mathbf{c}_1=1$ and $\mathbf{r}_2\cdot\mathbf{c}_2=1$. It will be observed that $_2\cdot\mathbf{c}_1=0$, a general result to be established in the exercises. sider the behavior of the matrix $\mathbf{c}_1\mathbf{r}_1$.

$$\mathbf{c}_1\mathbf{r}_1=\begin{bmatrix}1\\ 1\end{bmatrix}[-1 \quad 2]=\begin{bmatrix}-1 & 2\\ -1 & 2\end{bmatrix}$$

his matrix is the matrix itself, for example its square.

$$\begin{bmatrix}-1 & 2\\ -1 & 2\end{bmatrix}\begin{bmatrix}-1 & 2\\ -1 & 2\end{bmatrix}=\begin{bmatrix}-1 & 2\\ -1 & 2\end{bmatrix}$$

This property is called idempotence (Latin for self-power). It arises because $(\mathbf{c}_1\mathbf{r}_1)(\mathbf{c}_1\mathbf{r}_1)=\mathbf{c}_1(\mathbf{r}_1\cdot\mathbf{c}_1)\mathbf{r}_1=\mathbf{c}_1\mathbf{r}_1$. The same thing is of course true of $\mathbf{c}_2\mathbf{r}_2$. Now notice that the matrix $\mathbf{c}_1\mathbf{r}_1$ is "perpendicular" to the matrix $\mathbf{c}_2\mathbf{r}_2$, that is

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

since $\mathbf{r}_2$ and $\mathbf{c}_2$ are perpendicular.

Sylvester's theorem says that any function f of the matrix **A** may be written

$$f(\mathbf{A})=f(\lambda_1)\mathbf{c}_1\mathbf{r}_1+f(\lambda_2)\mathbf{c}_2\mathbf{r}_2$$

The simplest example is $f(\mathbf{A})=\mathbf{A}$

$$\begin{aligned}\mathbf{A}&=\lambda_1\mathbf{c}_1\mathbf{r}_1+\lambda_2\mathbf{c}_2\mathbf{r}_2\\ &=1\begin{bmatrix}-1 & 2\\ -1 & 2\end{bmatrix}+2\begin{bmatrix}2 & -2\\ 1 & -1\end{bmatrix}=\begin{bmatrix}3 & -2\\ 1 & 0\end{bmatrix}\end{aligned} \tag{5-2-7}$$

Another example is

$$\mathbf{A}^2=1^2\begin{bmatrix}-1 & 2\\ -1 & 2\end{bmatrix}+2^2\begin{bmatrix}2 & -2\\ 1 & -1\end{bmatrix}=\begin{bmatrix}7 & -6\\ 3 & -2\end{bmatrix}$$

The inverse is

$$\mathbf{A}^{-1}=1^{-1}\begin{bmatrix}-1 & 2\\ -1 & 2\end{bmatrix}+2^{-1}\begin{bmatrix}2 & -2\\ 1 & -1\end{bmatrix}=\tfrac{1}{2}\begin{bmatrix}0 & 2\\ -1 & 3\end{bmatrix}$$

The identity matrix may be expanded in terms of the eigenvectors of the matrix **A**.

$$\mathbf{A}^0=\mathbf{I}=1^0\begin{bmatrix}-1 & 2\\ -1 & 2\end{bmatrix}+2^0\begin{bmatrix}2 & -2\\ 1 & -1\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

Before illustrating some more complicated functions let us see what it takes to prove Sylvester's theorem. We will need one basic result which is in all the books on matrix theory, namely, that most matrices (see exercises) can be diagonalized. In terms of our 2×2 example this takes the form

$$\begin{bmatrix}\mathbf{r}_1\\ \hline \mathbf{r}_2\end{bmatrix}\mathbf{A}[\mathbf{c}_1\mid\mathbf{c}_2]=\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix} \tag{5-2-8}$$

where

$$\begin{bmatrix}\mathbf{r}_1\\ \hline \mathbf{r}_2\end{bmatrix}[\mathbf{c}_1\mid\mathbf{c}_2]=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \tag{5-2-9}$$

Since a matrix commutes with its inverse, (5-2-9) implies

$$[\mathbf{c}_1\mid\mathbf{c}_2]\begin{bmatrix}\mathbf{r}_1\\ \hline \mathbf{r}_2\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \tag{5-2-10}$$

2-8) by the row matrix and premultiply by the column matrix. ve get

$$\mathbf{A} = [\mathbf{c}_1 \mid \mathbf{c}_2]\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}\begin{bmatrix}\mathbf{r}_1\\ \mathbf{r}_2\end{bmatrix} \qquad (5\text{-}2\text{-}11)$$

) is (5-2-7) in disguise, as we can see by writing (5-2-11) as

$$\mathbf{A} = [\mathbf{c}_1 \mid \mathbf{c}_2]\left\{\begin{bmatrix}\lambda_1 & 0\\ 0 & 0\end{bmatrix} + \begin{bmatrix}0 & 0\\ 0 & \lambda_2\end{bmatrix}\right\}\begin{bmatrix}\mathbf{r}_1\\ \mathbf{r}_2\end{bmatrix}$$

$$= [\mathbf{c}_1 \mid \mathbf{c}_2]\begin{bmatrix}\lambda_1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\mathbf{r}_1\\ \mathbf{r}_2\end{bmatrix} + [\mathbf{c}_1 \mid \mathbf{c}_2]\begin{bmatrix}0 & 0\\ 0 & \lambda_2\end{bmatrix}\begin{bmatrix}\mathbf{r}_1\\ \mathbf{r}_2\end{bmatrix}$$

$$= \lambda_1\mathbf{c}_1\mathbf{r}_1 + \lambda_2\mathbf{c}_2\mathbf{r}_2$$

ve have

$$\mathbf{A}^2 = (\lambda_1\mathbf{c}_1\mathbf{r}_1 + \lambda_2\mathbf{c}_2\mathbf{r}_2)(\lambda_1\mathbf{c}_1\mathbf{r}_1 + \lambda_2\mathbf{c}_2\mathbf{r}_2)$$

normality of $\mathbf{c}_1\mathbf{r}_1$ and $\mathbf{c}_2\mathbf{r}_2$ this reduces to

$$\mathbf{A}^2 = \lambda_1^2\mathbf{c}_1\mathbf{r}_1 + \lambda_2^2\mathbf{c}_2\mathbf{r}_2$$

-2-11) can be used to prove Sylvester's theorem for any polynomial Clearly, there is nothing peculiar about 2 × 2 matrices either. This Likewise, one may consider infinite series functions in **A**. Since ion can be made up of infinite series, we can consider also trans- ons like sine, cosine, exponential.

s arise naturally as the solutions to differential equations. Con- differential equation

$$\frac{d}{dx}\mathbf{E} = \mathbf{A}\mathbf{E} \qquad (5\text{-}2\text{-}12)$$

verify the power series solution

$$\mathbf{E} = \mathbf{I} + \mathbf{A}x + \frac{\mathbf{A}^2x^2}{2!} + \cdots$$

r series definition of an exponential function. If the matrix **A** is majority which can be diagonalized, then the exponential can be ressed by Sylvester's theorem. For the numerical example we have , we have

$$\mathbf{E} = e^x\begin{bmatrix}-1 & 2\\ -1 & 2\end{bmatrix} + e^{2x}\begin{bmatrix}2 & -2\\ 1 & -1\end{bmatrix}$$

matrix is a solution to the differential equation (5-2-12) without aries. It frequently happens that physics gives one a differential

$$\frac{d}{dx}\begin{bmatrix}y_1\\ y_2\end{bmatrix} = \mathbf{A}\begin{bmatrix}y_1\\ y_2\end{bmatrix} \qquad (5\text{-}2\text{-}13)$$

subject to two boundary conditions on either of y_1 or y_2 or a combination. One may verify that

$$\begin{bmatrix}y_1\\ y_2\end{bmatrix} = e^{\mathbf{A}x}\begin{bmatrix}k_1\\ k_2\end{bmatrix}$$

is the solution to (5-2-13) for arbitrary constants k_1 and k_2. Boundary conditions are then used to determine the numerical values of k_1 and k_2. Note that k_1 and k_2 are just $y_1(x = 0)$ and $y_2(x = 0)$.

An interesting situation arises with the square root of a matrix. A 2 × 2 matrix like **A** will have four square roots because there are four possible combinations for choice of plus or minus signs on $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$. In general, an $n \times n$ matrix has 2^n square roots. An important application arises in a later chapter, where we will deal with the differential operator $(k^2 + \partial^2/\partial x^2)^{1/2}$. The square root of an operator is explained in very few books and few people even know what it means. The best way to visualize the square root of this differential operator is to relate it to the square root of the matrix **M** where

$$\mathbf{M} = k^2\begin{bmatrix}1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1\end{bmatrix} + \frac{1}{(\Delta x)^2}\begin{bmatrix}? & ? & & & \\ 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & 1 & -2 & 1\\ & & & ? & ?\end{bmatrix}$$

The right-hand matrix is a second difference approximation to a second partial derivative. Let us define

$$\mathbf{M} = k^2\mathbf{I} + \mathbf{T}$$

Clearly we wish to consider **M** generalized to a very large size so that the end effects may be minimized. In concept, we can make **M** as large as we like and for any size we can get $2^{\mathbf{M}}$ square roots. In practice there will be only two square roots of interest, one with the plus roots of all the eigenvalues and the other with all the minus roots. How can we find these "principal value" square roots? An important case of interest is where we can use the binomial theorem so that

$$(k^2\mathbf{I} + \mathbf{T})^{1/2} = \pm k\left(\mathbf{I} + \frac{\mathbf{T}}{k^2}\right)^{1/2}$$

$$= \pm k\left(\mathbf{I} + \frac{\mathbf{T}}{2k^2} - \frac{\mathbf{T}^2}{8k^4} + \cdots\right)$$

The result is justified by merely squaring the assumed square root. Alternatively, it may be justified by means of Sylvester's theorem. It should be noted that on squaring the assumed square root one utilizes the fact that **I** and **T** commute. We are led to the idea that the square root of the differential operator may be interpreted as

$$\left(k^2 + \frac{\partial^2}{\partial x^2}\right)^{1/2} = k + \frac{1}{2k}\frac{\partial^2}{\partial x^2} + \cdots$$

k is not a function of x. If k is a function of x, the square root of operator still has meaning but is not so simply computed with the rem.

(5-2-6b) by $\mathbf{r}_1$ and postmultiply (5-2-6c) by $\mathbf{c}_2$, then subtract. Is $\lambda_1 \neq \lambda_2$ a ndition for $\mathbf{r}_1$ and $\mathbf{c}_2$ to be perpendicular? Is it a sufficient condition? ayley-Hamilton theorem, that is, if

$$0 = f(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = p_0 + p_1\lambda + p_2\lambda^2 + \cdots p_n\lambda^n$$

$$f(\mathbf{A}) = p_0 + p_1\mathbf{A} + p_2\mathbf{A}^2 + \cdots + p_n\mathbf{A}^n = 0$$

for a general 2×2 matrix $\mathbf{A}$, for which

$$\lambda_1 \neq \lambda_2,$$
$$\mathbf{c}_1\mathbf{r}_1 = (\lambda_2\mathbf{I} - \mathbf{A})/(\lambda_2 - \lambda_1)$$

λ_2 are eigenvalues of $\mathbf{A}$. What is the general form for $\mathbf{c}_2\mathbf{r}_2$?
etric matrix it can be shown that there is always a complete set of eigen-
roblem sometimes arises with nonsymmetric matrices. Study the matrix

$$\begin{bmatrix} 1 & 1-\varepsilon^2 \\ -1 & 3 \end{bmatrix}$$

ee why one eigenvector is lost. This is called a defective matrix. (This om T. R. Madden.)
ty of wave-propagation problems in a stratified medium reduce to the

$$\frac{d}{dx}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 & b \\ a & 0 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

dependence of the solution when ab is positive? When ab is negative? d b are independent of x. Use Sylvester's theorem. What would it take ctive matrix? What are the solutions in the case of a defective matrix? atrix of the form $\mathbf{I} + \mathbf{v}\mathbf{v}^T$ where $\mathbf{v}$ is a column vector and $\mathbf{v}^T$ is its transpose. $)^{-1}$ in terms of a power series in $\mathbf{v}\mathbf{v}^T$. [Note that $(\mathbf{v}\mathbf{v}^T)^N$ collapses to $\mathbf{v}\mathbf{v}^T$ g factor, so the power series reduces considerably.]
ng "cross-product" matrix often arises in electrodynamics. Let B_z)

$$\mathbf{U} = \frac{1}{\sqrt{\mathbf{B}\cdot\mathbf{B}}}\begin{bmatrix} 0 & -B_z & B_y \\ B_z & 0 & -B_x \\ -B_y & B_x & 0 \end{bmatrix}$$

elements of $\mathbf{I} + \mathbf{U}^2$.
$\mathbf{U}(\mathbf{I} + \mathbf{U}^2) = 0$ or $\mathbf{U}^3 = -\mathbf{U}$.

(c) Let $\mathbf{v}$ be an arbitrary vector. In what geometrical directions do $\mathbf{U}\mathbf{v}$, $\mathbf{U}^2\mathbf{v}$, and $(\mathbf{I} + \mathbf{U}^2)\mathbf{v}$ point?
(d) What are the eigenvalues of $\mathbf{U}$. [HINT: Use part (b).]
(e) Why cannot $\mathbf{U}$ be canceled from $\mathbf{U}^3 = -\mathbf{U}$?
(f) Verify that the idempotent matrices of $\mathbf{U}$ are

$$\mathbf{c}_1\mathbf{r}_1 = (\mathbf{I} + \mathbf{U}^2)$$

$$\mathbf{c}_2\mathbf{r}_2 = \tfrac{1}{2}(i\mathbf{U} - \mathbf{U}^2)$$

$$\mathbf{c}_3\mathbf{r}_3 = \tfrac{1}{2}(-i\mathbf{U} - \mathbf{U}^2)$$

5-3 MATRIX FILTERS, SPECTRA, AND FACTORING

Two time series can be much more interesting than one because of the possibility of interactions between them. The general linear model for two series is depicted in Fig. 5-2

The filtering operation in the figure can be expressed as a matrix times vector operation, where the elements of the matrix and vectors are Z transform polynomials. That is,

$$\begin{bmatrix} Y_1(Z) \\ Y_2(Z) \end{bmatrix} = \begin{bmatrix} B_{11}(Z) & B_{12}(Z) \\ B_{21}(Z) & B_{22}(Z) \end{bmatrix}\begin{bmatrix} X_1(Z) \\ X_2(Z) \end{bmatrix} \tag{5-3-1}$$

One fact which is obvious but unfamiliar is that a matrix with polynomial elements is exactly the same thing as a polynomial with matrix coefficients. This is illustrated by the example:

$$\begin{bmatrix} 1+Z+2Z^2 & Z \\ 1 & 1+Z^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}Z + \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}Z^2$$

Now we can address ourselves to the inverse problem; given a filter $\mathbf{B}$ and the outputs $\mathbf{Y}$ how can we find the inputs $\mathbf{X}$? The solution is analogous to that of single time series. Let us regard $\mathbf{B}(Z)$ as a matrix of polynomials. One knows, for example, that the inverse of any 2×2 matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{is} \quad \frac{\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}}{ad - bc}$$

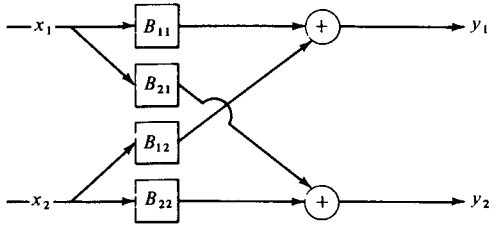

FIGURE 5-2
Two time series x_1 and x_2 input to a matrix of four filters illustrates the general linear model of multichannel filtering.

$= \mathbf{BX}$ may be solved for $\mathbf{X}$ as $\mathbf{X} = \mathbf{B}^{-1}\mathbf{Y}$ where

$$= \begin{bmatrix} B_{22}(Z) & -B_{12}(Z) \\ -B_{21}(Z) & B_{11}(Z) \end{bmatrix} \frac{1}{B_{11}(Z)B_{22}(Z) - B_{12}(Z)B_{21}(Z)}$$

or is a scalar. We have treated scalar denominators before. If all tside the unit circle, we can use an ordinary power series for the ise, it is not minimum-phase and we use a Laurent series.

generalizes to many time series, the numerator matrix is the so- natrix and the denominator is the determinant. The adjoint matrix vithout the use of any division operations. In other words, elements natrix are in the form of sums of products. For this reason, we may terion for a minimum-phase matrix wavelet is that the determinant rm has no zeros inside the unit circle.

(5-3-1) is a useful description of Fig. 5-2 in most applications. ne applications (where the filter is an unknown to be determined), rm of (5-3-1) is more useful. If b_{12} was interchanged with b_{21} in uld use the "row data" expression

$$[Y_1(Z) \quad Y_2(Z)] = [X_1(Z) \quad X_2(Z)] \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \tag{5-3-2}$$

we have generalized the concept of filtering from scalar-valued time -valued series, it is natural to generalize the idea of spectrum. For ime functions, the spectrum is a matrix called the *spectral matrix* by

$$\mathbf{R}(\omega) = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} = \begin{bmatrix} \overline{Y}_1\left(\frac{1}{Z}\right) \\ \overline{Y}_2\left(\frac{1}{Z}\right) \end{bmatrix} [Y_1(Z) \quad Y_2(Z)] \tag{5-3-3}$$

$$= \begin{bmatrix} \overline{Y}_1 Y_1 & \overline{Y}_1 Y_2 \\ \overline{Y}_2 Y_1 & \overline{Y}_2 Y_2 \end{bmatrix}$$

ced that the vector times vector operation defining (5-3-3) is an t" rather than the more usually occurring "inner product." The e spectral matrix $\mathbf{R}$ contain the usual auto-spectrum of each channel. ontain the cross spectrum. Because (5-3-3) is an outer product, the ar. Now, instead of taking $[Y_1(Z) \quad Y_2(Z)]$ to have a time function nount of energy, let us suppose the filter inputs to (5-3-2), namely e made up of random numbers, independently drawn from some ction at every point in time. In this case, $y_1(t)$ and $y_2(t)$ are random their spectral matrix is defined like (5-3-3) but taking an expectation he ensemble). We have

$$\mathbf{R}(\omega) = E\begin{bmatrix} \overline{Y}_1 \\ \overline{Y}_2 \end{bmatrix} [Y_1 \quad Y_2]$$

substituting from (5-3-2)

$$\mathbf{R}(\omega) = E\begin{bmatrix} \bar{B}_{11} & \bar{B}_{21} \\ \bar{B}_{12} & \bar{B}_{22} \end{bmatrix} \begin{bmatrix} \overline{X}_1 \\ \overline{X}_2 \end{bmatrix} [X_1 \quad X_2] \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

Now, grouping the ensemble summation with the random variables, we get

$$\mathbf{R} = \begin{bmatrix} \bar{B}_{11} & \bar{B}_{21} \\ \bar{B}_{12} & \bar{B}_{22} \end{bmatrix} \left\{ E\begin{bmatrix} \overline{X}_1 \\ \overline{X}_2 \end{bmatrix} [X_1 \quad X_2] \right\} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \tag{5-3-4}$$

Next, we explicitly introduce the assumption that the random numbers $x_1(t)$ are drawn independently of $x_2(t)$, thus $E(\overline{X}_2(1/z)X_1(z)) = 0$ and the assumption that $x_i(t)$ is white $E[x_i(t)x_i(t+s)] = 0$ if $s \neq 0$ and of unit variance $E[x_i(t)^2] = 1$. Thus (5-3-4) becomes

$$\mathbf{R} = \begin{bmatrix} \bar{B}_{11} & \bar{B}_{21} \\ \bar{B}_{12} & \bar{B}_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

$$= \begin{bmatrix} \bar{B}_{11} & \bar{B}_{21} \\ \bar{B}_{12} & \bar{B}_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \tag{5-3-5}$$

Of course, in practice the spectral matrix must be estimated, say $\hat{\mathbf{R}}$, from finite samples of data. This means that ensemble summation must be simulated. If the ensemble sum in (5-3-4) is simulated by summation over one point (no summation), then (5-3-4) is a singular matrix like (5-3-3). As discussed earlier, the accuracy of the elements of the spectral matrix improves with the square root of the number of ensemble elements summed over.

Single-channel spectral factorization gives insight into numerous important problems in mathematical physics. We have seen that the concepts of filter and spectrum extend in quite a useful fashion to multichannel data. It was only natural that a great deal of effort should have gone to spectral factorization of multichannel data. This effort has been successful. However, in retrospect, from the point of view of computer modeling and interpretation of observed waves, it must be admitted that multichannel spectral factorization has not been especially useful. Nevertheless a brief summary of results will be given.

The root method The author extended the single-channel root method to the multichannel case [Ref. 19]. The method is even more cumbersome in the multichannel case. A most surprising thing about the solution is that it includes a much broader result: that a polynomial with matrix coefficients may be factored. For example,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + Z\begin{bmatrix} -3 & -1 \\ 14 & -11 \end{bmatrix} + Z^2\begin{bmatrix} -4 & 4 \\ -58 & 28 \end{bmatrix}$$

factors 6 ways to

$$\left\{I + Z\begin{bmatrix} 2 & -1 \\ 20 & -7 \end{bmatrix}\right\}\left\{I + Z\begin{bmatrix} -5 & 0 \\ -6 & -4 \end{bmatrix}\right\} \quad \left\{I + Z\begin{bmatrix} -4 & 0 \\ -10 & -2 \end{bmatrix}\right\}\left\{I + Z\begin{bmatrix} 1 & -1 \\ 24 & -9 \end{bmatrix}\right\}$$

$$\left\{I + Z\begin{bmatrix} 0 & -1 \\ 10 & -7 \end{bmatrix}\right\}\left\{I + Z\begin{bmatrix} -3 & 0 \\ 4 & -4 \end{bmatrix}\right\} \quad \left\{I + Z\begin{bmatrix} -4 & 0 \\ -4 & -3 \end{bmatrix}\right\}\left\{I + Z\begin{bmatrix} 1 & -1 \\ 18 & -8 \end{bmatrix}\right\}$$

$$\left\{I + Z\begin{bmatrix} -1 & -1 \\ 8 & -7 \end{bmatrix}\right\}\left\{I + Z\begin{bmatrix} -2 & 0 \\ 6 & -4 \end{bmatrix}\right\} \quad \left\{I + Z\begin{bmatrix} -4 & 0 \\ 2 & -5 \end{bmatrix}\right\}\left\{I + Z\begin{bmatrix} 1 & -1 \\ 12 & -6 \end{bmatrix}\right\}$$

thod The only really practical method for finding an invertible vith a given spectrum is the multichannel Toeplitz method. The a is developed in a later section on multichannel time series pre-

Hilbert transform methods A number of famous mathematicians rt Wiener have worked on the problem from the point of view of p-log or the Hilbert transform method. The principal stumbling $p(A + B)$ does not equal $\exp(A)\exp(B)$ unless A and B happen it is, $AB = BA$. This is usually not the case. Although many have appeared on the subject (some stating that they solved the uthor is unaware of anyone who ever wrote a computer program fast Fourier transform speeds as does the single-channel Hilbert od.

atrix filter where the two outputs $y_1(t)$ and $y_2(t)$ are the same but for a Clearly **X** cannot be recovered from **Y**. Show that the determinant of the Find another example in which the determinant is zero at one frequency lsewhere. Explain in the time domain in what sense the input cannot rom the output.

ometer which measures temperature plus α times pressure and a pressure easures pressure plus β times the time rate of change of the temperature, x filter which converts the observed series to temperature and pressure. ther the time derivative approximation $1 - Z$ or $2(1 - Z)/(1 + Z)$.]

$$\mathbf{B}(Z) = \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix} + Z\begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$$

cients of powers of Z in $\mathbf{B}(Z)\mathbf{A}(Z) = \mathbf{I}$, to recursively develop the co- $(Z) = [\mathbf{B}(Z)]^{-1}$.

verse of

$$\begin{bmatrix} 1 + 2Z & Z \\ 3 & 2 \end{bmatrix}$$

Laurent series as is necessary.

ant of a polynomial with matrix cofficients may be independent of Z. atrix filters, this may mean that an inverse filter may have only a finite wers in Z instead of the infinite series one always has with scalar filters. ost nontrivial example you can find?

V PROCESSES

ess is another mathematical model for a time series. Until now it use in geophysics, but we will include it anyway because it might nd it is easily explained with the methods previously developed.

Suppose that x_t could take on only integer values. A given value is called a state. As time proceeds, transitions are made from the jth state to the ith state according to a probability matrix p_{ij}. The system has no memory. The next state is probabilistically dependent on the current state but independent of the previous states. The classic example is of a frog in a lily pond. As time goes by, the frog jumps from one lily pad to another. He may be more likely to jump to a near one than to a far one. He may prefer big to small pads, and he doesn't remember the last pad he was on. The state of the system is the number of the pad the frog currently occupies. The transitions are his jumps.

To begin with, one defines a state probability $\pi_i(k)$, the probability that the system will occupy state i after k transitions if its state is known at $k = 0$. We also define the transition matrix P_{ij}. Then

$$\boldsymbol{\pi}(k + 1) = \mathbf{P}\boldsymbol{\pi}(k) \qquad (5\text{-}4\text{-}1)$$

The initial-state probability vector is $\boldsymbol{\pi}(0)$. Since the initial state is known, then $\boldsymbol{\pi}(0)$ is all zeros except for a one (1) in the position corresponding to the initial state. For example, see the state-transition diagram of Fig. 5-3.

The diagram corresponds to the probability matrix

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{bmatrix}_{k+1} = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 1 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{bmatrix}_k$$

Since at each time a transition must occur, we have that the sum of the elements in a column must be unity. In other words, the row vector [1 1 1 1] is an eigenvector of **P** with unit eigenvalue. Let us define the Z transform of the probability vector as

$$\boldsymbol{\Pi}(Z) = \boldsymbol{\pi}(0) + Z\boldsymbol{\pi}(1) + Z^2\boldsymbol{\pi}(2) + \cdots \qquad (5\text{-}4\text{-}2)$$

In terms of Z transforms (5-4-1) becomes

$$\begin{aligned} [\boldsymbol{\pi}(1) + Z\boldsymbol{\pi}(2) + \cdots] &= \mathbf{P}[\boldsymbol{\pi}(0) + Z\boldsymbol{\pi}(1) + \cdots] \\ Z^{-1}[\boldsymbol{\Pi}(Z) - \boldsymbol{\pi}(0)] &= \mathbf{P}\boldsymbol{\Pi}(Z) \\ (\mathbf{I} - Z\mathbf{P})\boldsymbol{\Pi}(Z) &= \boldsymbol{\pi}(0) \\ \boldsymbol{\Pi}(Z) &= (\mathbf{I} - Z\mathbf{P})^{-1}\boldsymbol{\pi}(0) \qquad (5\text{-}4\text{-}3) \end{aligned}$$

Thus we have expressed the general solution to the problem as a function of the matrix **P** times an initial-state vector. There will be values of Z for which the

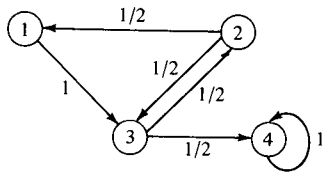

FIGURE 5-3
An example of a state-transition diagram.

o $(\mathbf{I} - Z\mathbf{P})$ does not exist. These values Z_j are given by $\det(\mathbf{I} - Z_j\mathbf{P})$ $Z_j^{-1}\mathbf{I}) = 0$. Clearly the Z_j^{-1} are the eigenvalues of $\mathbf{P}$. Utilizing rem, then, we have

$$(\mathbf{I} - Z\mathbf{P})^{-1} = \sum_j \frac{\mathbf{c}_j\mathbf{r}_j}{1 - \dfrac{Z}{Z_j}} \qquad (5\text{-}4\text{-}4)$$

ion to (5-4-4) is required if there are repeated eigenvalues. Equation ially a partial fraction expansion. A typical term has the form

$$\frac{1}{1 - \dfrac{Z}{Z_j}} = 1 + \frac{Z}{Z_j} + \left(\frac{Z}{Z_j}\right)^2 + \cdots$$

s at successive powers of Z decline with time in the form $(Z_j^{-1})^t$. f probabilities are to be bounded, the roots $1/Z_j$ must be inside the ll minimum phase). We have already shown that one of the roots nity. This leads to the "steady-state" solution $1^t = 1$. In our ıple, one can see by inspection that the steady-state probability $0 \quad 1]^T$ so the general solution is of the form

$$\boldsymbol{\pi}(t) = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} [1 \quad 1 \quad 1 \quad 1] + \sum_{j=2}^{4} Z_j^{-t}\mathbf{c}_j\mathbf{r}_j \right\} \boldsymbol{\pi}(0)$$

word of caution must be added. Occasionally defective matrices e set of eigenvectors) and for these the Sylvester theorem does not cases, the solutions turn out to contain not only terms like Z_j^{-t} like tZ^{-t} and t^2Z^{-t}. It is the same situation as that applying to ntial equations with constant coefficients. Ordinarily, the solutions $(r_i)^t$ where r_i is the ith root of the *indicial* equation but the presence ts gives rise to solutions like tr_i^t. A mathematical survey of the by Seneta [Ref. 20].

6

DATA MODELING BY LEAST SQUARES

The reconciliation of theory and data is the essence of science. An ubiquitous tool in this task is the method of least-squares fitting. Elementary calculus books generally consider the fitting of a straight line to scattered data points. Such an elementary application gives scant hint of the variety of practical problems which can be solved by the method of least squares. Some geophysical examples which we will consider include locating earthquakes, analyzing tides, expanding the earth's gravity and magnetic fields in spherical harmonics, and doing interesting things with time series. When the past of a time series is available, one may find that least squares can be used to determine a filter which predicts some future values of the time series. When a time series which has been highly predictable for a long stretch of time suddenly becomes much less predictable an "event" is said to have occurred. A filter which emphasizes such events is called a *prediction-error filter*. If one is searching for a particular dispersed wavelet in a time series, it may help to design a filter which compresses the wavelet into some more recognizable shape, an impulse for example. Such a wave-shaping filter may be designed by least squares. With multiple time series which arise from several sensors detecting waves in space, least squares may be used to find filters which respond only to certain directions and wave speeds.

Before we begin with the general theory, let us take up a simple example in