
Function fitting

The objective analysis of meteorological fields was first seriously discussed at the Princeton Institute for Advanced Study by an informal group consisting of John von Neumann, Jule Charney, Joseph Smagorinsky, George Platzman, and others in the late 1940s and early 1950s. The analysis method that seemed most promising for the primitive computers of the time was polynomial fitting, and two forms of this method were studied: regional polynomial fitting (Panofsky 1949) and local quadratic fitting (Gilchrist and Cressman 1954).

Both of these procedures are examples of a spatial analysis method that will be referred to here as function fitting. The analysis is presumed to be expanded in a finite series of ordered mathematical basis functions with unknown expansion coefficients. This series is evaluated at the observation stations, and the squared difference between the observations and the evaluated series is minimized. This least squares minimization leads to a linear relation for the unknown expansion coefficients, which can then be determined by inverting a matrix and the spatial analysis obtained by evaluating the series at the analysis gridpoints.

The early function-fitting techniques of Panofsky, Cressman, and Gilchrist were quickly superseded by the successive-correction method (Chapter 3) and later by statistical interpolation (Chapters 4 and 5). However, function fitting has been used sporadically in meteorological analysis over the years, and advanced methods are under active investigation at the present time.

The principle of least squares is the basis for both function fitting and statistical interpolation (Chapters 4 and 5). Before discussing this important principle, however, we illustrate a simple function-fitting algorithm inspired by the historic study of Gilchrist and Cressman (1954).

2.1 Local polynomial fitting

Figure 2.1 illustrates an array of 16 analysis gridpoints whose absolute spatial locations are denoted r_i . Around each gridpoint is drawn a circle called the radius

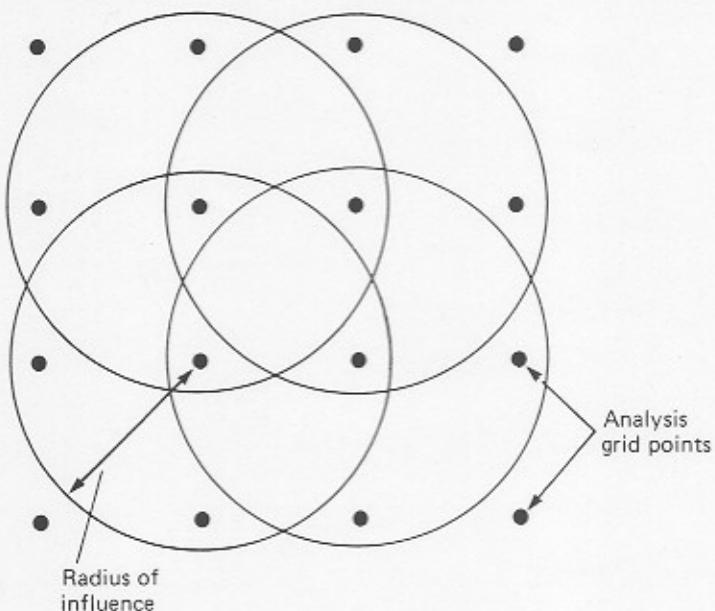


Figure 2.1 An array of 16 analysis gridpoints showing the regions of influence of each gridpoint.

of influence (R_i), and the area circumscribed is called the region of influence. In local fitting, it is assumed that the analysis at the i th gridpoint is influenced by all the observations that lie within the region of influence of that particular gridpoint.

Consider the spatial analysis of the dependent variable f . Define a local coordinate system (x, y) at the i th gridpoint such that $x = 0, y = 0$ at the gridpoint itself. Consider observations $f_O(x_k, y_k)$ (which usually contain errors) surrounding the i th gridpoint such that $x_k^2 + y_k^2 \leq R_i^2$, and suppose that there are K_i such observations.

Define the analyzed value of f within the i th region of influence as $f_A(x, y)$ and assume that it can be represented by a two-dimensional polynomial expansion of the form

$$f_A(x, y) = \sum_m \sum_n c_{mn} x^m y^n, \quad (m + n \leq M), \quad (m, n \geq 0) \quad (2.1.1)$$

where the c_{mn} are the (as yet undetermined) real expansion coefficients. Note that at the analysis gridpoint itself $x = y = 0$ and $f_A(0, 0) = c_{00}$. Form the following quadratic expression

$$I = \frac{1}{2} \sum_{k=1}^{K_i} \left[\sum_m \sum_n c_{mn} x_k^m y_k^n - f_O(x_k, y_k) \right]^2 \quad (2.1.2)$$

by formally evaluating $f_A(x, y)$ at each of the observation stations $1 \leq k \leq K_i$ in the i th region of influence. I in (2.1.2) is minimized by differentiating it with respect to each of the coefficients c_{mn} in turn and setting the results equal to zero:

$$\frac{\partial I}{\partial c_{mn}} = \sum_{k=1}^{K_i} x_k^m y_k^n \left[\sum_{\mu} \sum_{\nu} c_{\mu\nu} x_k^{\mu} y_k^{\nu} - f_O(x_k, y_k) \right] = 0$$

or

$$\sum_{\mu} \sum_{\nu} c_{\mu\nu} \sum_{k=1}^{K_i} x_k^{\mu+\nu} y_k^{\mu+\nu} = \sum_{k=1}^{K_i} x_k^m y_k^n f_O(x_k, y_k), \quad \mu + \nu \leq M \quad (2.1.3)$$

for all $m + n \leq M$. μ and ν are dummy indices.

In the study of Gilchrist and Cressman (1954), $M = 2$ and (2.1.1) was a quadratic polynomial. Define the operator

$$\overline{(\quad)} = \frac{1}{K_i} \sum_{k=1}^{K_i} (\quad)$$

Then the six expansion coefficients c_{00} , c_{10} , c_{01} , c_{20} , c_{02} , and c_{11} are given by the matrix relation

$$\begin{pmatrix} 1 & \overline{x_k} & \overline{y_k} & \overline{x_k^2} & \overline{y_k^2} & \overline{x_k y_k} \\ \overline{x_k} & \overline{x_k^2} & \overline{x_k y_k} & \overline{x_k^3} & \overline{x_k y_k^2} & \overline{x_k^2 y_k} \\ \overline{y_k} & \overline{x_k y_k} & \overline{y_k^2} & \overline{x_k^2 y_k} & \overline{y_k^3} & \overline{x_k y_k^2} \\ \overline{x_k^2} & \overline{x_k^3} & \overline{x_k^2 y_k} & \overline{x_k^4} & \overline{x_k^2 y_k^2} & \overline{x_k^3 y_k} \\ \overline{y_k^2} & \overline{x_k y_k^2} & \overline{y_k^3} & \overline{x_k^2 y_k^2} & \overline{y_k^4} & \overline{x_k y_k^3} \\ \overline{x_k y_k} & \overline{x_k^2 y_k} & \overline{x_k y_k^2} & \overline{x_k^3 y_k} & \overline{x_k y_k^3} & \overline{x_k^2 y_k^2} \end{pmatrix} \begin{pmatrix} c_{00} \\ c_{10} \\ c_{01} \\ c_{20} \\ c_{02} \\ c_{11} \end{pmatrix} = \begin{pmatrix} \overline{f_O(x_k, y_k)} \\ \overline{x_k f_O(x_k, y_k)} \\ \overline{y_k f_O(x_k, y_k)} \\ \overline{x_k^2 f_O(x_k, y_k)} \\ \overline{y_k^2 f_O(x_k, y_k)} \\ \overline{x_k y_k f_O(x_k, y_k)} \end{pmatrix} \quad (2.1.4)$$

where the elements of the matrix on the left-hand side and of the column vector on the right-hand side are all known.

The c_{mn} can be determined by inverting the 6×6 left-hand matrix (provided it is nonsingular). The analyzed value at the i th gridpoint is equal to c_{00} , as noted following (2.1.1). The process is repeated for each of the analysis gridpoints r_i .

This simple illustration considers univariate analysis (a single dependent variable) and a circular region of influence. In actual fact, Gilchrist and Cressman's (1954) algorithm differed somewhat because it was multivariate (included observations of two or more dependent variables) and had a square region of influence. The most important step in this illustration is the least squares minimization (2.1.2-3), which is examined in more detail in the next section.

2.2 Least squares estimation

If the astronomical observations and other quantities on which the computation of orbits is based were absolutely correct, the elements also, whether deduced from three or four observations, would be strictly accurate (so far indeed as the motion is supposed to take place exactly according to the laws of Kepler) and, therefore, if other observations were used, they might be confirmed, but not corrected. But since our measurements and observations are nothing more than approximations to the truth, the same must be true of all calculations resting upon them, and the highest aim of all computation made concerning concrete phenomena must be to approximate, as nearly as practicable, to the truth. But this can be accomplished in no other way than by a suitable combination of more observations than the number absolutely requisite for the determination of the

unknown quantities. This problem can only be properly undertaken when an approximate knowledge of the orbit has been already attained, which is afterward to be corrected so as to satisfy all of the observations in the most accurate manner possibly. (Gauss 1809, tr. 1963)

The earliest stimulus for the development of estimation theory was provided by astronomical studies in which planet and comet motion was inferred from telescopic measurement data. The motions of these bodies can be completely characterized by six parameters, and the problem that was considered was that of estimating the values of these parameters from the measurement data. To solve this problem, the method of least squares was invented by Gauss in 1795.

The least squares method is discussed in detail in Whittaker and Robinson (1924) and Tarantola (1987), but it can be simply illustrated as follows. Suppose there exist N observations s_1, s_2, \dots, s_N of the variable s . Assume that these observations were taken with different types of instruments and that the observation error associated with each measurement is given by $\varepsilon_n = s_n - s$. Assume that the observation errors are *random, unbiased, and normally distributed* [Equation (D6) of Appendix D]. Then the probability that the error of the n th observation lies between ε_n and $\varepsilon_n + d\varepsilon_n$ is

$$p(\varepsilon_n) = \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left[-\frac{\varepsilon_n^2}{2\sigma_n^2}\right] \quad (2.2.1)$$

where

$$\sigma_n^2 = \langle(s_n - s)^2\rangle = \langle\varepsilon_n^2\rangle = \int_{-\infty}^{\infty} \varepsilon_n^2 p(\varepsilon_n) d\varepsilon_n, \quad \langle\varepsilon_n\rangle = 0$$

and $\langle \rangle$ is the expectation operator defined in Appendix D.

Suppose there is only one observation, $N = 1$. Then the most probable value of s is that value for which the probability $p(\varepsilon_1)$ is a maximum. From (2.2.1), this occurs at $\varepsilon_1 = 0$, or $s = s_1$. Now, when there are N observations, the joint probability that ε_1 lies between ε_1 and $\varepsilon_1 + d\varepsilon_1$, ε_2 lies between ε_2 and $\varepsilon_2 + d\varepsilon_2$, \dots , ε_N lies between ε_N and $\varepsilon_N + d\varepsilon_N$ is the product of all the individual probabilities (2.2.1). Thus,

$$p(\varepsilon_1, \dots, \varepsilon_N) = p(\varepsilon_1)p(\varepsilon_2)\cdots p(\varepsilon_N) \\ = \prod_{n=1}^N \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left(-\frac{\varepsilon_n^2}{2\sigma_n^2}\right) = \left[\prod_{n=1}^N \frac{1}{\sigma_n \sqrt{2\pi}} \right] \exp\left[-\sum_{n=1}^N \frac{(s_n - s)^2}{2\sigma_n^2}\right]$$

where \prod is the product operator. In this case, the most probable value of s is that value for which the probability $p(\varepsilon_1, \dots, \varepsilon_N)$ is a maximum. This obviously occurs when the summation inside the preceding exponential is a minimum. The most probable value, denoted s_a , is often called the *maximum likelihood estimate* of s , and it must minimize

$$I = \frac{1}{2} \sum_{n=1}^N \sigma_n^{-2} (s_a - s_n)^2 = \frac{(s_a - s_1)^2}{2\sigma_1^2} + \frac{(s_a - s_2)^2}{2\sigma_2^2} + \cdots + \frac{(s_a - s_N)^2}{2\sigma_N^2} \quad (2.2.2)$$

Equation (2.2.2) can be minimized by differentiating with respect to s_a and setting the result to zero,

$$\sum_{n=1}^N \sigma_n^{-2} (s_a - s_n) = \frac{s_a - s_1}{\sigma_1^2} + \frac{s_a - s_2}{\sigma_2^2} + \cdots + \frac{s_a - s_N}{\sigma_N^2} = 0$$

or

$$s_a = \frac{\sum_{n=1}^N \sigma_n^{-2} s_n}{\sum_{n=1}^N \sigma_n^{-2}} \quad (2.2.3)$$

In other words, the most probable value (or maximum likelihood estimate) of s is given by a weighted average of the observations, where the weights are inversely proportional to the expected observation error variances of each observation. Note that (2.2.3) is a linear estimate and that it is also unbiased as $\langle s_n \rangle = \langle s \rangle$, $1 \leq n \leq N$.

Define the error of the estimate as $\varepsilon_a = s_a - s$. Then the expected error variance of the estimate s_a is

$$\langle \varepsilon_a^2 \rangle = \left\langle \left[\frac{\sum_{n=1}^N \sigma_n^{-2} (s_n - s)}{\sum_{n=1}^N \sigma_n^{-2}} \right]^2 \right\rangle = \left[\sum_{n=1}^N \sigma_n^{-2} \right]^{-1} \quad (2.2.4)$$

as $\langle \varepsilon_n \varepsilon_\eta \rangle = 0$ for $n \neq \eta$, because ε_n is random.

In the case in which all observations were taken with the same type of instrument, $\sigma_n^2 = \sigma^2$, $1 \leq n \leq N$,

$$s_a = \frac{1}{N} \sum_{n=1}^N s_n \quad \text{and} \quad \langle \varepsilon_a^2 \rangle = \frac{\sigma^2}{N} \quad (2.2.5)$$

Here the estimate of s is the simple average of the observations, and the expected error variance is directly proportional to the observed error variance and inversely proportional to the number of observations.

The estimate s_a is unbiased. When the observations are biased, $\langle \varepsilon_n \rangle \neq 0$, and (2.2.3) implies s_a is biased. A minor modification of (2.2.2) produces an unbiased estimate. Presuming that the instrument biases $\langle \varepsilon_n \rangle$ are known, replace s_n in (2.2.2) by $s_n - \langle \varepsilon_n \rangle$ and redefine σ_n^2 as $\langle \varepsilon_n^2 \rangle - \langle \varepsilon_n \rangle^2$. Minimization with respect to s_a yields

$$s_a = \frac{\sum_{n=1}^N \sigma_n^{-2} (s_n - \langle \varepsilon_n \rangle)}{\sum_{n=1}^N \sigma_n^{-2}}$$

Subtraction of the true value s from both sides and application of the expectation operator gives $\langle \varepsilon_a \rangle = 0$. In most of the discussion of the next four chapters, it will be assumed that the observations are unbiased, because if they are not and the bias is known, then its removal is straightforward.

Equation (2.2.2) can be rewritten as

$$I = \sum_{n=1}^N w_n d_n^2 \quad (2.2.6)$$

where $w_n = 0.5\sigma_n^{-2}$ and $d_n = s - s_n$. The w_n is the Gauss precision modulus or weight. Since w_n is specified, it will be called the *a priori weight* and denoted with lowercase letters to distinguish it from the *a posteriori* weights of (1.7.1). The d_n is the *residual* of the n th observation. It should be noted that the least squares estimate obtained by minimizing (2.2.6) is a maximum likelihood estimate only if the observation errors are normally distributed and $w_n = 0.5\sigma_n^{-2}$.

The special case $N = 2$ arises repeatedly in this book. For future reference, we now rewrite this case using a special notation. Denote the two observations of s as s_a and s_b with corresponding error variances σ_a^2 and σ_b^2 , respectively. Then (2.2.2-4) can be rewritten as

$$\begin{aligned} I &= \frac{(s_a - s_o)^2}{2\sigma_o^2} + \frac{(s_b - s_o)^2}{2\sigma_b^2} \\ s_a &= \frac{\sigma_o^{-2}s_o + \sigma_b^{-2}s_b}{\sigma_o^{-2} + \sigma_b^{-2}} = \frac{\sigma_b^2s_o + \sigma_o^2s_b}{\sigma_o^2 + \sigma_b^2} = s_o + \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2}[s_b - s_o] \\ \langle e_a^2 \rangle &= \sigma_b^2 - \frac{\sigma_b^4}{\sigma_o^2 + \sigma_b^2} = \frac{\sigma_b^2\sigma_o^2}{\sigma_o^2 + \sigma_b^2} = (\sigma_o^{-2} + \sigma_b^{-2})^{-1} \end{aligned} \quad (2.2.7)$$

The simple example (2.2.2-7) defined a *scalar* or zero-dimensional application of the principle of least squares. When the variable has spatial and/or temporal dependence, then a *vector* version of the principle of least squares can be applied. Spatial dependence is considered here and temporal dependence in Sections 8.6 and 13.2.

Consider first the case $N = 1$. At a fixed time, define a dependent or state variable $f(\mathbf{r})$, where $\mathbf{r} = (x, y, z)$ is a three-dimensional spatial coordinate. Define $f_O(\mathbf{r}_k)$ to be an observation of f at the observation station \mathbf{r}_k , with expected observation error variance $\langle e_O^2(\mathbf{r}_k) \rangle$. Suppose there are K such observation stations and the observation errors are normally distributed, unbiased, and spatially uncorrelated $\langle e_O(\mathbf{r}_k)e_O(\mathbf{r}_l) \rangle = 0$, $k \neq l$. Designate the analyzed field of f as $f_A(\mathbf{r})$.

In this case, there is only a single observation $f_O(\mathbf{r}_k)$ at each observation station \mathbf{r}_k . Introduce the quadratic form

$$I = \sum_{k=1}^K w_k d_k^2 = \frac{1}{2} \sum_{k=1}^K \langle e_O^2(\mathbf{r}_k) \rangle^{-1} [f_O(\mathbf{r}_k) - f_A(\mathbf{r}_k)]^2 \quad (2.2.8)$$

Minimization of (2.2.8) with respect to each of the unknown analysis values $f_A(\mathbf{r}_k)$ leads to the trivial solution

$$f_A(\mathbf{r}_k) = f_O(\mathbf{r}_k), \quad 1 \leq k \leq K$$

Because there is only a single observation at each observation state, the most probable value of f at \mathbf{r}_k is equal to that observation, consistent with (2.2.3). In the case of perfect observations (no observation error), this is the correct solution. Usually, quadratic forms such as (2.2.8) are minimized subject to some explicit or implicit constraint, and the solution is not trivial. In Section 2.1, for example, the objective

analysis was constrained to have the polynomial representation (2.1.1). In Sections 2.6 and 2.7, other types of constraints are introduced.

The quadratic form (2.2.8) can be written in matrix form as

$$I = [\underline{f}_A - \underline{f}_O]^T \underline{w} [\underline{f}_A - \underline{f}_O]$$

or

$$I = 0.5[\underline{f}_A - \underline{f}_O]^T \underline{Q}^{-1} [\underline{f}_A - \underline{f}_O] \quad (2.2.9)$$

where \underline{f}_A and \underline{f}_O are column vectors of length K of the analyzed values $f_A(\mathbf{r}_k)$ and observed values $f_O(\mathbf{r}_k)$, respectively. The \underline{w} is the $K \times K$ diagonal matrix with elements w_k , T indicates matrix transpose, and \underline{Q} is the diagonal $K \times K$ matrix with elements $\langle \varepsilon_O^2(\mathbf{r}_k) \rangle$. The convention of double underlining for matrices and single underlining for column vectors is used throughout this text.

If the observation errors are spatially correlated $\langle \varepsilon_O(\mathbf{r}_k) \varepsilon_O(\mathbf{r}_l) \rangle \neq 0$, then \underline{Q} is a full covariance matrix (see Appendix D). The form of (2.2.9) does not change in this case, but clearly \underline{Q} must be nonsingular.

Equation (2.2.9) is the vector equivalent of (2.2.2) for the case $N = 1$. For future comparison with results in Chapters 4 and 5, the vector case for $N = 2$ is also derived here. That is, Equation (2.2.9) is generalized to include background estimates of the function f at the observing stations. Thus, suppose there exist background estimates $f_B(\mathbf{r}_k)$, $1 \leq k \leq K$. Here the subscript B indicates background. With each background estimate is an associated background error $\varepsilon_B(\mathbf{r}_k)$. Assume that background and observation errors are unbiased, random, normally distributed, and spatially correlated, but not with each other:

$$\begin{aligned} \langle \varepsilon_O(\mathbf{r}_k) \varepsilon_O(\mathbf{r}_l) \rangle &\neq 0, & \langle \varepsilon_B(\mathbf{r}_k) \varepsilon_B(\mathbf{r}_l) \rangle &\neq 0 \\ \langle \varepsilon_O(\mathbf{r}_k) \varepsilon_B(\mathbf{r}_l) \rangle &= 0, & \text{for all } k, l \end{aligned} \quad (2.2.10)$$

Then the vector generalization of (2.2.2) for $N = 2$ is

$$I = 0.5\{\underline{f}_A - \underline{f}_O\}^T \underline{Q}^{-1} \{\underline{f}_A - \underline{f}_O\} + \{\underline{f}_A - \underline{f}_B\}^T \underline{B}^{-1} \{\underline{f}_A - \underline{f}_B\} \quad (2.2.11)$$

where \underline{B} is the background error covariance matrix with elements $\langle \varepsilon_B(\mathbf{r}_k) \varepsilon_B(\mathbf{r}_l) \rangle$. \underline{f}_A , \underline{f}_O , and \underline{f}_B are column vectors of the analyzed, observed, and background values at the observation stations. The properties of the error covariance matrices \underline{B} and \underline{Q} are discussed in Appendix D and Chapter 4, but for now it is enough to note that they must both be nonsingular. By analogy with (2.2.8), Equation (2.2.11) can be written as

$$\begin{aligned} I = 0.5 \sum_{k=1}^K \sum_{l=1}^K & \{ [f_A(\mathbf{r}_k) - f_O(\mathbf{r}_k)] [f_A(\mathbf{r}_l) - f_O(\mathbf{r}_l)] \delta_{kl} \\ & + [f_A(\mathbf{r}_k) - f_B(\mathbf{r}_k)] [f_A(\mathbf{r}_l) - f_B(\mathbf{r}_l)] \tilde{\delta}_{kl} \} \end{aligned} \quad (2.2.12)$$

where δ_{kl} and $\tilde{\delta}_{kl}$ are elements of the matrices \underline{Q}^{-1} and \underline{B}^{-1} , respectively. If the background and observation errors were spatially uncorrelated, then \underline{B} and \underline{Q} would

be diagonal and only the terms for $l = k$ would be retained in (2.2.12). Differentiating (2.2.12) with respect to each of the $f_A(r_k)$, $1 \leq k \leq K$, and setting the result to zero gives

$$\frac{\partial I}{\partial f_A(r_k)} = 0 = \sum_{l=1}^K \{ [f_A(r_l) - f_O(r_l)] \tilde{\alpha}_{kl} + [f_A(r_l) - f_B(r_l)] \tilde{\beta}_{kl} \}, \quad 1 \leq k \leq K$$

or

$$\underline{B}^{-1}[f_A - f_B] + Q^{-1}[f_A - f_O] = 0 \quad (2.2.13)$$

Rearrangement of (2.2.13) gives

$$f_A = [\underline{B}^{-1} + Q^{-1}]^{-1}[\underline{B}^{-1}f_B + Q^{-1}f_O] \quad (2.2.14)$$

$$f_A - f_B = \underline{B}[\underline{B} + Q]^{-1}[f_O - f_B] \quad (2.2.15)$$

Equation (2.2.14) is a generalization of (2.2.3) for the case $N = 2$, and thus f_A is the column vector of maximum likelihood estimates of f at the observation stations. Following the terminology of Chapters 1, $f_A - f_B$ is the column vector of analysis increments (at the observation stations), and $f_O - f_B$ is the column vector of observation increments. Note that (2.2.15) is not, in itself, an objective analysis algorithm because it provides no mechanism for determining f_A at any location that is not an observation station.

It is also possible to determine the expected error variance of the analyzed values f_A . Define f_T to be the column vector of true values $f_T(r_k)$. Subtract f_T from both sides of (2.2.14),

$$f_A - f_T = [\underline{B}^{-1} + Q^{-1}]^{-1}[\underline{B}^{-1}(f_B - f_T) + Q^{-1}(f_O - f_T)] \quad (2.2.16)$$

Define $\varepsilon_A(r_k) = f_A(r_k) - f_T(r_k)$ to be the analysis error at observation station k , and ε_A to be the column vector with elements $\varepsilon_A(r_k)$ and ε_O with elements $\varepsilon_O(r_k)$. Then, (2.2.16) can be written as

$$\varepsilon_A = [\underline{B}^{-1} + Q^{-1}]^{-1}[\underline{B}^{-1}\varepsilon_B + Q^{-1}\varepsilon_O] \quad (2.2.17)$$

Right multiply both sides of (2.2.17) by ε_A^T (where superscript T stands for transpose) and take expectation values. As $\langle \varepsilon_B \varepsilon_O^T \rangle = \langle \varepsilon_O \varepsilon_B^T \rangle = 0$,

$$\begin{aligned} \langle \varepsilon_A \varepsilon_A^T \rangle &= [\underline{B}^{-1} + Q^{-1}]^{-1}[\underline{B}^{-1} \langle \varepsilon_B \varepsilon_B^T \rangle \underline{B}^{-1} \\ &\quad + Q^{-1} \langle \varepsilon_O \varepsilon_O^T \rangle Q^{-1}] [\underline{B}^{-1} + Q^{-1}]^{-1} \end{aligned}$$

But

$$\langle \varepsilon_B \varepsilon_B^T \rangle = \underline{B} \quad \text{and} \quad \langle \varepsilon_O \varepsilon_O^T \rangle = Q$$

so

$$\langle \varepsilon_A \varepsilon_A^T \rangle = [\underline{B}^{-1} + Q^{-1}]^{-1} \quad (2.2.18)$$

Define the analysis error covariance matrix A with elements $\langle \varepsilon_A(r_k) \varepsilon_A(r_l) \rangle$. Then

$$A = [\underline{B}^{-1} + Q^{-1}]^{-1}$$

or

$$\underline{A} = \underline{B}[\underline{B} + \underline{Q}]^{-1}\underline{Q} = \underline{B} - \underline{B}[\underline{B} + \underline{Q}]^{-1}\underline{B} \quad (2.2.19)$$

Equations (2.2.11, 2.2.15, and 2.2.19) are the vector equivalents of the scalar equations (2.2.7), and there is an obvious similarity in the functional form, with constants being replaced by matrices. The elements along the main diagonal of \underline{B} and \underline{Q} are the expected background and observation error variances at the observation stations. In the same way, the main diagonal elements of \underline{A} are the expected analysis error variances at the observation stations. Equations (2.2.15 and 2.2.19) can be derived more elegantly using the vector methods of Section 5.1.

Minimization with respect to quadratic forms such as (2.2.2) or (2.2.11) is referred to mathematically as minimization in the l_2 -norm sense, with the subscript "2" referring to the power 2 in (2.2.2). Minimization in the l_2 sense leads to linear analysis equations (Section 1.7). However, it should be mentioned that minimization with respect to other norms is also possible. Thus an l_q -norm form corresponding to (2.2.2) might be,

$$I_q = \frac{1}{q} \sum_{n=1}^N \frac{|s_a - s_n|^q}{\sigma_n^q} \quad (2.2.20)$$

where $| |$ indicates absolute value. Tarantola (1987) discusses minimization using l_q -norm criteria, for $1 \leq q \leq \infty$. Minimization of (2.2.20) would be appropriate if the error probability distribution were not normal (2.2.1) but of the form

$$p(\varepsilon_n) = c_q \exp\left[\frac{-|\varepsilon_n|^q}{q\sigma_n^q} \right] \quad (2.2.21)$$

where c_q is a constant determined from the requirement that the integral of $p(\varepsilon_n)$ between $\pm\infty$ must equal one. The cases $q = 1$ and $q = \infty$ are particularly important. When $q = 1$, the error distribution (2.2.21) is "long-tailed" in that there is a much higher probability of large errors than in (2.2.1). Results obtained using a minimum l_1 -norm criterion are less sensitive to large errors than in the l_2 -norm case. This makes the l_1 -norm attractive for the quality control of data in a number of disciplines (Barrodale 1968).

Claerbout and Muir (1973) compared the l_1 - and l_2 -norms using the following analogy. "When a traveller reaches a fork in the road, the l_1 -norm tells him to take one way or the other, but the l_2 -norm instructs him to head off into the bushes. Likewise, a hunter when seeing two birds in the sky, might not choose to shoot at the mid-point between them." In other words, for the case $N = 2$, discussed in (2.2.7), the value s_a produced by an l_1 -norm minimization would usually be close to either s_a or s_b , whereas with an l_2 -norm approach it might be close to neither.

The l_2 -norm procedures are emphasized in this book because they lead to simple linear forms that are relatively easy to implement and because the normal error distribution (2.2.1) is thought to be appropriate for many meteorological variables (provided grossly erroneous observations have been rejected by quality-control

mechanisms). This background material on the theory of least squares estimation will be useful in Chapters 4 and 5, but a more immediate application is in developing function-fitting algorithms.

2.3 The Gram matrix

Section 2.1 gave a simple demonstration of function fitting. In that illustration, the function (2.1.1) was fitted to all the observations in a region of influence surrounding a given analysis gridpoint, and the functional fit was evaluated only at the gridpoint itself (i.e., at the centre of the data cluster). This type of fitting is referred to as *local fitting*.

In another type of function fitting, called *global fitting*, a single function is fitted to *all observations in the analysis domain*, and the analysis anywhere in the domain can be obtained by evaluating this function. A particularly important example of global fitting (see Section 2.6) occurs when the analysis domain is the entire earth's atmosphere.

The least squares minimization theory of the previous section is now used to explore function-fitting algorithms in more detail. Suppose we want to analyze a dependent variable $f(\mathbf{r})$ where \mathbf{r} is a one-, two- or three-dimensional spatial coordinate. Assume that the analyzed field $f_A(\mathbf{r})$ can be represented by a finite series of ordered basis functions $h_0(\mathbf{r}), \dots, h_m(\mathbf{r}), \dots, h_M(\mathbf{r})$, with (as yet unknown) expansion coefficients c_m , $0 \leq m \leq M$,

$$f_A(\mathbf{r}) = \sum_{m=0}^M c_m h_m(\mathbf{r}) \quad (2.3.1)$$

The notation adopted in (2.3.1) is purposely compact and must be used with caution if the expansion is two or three dimensional. In fact, there are two summation indices in two dimensions (as in Section 2.1) and three in three dimensions. The index m in (2.3.1) is assumed to be a multidimensional index that combines the individual indices for each dimension. For example, in Section 2.1, the six expansion coefficients $c_{00}, c_{10}, c_{01}, c_{20}, c_{02}$, and c_{11} would be relabeled c_0, c_1, c_2, c_3, c_4 , and c_5 . The corresponding basis functions would be denoted $h_0 = 1, h_1 = x, h_2 = y, h_3 = x^2, h_4 = y^2$, and $h_5 = xy$.

Assume that there are K observations $f_O(\mathbf{r}_k)$ over the domain and that we want a global fit to these observations. This can be achieved by minimizing the following quadratic form:

$$I = \sum_{k=1}^K w_k d_k^2 = \sum_{k=1}^K w_k \left[\sum_{m=0}^M c_m h_m(\mathbf{r}_k) - f_O(\mathbf{r}_k) \right]^2 \quad (2.3.2)$$

where $w_k = 0.5 \langle \varepsilon_O^2(\mathbf{r}_k) \rangle^{-1}$ and the observation errors are random, normal, unbiased, and spatially uncorrelated.

Equation (2.3.2) can be minimized by differentiating with respect to each coefficient

c_m and setting $\partial I / \partial c_m = 0$ for each m . This gives

$$\sum_{\mu=0}^M c_{\mu} \sum_{k=1}^K w_k h_m(\mathbf{r}_k) h_{\mu}(\mathbf{r}_k) = \sum_{k=1}^K w_k f_O(\mathbf{r}_k) h_m(\mathbf{r}_k), \quad 0 \leq m \leq M, \quad (2.3.3)$$

where μ is a dummy index as before. Equations (2.3.3) are the *normal equations*.

Equation (2.3.1) evaluated at the observation stations \mathbf{r}_k can be written in matrix form:

$$\mathbf{f}_A = \underline{\mathbf{H}} \mathbf{c} \quad (2.3.4)$$

where \mathbf{f}_A is the column vector of length K whose elements are $f_A(\mathbf{r}_k)$; \mathbf{c} is the column vector of length $M + 1$ of expansion coefficients c_m ; $\underline{\mathbf{H}}$ is the $K \times (M + 1)$ rectangular matrix whose elements $h_{km} = h_m(\mathbf{r}_k)$. Then the normal equations (2.3.3) can be written as

$$\underline{\mathbf{G}} \mathbf{c} = \underline{\mathbf{H}}^T \underline{\mathbf{Q}}^{-1} \underline{\mathbf{H}} \mathbf{c} = \underline{\mathbf{H}}^T \underline{\mathbf{Q}}^{-1} \mathbf{f}_O \quad (2.3.5)$$

where $\underline{\mathbf{Q}}$ is the diagonal $K \times K$ square matrix whose elements are $\langle \varepsilon_O^2(\mathbf{r}_k) \rangle$ and \mathbf{f}_O is the column vector of length K whose elements are $f_O(\mathbf{r}_k)$. The matrix $\underline{\mathbf{G}} = \underline{\mathbf{H}}^T \underline{\mathbf{Q}}^{-1} \underline{\mathbf{H}}$ is a $(M + 1) \times (M + 1)$ square matrix called the *Gram matrix*. The matrix on the left-hand side of (2.1.4) is such a matrix. $\underline{\mathbf{G}}$ is a square matrix (unlike $\underline{\mathbf{H}}$) and is, in principle, invertible:

$$\mathbf{c} = (\underline{\mathbf{H}}^T \underline{\mathbf{Q}}^{-1} \underline{\mathbf{H}})^{-1} \underline{\mathbf{H}}^T \underline{\mathbf{Q}}^{-1} \mathbf{f}_O \quad (2.3.6)$$

The analyzed field $f_O(\mathbf{r})$ can then be obtained at arbitrary locations in the domain by inserting (2.3.6) into (2.3.1). In global fitting, unlike local fitting, all the coefficients, c_m , $0 \leq m \leq M$ are required.

Because there are K observations, up to K coefficients can be determined from (2.3.6). In other words, $M \times 1 \leq K$. In the special case $K = M + 1$, $\underline{\mathbf{H}}$ is a square matrix, and the normal equations (2.3.3) are said to be *fully determined*. In general, $M + 1 \leq K$ and the normal equations are said to be *overdetermined*; that is, there are more independent observations than there are coefficients to be determined.

In the fully determined case, (2.3.5) becomes $\underline{\mathbf{H}} \mathbf{c} = \mathbf{f}_O$, and the specified weights are irrelevant. We illustrate this case using one-dimensional polynomial basis functions:

$$h_0(x) = 1, h_1(x) = x, \dots, h_M(x) = x^M \quad (2.3.7)$$

Then, for the observation points x_k , $1 \leq k \leq K$, the matrix $\underline{\mathbf{H}}$ becomes

$$\underline{\mathbf{H}} = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^M \\ \vdots & & & & \\ 1 & x_k & x_k^2 & \cdots & x_k^M \end{pmatrix} \quad (2.3.8)$$

In the special case $K = M + 1 = 2$, (2.3.6) becomes

$$\begin{vmatrix} c_0 \\ c_1 \end{vmatrix} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \end{pmatrix}^{-1} \begin{vmatrix} f_O(x_1) \\ f_O(x_2) \end{vmatrix} \quad (2.3.9)$$

or

$$(2.3.3) \quad c_0 = \frac{x_2 f_O(x_1) - x_1 f_O(x_2)}{x_2 - x_1}, \quad c_1 = \frac{f_O(x_2) - f_O(x_1)}{x_2 - x_1}$$

ns.
matrix

2.3.4)
lumn
gular
ritten

2.3.5)

f_O is
 H is
and
iple,

3.6)

main
nts,

om
are
ral,
are

hts
sis

.7)

8)

The analyzed field $f(x)$ from (2.3.4) becomes

$$(2.3.10) \quad f_A(x) = \frac{x_2 - x}{x_2 - x_1} f_O(x_1) + \frac{x - x_1}{x_2 - x_1} f_O(x_2)$$

which will be recognized as the Lagrange interpolation formula [Equation (E1) in Appendix E] for $K = 2$. In this fully determined case, the analysis $f_A(x)$ is required to exactly fit the observation stations x_k , $1 \leq k \leq K$. In fact, the residuals $d_k = f_A(x_k) - f_O(x_k)$ are identically equal to zero in this case. Thus, fully determined least squares minimization using polynomial basis functions is equivalent to Lagrange interpolation. The choice $M + 1 = K$ is appropriate when the observations are error free. Usually, the observations $f_O(x_k)$ are not error free, and a fully determined fit is not only inappropriate, it can lead to serious problems, as we demonstrate in the next section.

The Gram matrix G has special properties. Usually, G is a full matrix and must be inverted using the standard techniques of linear algebra. In the case for which the observation errors are spatially uncorrelated, the elements $g_{m\mu}$ of the Gram matrix may be written as

$$(2.3.11) \quad g_{m\mu} = \sum_{k=1}^K w_k h_m(\mathbf{r}_k) h_\mu(\mathbf{r}_k)$$

From (2.3.11), it is apparent that G is real and symmetric. An arbitrary $N \times N$ matrix Q is said to be positive semidefinite if (see Faddeev and Faddeeva 1963), for any arbitrary column vector \underline{z} with elements z_n , $1 \leq n \leq N$,

$$(2.3.12) \quad \underline{z}^T Q \underline{z} \geq 0 \quad (\text{provided not all } z_n \text{ are equal to zero})$$

When (2.3.12) is strictly positive, Q is said to be positive definite. It is straightforward to show that the Gram matrix is positive semidefinite:

$$(2.3.13) \quad \begin{aligned} \underline{z}^T G \underline{z} &= \sum_{m=0}^M \sum_{\mu=0}^M g_{m\mu} z_m z_\mu = \sum_{k=1}^K w_k \sum_{m=0}^M \sum_{\mu=0}^M h_m(\mathbf{r}_k) h_\mu(\mathbf{r}_k) z_m z_\mu \\ &= \frac{1}{2} \sum_{k=1}^K \langle \varepsilon_O^2(\mathbf{r}_k) \rangle^{-1} \left[\sum_{m=0}^M h_m(\mathbf{r}_k) z_m \right]^2 \geq 0 \end{aligned}$$

Consider now the eigenstructure of G . If \underline{e} is an eigenvector of G , then

$$(2.3.14) \quad G \underline{e} = \lambda \underline{e}$$

where λ is the eigenvalue. Two eigenvectors \underline{e}_m and \underline{e}_μ of a matrix are said to be orthogonal if $\underline{e}_m^T \underline{e}_\mu = \delta_{m\mu}$ and $\delta_{m\mu}$ is the Kronecker delta function (Here we have assumed that the eigenvectors have been normalized so that $\underline{e}_m^T \underline{e}_m = 1$). Because G is real and symmetric, its eigenvectors are orthogonal and its eigenvalues are real. If λ

is an eigenvalue of the positive semidefinite matrix \underline{G} and \underline{e} is the corresponding eigenvector, then

$$\frac{\underline{e}^T \underline{G} \underline{e}}{\underline{e}^T \underline{e}} = \lambda \geq 0 \quad (2.3.15)$$

because both numerator and denominator are positive. Thus, the eigenvalues of \underline{G} are real and nonnegative (positive if \underline{G} is positive definite).

The Gram matrix can always be inverted if all of its eigenvalues are positive. However, if one or more of its eigenvalues is vanishingly small, it can be numerically singular. Under what circumstances does this happen? If two of the basis functions $h_m(\mathbf{r})$, $h_\mu(\mathbf{r})$ are similar but not identical, then two rows of the Gram matrix will become similar, and the matrix can be numerically singular. This can also happen if one of the rows can be approximately represented by a linear combination of the other rows.

A matrix that is difficult to invert is called ill-conditioned, and the matrix inverse is not very accurate. A measure of the accuracy of the matrix inverse is given by the condition number defined by

$$N_c = \frac{\lambda_1}{\lambda_s} \quad (2.3.16)$$

where λ_1 is the largest eigenvalue of \underline{G} , and λ_s is the smallest. A high condition number implies a λ_s that is very close to zero and near linear dependency of the rows (and columns) of the matrix. As shown by Dixon et al. (1972), the relative error in the coefficients \underline{c} of (2.3.6) is directly proportional to the condition number of the Gram matrix \underline{G} . Thus, if \underline{G}^{-1} is not very accurate, then the resulting objective analyses are not very satisfactory. Unfortunately, the matrices encountered in function fitting often have high condition numbers. For the polynomial basis functions (2.3.7), the condition number is high except when M is small. As discussed in Dorny (1975), $M = 7$ is about the practical limit for polynomial fitting.

Even when \underline{G} is well-conditioned, it may be difficult to invert in practice. The order of the Gram matrix is equal to the number of degrees of freedom in the expansion. There is a practical limit to the sizes of full matrices that can be inverted because of the finite memory size of the computer and because the computation time for matrix inversion increases as the cube of the order. This is not a problem in local fitting, where M is generally small, but it is a serious limitation in global fitting.

In principle, both of these problems can be circumvented by a Gram-Schmidt orthogonalization process. If the elements $g_{m\mu}$, $m \neq \mu$, defined in (2.3.11) can be made equal to zero, then the Gram matrix is diagonal and trivial to invert for any value of M . The off-diagonal elements of \underline{G} vanish if the basis functions are orthogonal to each other over the observation network \mathbf{r}_k , $1 \leq k \leq K$, with respect to the prescribed weights w_k . Orthogonal basis functions $p_m(\mathbf{r})$ can be constructed from nonorthogonal

basis functions $h_m(\mathbf{r})$ as follows. Define an inner product

$$[p, q] = \sum_{k=1}^K w_k p(\mathbf{r}_k)q(\mathbf{r}_k) \quad (2.3.17)$$

Then (2.3.3) can be written as

$$\sum_{\mu=0}^M c_{\mu} [h_m, h_{\mu}] = [h_m, f_0] \quad (2.3.18)$$

and $[h_m, h_{\mu}]$ is not usually equal to zero when $m \neq \mu$. A new set of basis functions $p_m(\mathbf{r})$, $0 \leq m \leq M$, can be constructed:

$$\begin{aligned} p_0(\mathbf{r}) &= h_0(\mathbf{r}), \\ p_1(\mathbf{r}) &= h_1(\mathbf{r}) - \frac{[p_0, h_1]}{[p_0, p_0]} p_0(\mathbf{r}) \\ p_m(\mathbf{r}) &= h_m(\mathbf{r}) - \sum_{\mu=0}^{m-1} \frac{[p_{\mu}, h_m]}{[p_{\mu}, p_{\mu}]} p_{\mu}(\mathbf{r}) \end{aligned} \quad (2.3.19)$$

Equation (2.3.18) can be rewritten in terms of the new basis functions $p_m(\mathbf{r})$ with different expansion coefficients \tilde{c}_m . Thus,

$$\sum_{\mu=0}^M \tilde{c}_{\mu} [p_m, p_{\mu}] = [p_m, f_0] \quad \text{or} \quad \tilde{c}_m = [p_m, p_m]^{-1} [p_m, f_0] \quad (2.3.20)$$

from (2.3.19). The Gram-Schmidt process (2.3.19) for a discrete network is completely analogous to classical Gram-Schmidt orthogonalization for continuous variables.

The process (2.3.19) can be laborious if M is very large. Also, as pointed out by Dorni (1975), the $p_m(\mathbf{r})$ must be constructed very carefully if the $h_m(\mathbf{r})$ are almost linearly dependent. The Gram-Schmidt procedure is attractive for a time-invariant observation network \mathbf{r}_k that will be used many times. For then, the procedure (2.3.19) can be done once and for all because it does not depend on the observations, only on their locations and expected observation errors. The Gram-Schmidt procedure has been applied to polynomial fitting for meteorological spatial analysis by Dixon et al. (1972).

2.4 Underfitting, overfitting, and other problems

The previous section discussed two technical problems in the inversion of the Gram matrix: possible ill-conditioning (particularly for polynomial basis functions) and the order of the matrix being too large to invert (for global fitting). Unfortunately, there are other problems, and accurate inversion of the Gram matrix does not necessarily guarantee a satisfactory objective analysis. This section examines some of the remaining problems of function-fitting algorithms, with particular emphasis on underfitting and overfitting.

Problems of underfitting and overfitting are examined using trigonometric basis functions. The one-dimensional case will be considered; extension to two and three dimensions is straightforward. Consider the analysis domain $-\pi \leq x \leq \pi$, with periodic boundary conditions $f(-\pi) = f(\pi)$. Assume the analysis $f_A(x)$ is expanded in trigonometric basis functions as follows:

$$f_A(x) = \frac{a_0}{2} + \sum_{m=1}^M [a_m \cos mx + b_m \sin mx] \quad (2.4.1)$$

The subsequent developments are considerably simplified if complex Fourier notation is used. Define $c_m = 0.5(a_m - ib_m)$, where $i = \sqrt{-1}$. Because $f_A(x)$ is real, c_m is conjugate symmetric, $c_{-m} = c_m^*$, where (*) indicates complex conjugation. Then

$$f_A(x) = \sum_{m=-M}^M c_m e^{imx} \quad (2.4.2)$$

The spatial scale of the trigonometric basis functions decreases as m increases. Now consider the observations $f_O(x_k)$, $1 \leq k \leq K$, and define a quadratic form similar to (2.3.2):

$$I = \sum_{k=1}^K w_k d_k^2 = \sum_{k=1}^K w_k \left[\sum_{m=-M}^M c_m e^{imx_k} - f_O(x_k) \right]^2 \quad (2.4.3)$$

Minimizing (2.4.3) with respect to each of the $2M + 1$ coefficients c_m gives the normal equations,

$$\sum_{\mu=-M}^M c_{\mu} \sum_{k=1}^K w_k e^{i(\mu-m)x_k} = \sum_{k=1}^K w_k f_O(x_k) e^{-imx_k} \quad (2.4.4)$$

where μ is a dummy index as before and we have made use of the fact that c_m is conjugate symmetric. The Gram matrix in (2.4.4) can be thought of as a $(2M + 1) \times (2M + 1)$ real, symmetric, positive semidefinite matrix whose elements can be determined from the real form of the trigonometric expansion (2.4.1).

A close correspondence exists between (2.4.4) and a truncated Fourier expansion for the case in which the specified weight functions w_k are all equal. If the sums $\sum_{k=1}^K (\cdot)$ in (2.4.3–4) are replaced by integrals $\int_{-\pi}^{\pi} (\cdot) dx$, then (2.4.4) becomes

$$c_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_O(x) e^{-imx} dx \quad (2.4.5)$$

In deriving (2.4.5), we have used the identity

$$\int_{-\pi}^{\pi} e^{i(\mu-m)x} dx = 2\pi \delta_{m\mu} \quad (2.4.6)$$

Note that (2.4.5) is the classical expression for the coefficients of a Fourier series.

Figure 2.2 illustrates a least squares fit using trigonometric basis functions. Over

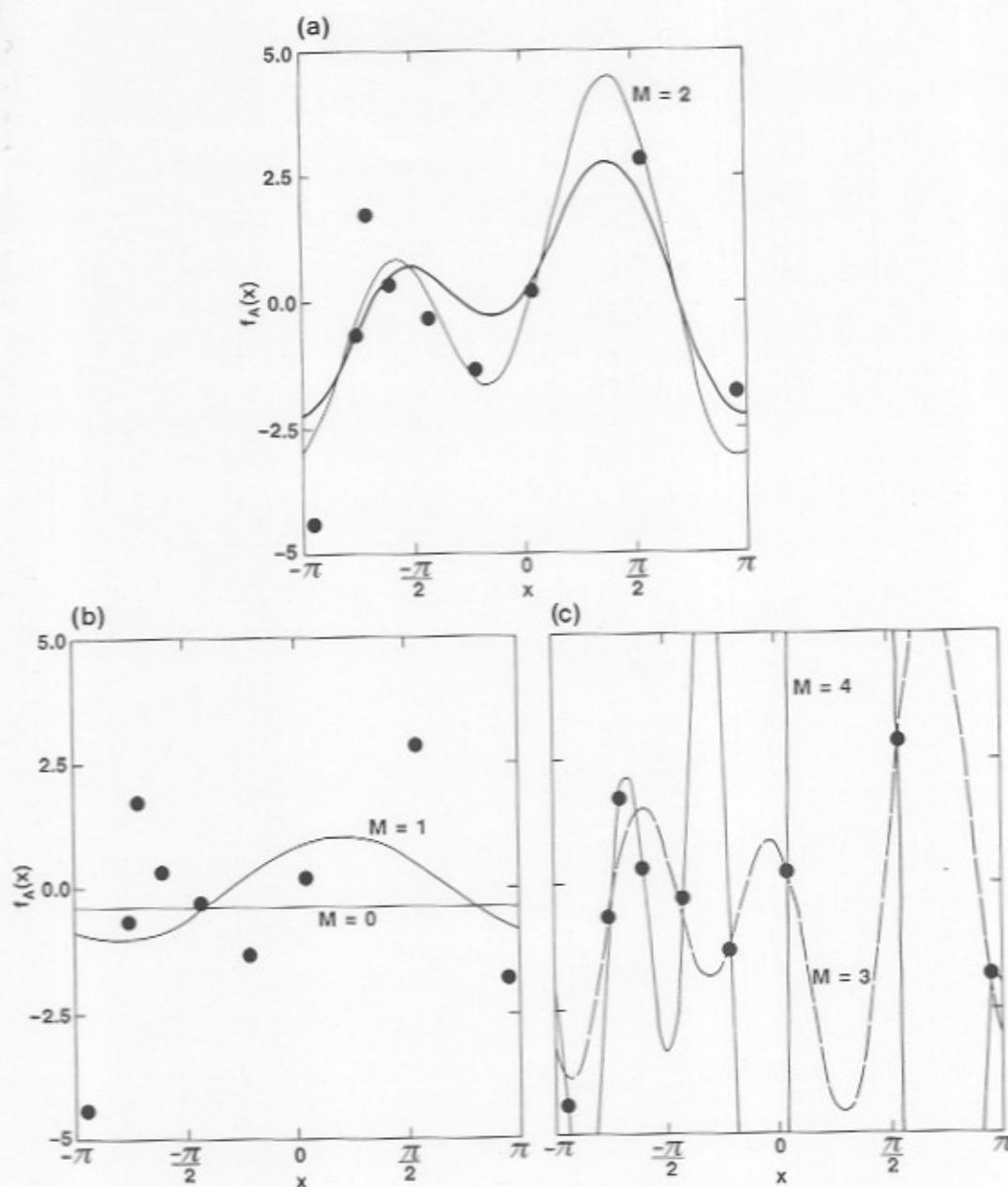


Figure 2.2 Underfitting and overfitting. The signal is shown by the heavy curve in (a); the observations (including error) are shown by solid dots. Fits for $M = 0, 1, 2, 3$, and 4 are shown in (a), (b), and (c).

the interval $-\pi \leq x \leq \pi$, the signal or truth is specified to be

$$f_T(x) = 0.5 + 1.2 \cos(x) + 0.8 \sin(x) - 1.3 \cos(2x) + 0.6 \sin(2x) \quad (2.4.7)$$

where subscript T stands for the truth. The signal is shown by the dark solid line in Figure 2.2(a).

All the observations $f_O(x_k)$ are assumed to have the same expected observation

error variance and were created by adding to $f_T(x_k)$ random errors drawn from a normal distribution (Appendix D) with mean zero and standard deviation equal to 1. There are nine observations, indicated by the dots in Figure 2.2(a). The observation location is shown by its position on the x axis (abscissa), and the values of the observation $f_O(x_k)$ by its value on the ordinate. Note that the observations are not distributed uniformly; they are relatively dense for $x < 0$ and relatively sparse for $x > 0$.

The signal (2.4.7) has 5 degrees of freedom $-2 \leq m \leq 2$. Figure 2.2(a) shows the analysis of nine observations using the fitting algorithm – [(2.4.4) and (2.4.2 with all w_k equal)] – for $M = 2$ (lightweight curve). From (a), we see that the analysis errors are largest when the actual observation error is large or when the observation density is low, in accord with intuition.

Now consider the case $M < 2$. That is, the correct basis functions are used for the analysis, but there are fewer degrees of freedom than in the signal. This is called *underfitting* and is illustrated in Figure 2.2(b). The signal is also given by (2.4.7), and the observation values and locations are indicated as in (a). Two cases of underfitting are shown: $M = 0$ and $M = 1$. The analysis for $M = 0$ is a constant and is a very poor fit to the signal. The case $M = 1$ (3 degrees of freedom) is a better fit but not as good as for $M = 2$.

Another case is $M > 2$, for which there are more degrees of freedom in the analysis than in the signal. This is called *overfitting* and is illustrated in Figure 2.2(c). Two cases are shown: $M = 3$ (7 degrees of freedom) and $M = 4$ (9 degrees of freedom). Note that for $M = 4$, there are exactly the same number of degrees of freedom in the expansion as there are observations, giving a fully determined fit. Every observation is fitted exactly for $M = 4$. The result is an extremely poor analysis in the regions with sparse observations. It is clear from (c) why an overdetermined fit is preferable to a fully determined fit.

In actual practice, of course, the signal is never known. However, the spectrum of the signal (Figure 1.4, for example) might be known, at least approximately. The optimum value of the truncation limit M would be chosen on the basis of knowledge of the spectrum and the characteristics of the observing system. There is usually a trade-off between choosing M too small and too large. If M is too small, the analysis will be smooth and will not fit the observations very well. If M is too large, the analysis will fit the observations well but is likely to be totally unacceptable in data voids. Overfitting is extremely dangerous.

Overfitting and underfitting can occur with any set of basis functions. Polynomial fitting has additional problems. All polynomials tend to $\pm\infty$ at large distances, and extrapolation is very risky, as noted in Appendix E. Local polynomial fitting is more reliable than global polynomial fitting, but there are difficulties when the data density is changing rapidly (at the edge of a data void).

Polynomial fitting has still another problem. Suppose one wants to calculate spatial derivatives of the analyzed field – for example, to calculate the geostrophic wind (7.4.5) from a geopotential analysis or the vorticity from a streamfunction analysis

(Section 6.4). First derivatives of the analyzed field are required in the first instance and second derivatives in the second instance.

The j th spatial derivative (in one dimension) of trigonometric and polynomial expansions is given, respectively, by

$$\frac{d^j f_A(x)}{dx^j} = \sum_{m=-M}^M c_m (im)^j e^{imx} \quad \text{and} \quad \frac{d^j f_A(x)}{dx^j} = \sum_{m=0}^{M-j} c_{m+j} \frac{(m+j)!}{m!} x^m \quad (2.4.8)$$

where (!) indicates factorial. The higher derivatives of an analysis based on a trigonometric expansion give more weight to the smaller-scale basis functions ($|m|$ large). Thus, the spatial derivatives of the analysis have more small-scale variance and are finer grained (less smooth) than the analysis itself. This is in accord with intuition and experience.

In the case of a polynomial expansion, the maximum order of the polynomial is reduced by spatial differentiation; and when $j > M$, the j th derivative of the analysis is reduced to zero. Thus, the derived fields obtained by differentiation of an analysis based on polynomial basis functions have fewer local maxima and minima than the analysis itself. Such derived fields lack detail and are, in general, smoother than the analyzed field. The derived fields, in this case, are unrealistic.

2.5 The a posteriori analysis weights

We can obtain considerable insight into any spatial analysis procedure by examining the a posteriori analysis weights defined in (1.7.1). Examination of the a posteriori weights is one way to compare the characteristics of various objective analysis algorithms. Another way to compare the spectral responses of different algorithms is discussed in Chapter 3.

An expression for the a posteriori weights (assuming that the observation errors are spatially uncorrelated) is obtained as follows. Define $\tilde{g}_{m\mu}$ to be an arbitrary element of \mathbf{G}^{-1} , the inverse of the Gram matrix. Then, from (2.3.3),

$$c_m = \sum_{\mu=0}^M \sum_{k=1}^K w_k \tilde{g}_{m\mu} h_\mu(\mathbf{r}_k) f_O(\mathbf{r}_k), \quad 0 \leq m \leq M \quad (2.5.1)$$

At an analysis point \mathbf{r}_i , the analyzed value is

$$f_A(\mathbf{r}_i) = \sum_{m=0}^M c_m h_m(\mathbf{r}_i) = \sum_{k=1}^K W_{ik} f_O(\mathbf{r}_k)$$

where

$$W_{ik} = w_k \sum_{m=0}^M \sum_{\mu=0}^M \tilde{g}_{m\mu} h_m(\mathbf{r}_i) h_\mu(\mathbf{r}_k) \quad (2.5.2)$$

is the a posteriori weight given to the observation $f_O(\mathbf{r}_k)$ in the analysis at \mathbf{r}_i .

Because the $\tilde{g}_{m\mu}$ are the elements of the inverse of the Gram matrix, which is

usually a full matrix, we cannot get a simple analytic expression for W_{ik} except in very special circumstances, such as the following. Define the equally spaced one-dimensional observation network over the analysis domain $-\pi \leq x \leq \pi$:

$$x_k = -\pi + k \Delta x \quad \text{where} \quad \Delta x = \frac{2\pi}{K}, \quad 1 \leq k \leq K \quad (2.5.3)$$

The mathematics is slightly simplified if K is chosen to be odd. Over this network, the normal equations for one-dimensional trigonometric basis functions (2.4.4), when all the specified weights w_k are equal, become

$$c_m = \frac{1}{2\pi} \sum_{k=1}^K f_O(x_k) \exp(-imx_k) \Delta x, \quad -M \leq m \leq M \quad (2.5.4)$$

The off-diagonal elements in (2.5.4) vanish because of the following identity (Jenkins and Watt 1968), which is valid provided that $(2m+1)$ and $(2\mu+1)$ are both less than or equal to K :

$$\sum_{k=1}^K \exp[i(\mu-m)(-\pi + k \Delta x)] \Delta x = 2\pi \delta_{m\mu} \quad (2.5.5)$$

Substitution of (2.5.4) into

$$f_A(x_i) = \sum_{m=-M}^M c_m e^{imx_i} = \sum_{k=1}^K W_{ik} f_O(x_k)$$

yields

$$W_{ik} = \frac{\Delta x}{2\pi} \sum_{m=-M}^M e^{im(x_i - x_k)} = \frac{\Delta x}{\pi} \left[0.5 + \sum_{m=1}^M \cos m(x_k - x_i) \right] \quad (2.5.6)$$

The a posteriori weights W_{ik} indicate the weight that is given to an observation displaced by a distance $x_k - x_i$ from the analysis gridpoint. A number of properties of W_{ik} can be obtained by examining (2.5.6). The a posteriori weight is symmetric in the univariate case, $W_{ik} = W_{ki}$. The application of (2.5.5) shows that the sum of the weights is equal to one, $\sum_{k=1}^K W_{ik} = 1$. It is also simple to show that $W_{ik} \leq W_{ii}$, where W_{ii} is the weight given to an observation if it occurs at the analysis gridpoint itself. In a fully determined fit, $W_{ik} = \delta_{ik}$ (see Exercise 2.7). For fixed M , $\lim_{K \rightarrow \infty} W_{ik} = 0$; so, as the number of observations increases (and the observation spacing Δx decreases), the weight given to an individual observation decreases.

$W_{ik}/\Delta x$ is purely a function of the displacement between observation stations and the analysis gridpoint and does not depend on K . From (2.5.6), $\lim_{M \rightarrow \infty} W_{ii}/\Delta x = \infty$. In Figure 2.3 we have plotted the function $W_{ik}/\Delta x$ (ordinate) for the equally spaced observation network using trigonometric basis functions (2.5.6). The abscissa is the distance between observation and analysis point, $x_k - x_i$, in the analysis domain $-\pi$ to π . Two cases are shown: $M = 2$ (5 degrees of freedom) and $M = 10$ (21 degrees of freedom).

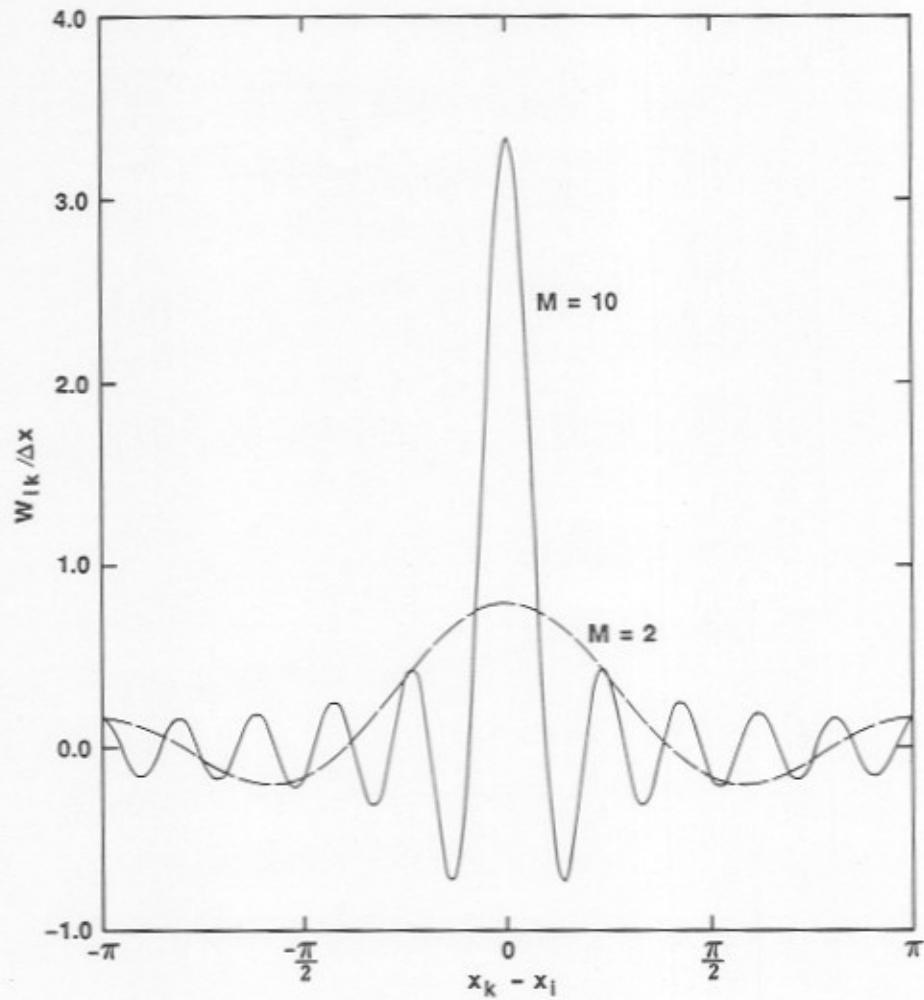


Figure 2.3 The a posteriori weight $W_{ik}/\Delta x$ as a function of $x_k - x_i$ for the equally spaced one-dimensional network in the univariate case.

In the case $M = 2$, relatively large weights are given to observations that lie considerable distances from the analysis gridpoint. For $M = 10$, the a posteriori weight tends to be small except in the immediate neighborhood of the analysis gridpoint. Specifying $M = 2$ produces a relatively large-scale analysis, which implies a relatively broad a posteriori weight function. Higher values of M imply increasingly localized a posteriori weight functions. Note that the a posteriori weight functions can be negative.

Now, the equally spaced case is rather special: A more realistic situation is shown in Figure 2.4. Again trigonometric basis functions (2.4.2) are used with $M = 4$ (9 degrees of freedom) over the analysis domain $-\pi$ to π . The observation network, this time, is irregular, and the 13 observation points x_k are indicated by large solid dots. It is evident that the observation network density is high for $x_k < 0$ and low for $x_k > 0$. The expected observation error variance $\langle \varepsilon_0^2(x_k) \rangle$ is assumed to be the

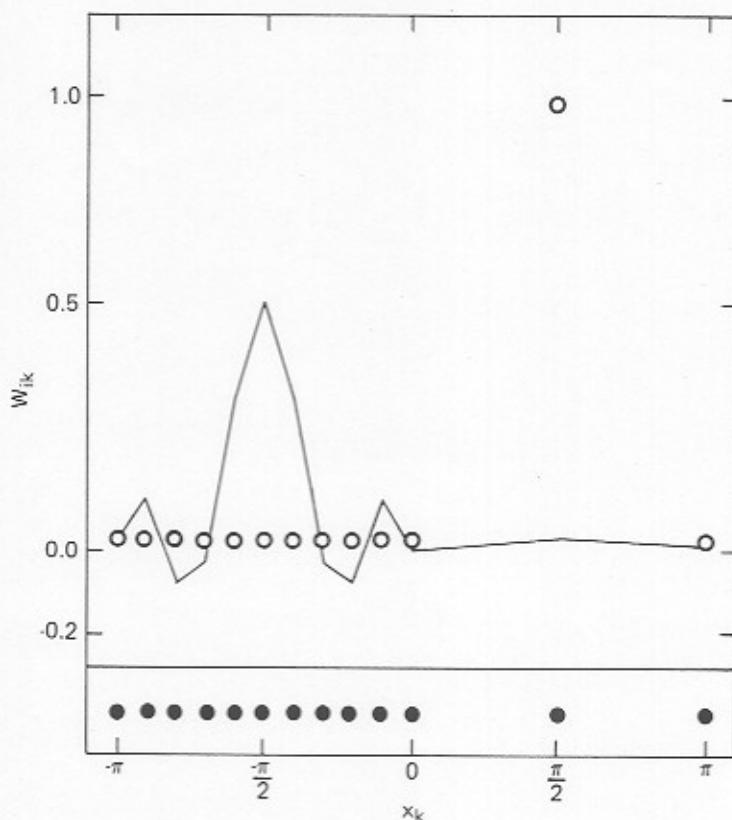


Figure 2.4 The a posteriori weight W_{ik} as a function of x_k for an irregular network with analysis gridpoints at $x_i = -\pi/2$ (solid curve) and $x_i = \pi/2$ (open circles). The observation locations are indicated by the solid dots.

same at all observation stations. The values of W_{ik} have been determined numerically for this case for two analysis points x_i . The first analysis point $x_i = -\pi/2$ happens to coincide with the sixth observation station from the left. The weight function W_{ik} , in this case, is shown by the continuous curve. The second analysis point chosen $x_i = \pi/2$ happens to coincide with the twelfth observation station from the left. The weights in this case are indicated by the open circles.

It can be seen that for the same value of M , the weight distribution differs markedly from location to location when the observation network is irregular. Where the network is relatively dense ($x_k < 0$), the a posteriori weight distribution is similar to that in Figure 2.3. On the other hand, where the observation density is low ($x_k > 0$), the analysis algorithm puts almost all the weight on the local observations.

2.6 Multivariate function fitting

One of the major themes of this book is the necessity for objective analyses to be consistent with the governing equations of the atmosphere. The dependent or state variables are coupled through the governing equations and therefore cannot be

analyzed in isolation. The equations that relate the dependent variables are called dynamic or physical constraints. (Mathematical constraints are introduced in the next section.) When several dependent variables are analyzed simultaneously, with implicit or explicit dynamic constraints, the process is referred to as multivariate objective analysis.

A series of increasingly complex dynamic constraints will be developed in Chapters 6–10; but for present purposes, we consider only very simple constraints. Simple, linear dynamic constraints can be accommodated easily into the function-fitting formalism. In fact, multivariate function fitting goes back to the time of Panofsky (1949).

The geostrophic relation is a simple linear constraint that is derived in most elementary meteorology textbooks (discussed in detail in Chapter 6). In cartesian coordinates on an *f*-plane, it is

$$u = -\frac{1}{f_0} \frac{\partial \Phi}{\partial y}, \quad v = \frac{1}{f_0} \frac{\partial \Phi}{\partial x} \quad (2.6.1)$$

where *u* and *v* are the eastward and northward velocity components, Φ the geopotential field, *x* and *y* are the eastward and northward spatial coordinates, and f_0 is the (constant) Coriolis parameter. Equation (2.6.1) is a low-order approximation to the governing equations with considerable validity in the extratropics, but it is completely inappropriate in the tropics.

Consider the simultaneous analysis of geopotential $\Phi_A(x, y)$ and horizontal wind components $u_A(x, y)$ and $v_A(x, y)$. Assume that the analyzed values are expanded in a truncated series of ordered two-dimensional basis functions:

$$\begin{aligned} \Phi_A(x, y) &= \sum_{mn} c_{mn} h_{mn}^\Phi(x, y) & u_A(x, y) &= \sum_{mn} c_{mn} h_{mn}^u(x, y) \\ v_A(x, y) &= \sum_{mn} c_{mn} h_{mn}^v(x, y) \end{aligned} \quad (2.6.2)$$

where the h_{mn}^Φ , h_{mn}^u , and h_{mn}^v are basis functions corresponding to Φ , *u*, and *v*, respectively. Note that the expansion coefficients in each of the three expansions are the same. The dynamic constraint (2.6.1) is imposed implicitly by demanding common expansion coefficients for each of the three variables and a relationship between the basis functions. For example, suppose the basis functions for the expansion Φ_A were two-dimensional trigonometric functions. Then the geostrophic constraint (2.6.1) would imply

$$h_{mn}^\Phi = e^{i(mx + ny)}, \quad h_{mn}^u = -\frac{in}{f_0} e^{i(mx + ny)}, \quad h_{mn}^v = \frac{im}{f_0} e^{i(mx + ny)} \quad (2.6.3)$$

Let us now introduce the more compact notation of (2.3.1). Define $\mathbf{r} = (x, y)$ to be the two-dimensional spatial coordinate, and write the analyzed values Φ_A , u_A , and

v_A in (2.6.2) as

$$\Phi_A(\mathbf{r}) = \sum_m c_m h_m^\Phi(\mathbf{r}), \quad u_A(\mathbf{r}) = \sum_m c_m h_m^u(\mathbf{r}), \quad v_A(\mathbf{r}) = \sum_m c_m h_m^v(\mathbf{r}) \quad (2.6.4)$$

where m is now a two-dimensional index. Suppose there are observations $\Phi_O(\mathbf{r}_k)$, $1 \leq k \leq K$, at the observation stations \mathbf{r}_k and $u_O(\mathbf{r}_l), v_O(\mathbf{r}_l)$, $1 \leq l \leq L$, at the observation stations \mathbf{r}_l . Define $\varepsilon_\Phi(\mathbf{r}_k)$ and $\varepsilon_v(\mathbf{r}_l)$ to be the observation errors of geopotential and wind components, respectively, and assume that

$$\langle \varepsilon_\Phi(\mathbf{r}_q) \varepsilon_\Phi(\mathbf{r}_p) \rangle = \langle \varepsilon_v(\mathbf{r}_q) \varepsilon_v(\mathbf{r}_p) \rangle = 0 \quad \text{for } p \neq q$$

and

$$\langle \varepsilon_\Phi(\mathbf{r}_p) \varepsilon_v(\mathbf{r}_q) \rangle = 0 \quad \text{for all } p, q$$

Thus, the observation error is spatially uncorrelated, and there is no correlation between observation errors of wind and geopotential. Then the following quadratic form can be defined as

$$I = \sum_{k=1}^K w_\Phi(\mathbf{r}_k) \left[\sum_m c_m h_m^\Phi(\mathbf{r}_k) - \Phi_O(\mathbf{r}_k) \right]^2 + \sum_{l=1}^L w_v(\mathbf{r}_l) \left\{ \left[\sum_m c_m h_m^u(\mathbf{r}_l) - u_O(\mathbf{r}_l) \right]^2 + \left[\sum_m c_m h_m^v(\mathbf{r}_l) - v_O(\mathbf{r}_l) \right]^2 \right\} \quad (2.6.5)$$

where $w_\Phi(\mathbf{r}_k) = 0.5 / \langle \varepsilon_\Phi^2(\mathbf{r}_k) \rangle$ and $w_v(\mathbf{r}_l) = 0.5 / \langle \varepsilon_v^2(\mathbf{r}_l) \rangle$.

Minimizing (2.6.5) with respect to each of the expansion coefficients c_m gives

$$\begin{aligned} \sum_\mu c_\mu & \left\{ \sum_{k=1}^K w_\Phi(\mathbf{r}_k) h_m^\Phi(\mathbf{r}_k) h_\mu^\Phi(\mathbf{r}_k) + \sum_{l=1}^L w_v(\mathbf{r}_l) [h_m^u(\mathbf{r}_l) h_\mu^u(\mathbf{r}_l) + h_m^v(\mathbf{r}_l) h_\mu^v(\mathbf{r}_l)] \right\} \\ & = \sum_{k=1}^K w_\Phi(\mathbf{r}_k) \Phi_O(\mathbf{r}_k) h_m^\Phi(\mathbf{r}_k) + \sum_{l=1}^L w_v(\mathbf{r}_l) [u_O(\mathbf{r}_l) h_m^u(\mathbf{r}_l) + v_O(\mathbf{r}_l) h_m^v(\mathbf{r}_l)] \end{aligned} \quad (2.6.6)$$

where μ is a dummy index as before. The analyzed values Φ_A , u_A , and v_A identically satisfy the imposed constraint (2.6.1) because the winds and geopotential are fitted simultaneously; the basis functions satisfy the relation; and the expansion coefficients are common. In the terminology to be introduced in Chapter 8, (2.6.1) is a strong constraint on the multivariate analysis because it is exactly satisfied.

Equation (2.6.6) is the multivariate equivalent of the univariate normal equations (2.3.3). The normal equations (2.6.6) can also be written in matrix form corresponding to (2.3.5). The Gram matrix in this case would have the same order as in the corresponding univariate matrix and, except for special networks, would be a full matrix. Over irregular observing networks, the multivariate fitting algorithm (2.6.4) and (2.6.6) has similar problems to those discussed in Sections 2.3 and 2.4: ill-conditioning, matrices too large to invert, unacceptable analyses in data voids, and so on.

(2.6.4) During the 1970s, a multivariate function-fitting algorithm developed by Flattery (1971) was used operationally at the National Meteorological Center in Washington. The analysis domain was the whole earth's atmosphere and the fit was global in the sense of Section 2.3. The analyzed geopotential and wind were assumed to be expanded in a truncated series of Hough functions.

Hough functions are the eigenfunctions of a linearized form of the governing equations on a sphere, known as the Laplace tidal equations. They will be discussed in detail in Chapter 9, but for the present, the following simplified description will suffice. Each Hough mode is a function of latitude ϕ and longitude λ and has three components – a zonal (eastward) wind component, a meridional (northward) wind component, and a geopotential component. The Hough modes have a distinct horizontal scale and an associated eigenfrequency; they are orthogonal over the sphere in the continuous case. The Hough modes are often divided into two classes. Low-frequency Rossby–Hough modes tend to satisfy the geostrophic relation (2.6.1) in the extratropics whereas the higher frequency Hough modes correspond to inertia–gravity waves (see Chapter 6).

(2.6.5) The frequency properties of Hough modes can be exploited as an implicit dynamical constraint. If the geopotential and wind are expanded in a series of Rossby–Hough modes only, the resulting objective analysis will satisfy a more general form of the geostrophic relation (2.6.1) that has some validity over the whole globe and not just in the extratropics. The constraints implied by a Rossby–Hough expansion are still linear, however.

(2.6.6) In principle, a multivariate fitting algorithm using Rossby–Hough functions can be based on (2.6.4) and (2.6.6) with $\mathbf{r} = (\lambda, \Phi)$ and with the basis functions $h_m^\phi(\mathbf{r})$, $h_m^u(\mathbf{r})$, and $h_m^v(\mathbf{r})$ being the components of the Rossby–Hough functions. Some practical difficulties exist, however, because the Gram matrix is too large to invert. Flattery (1971) circumvented this problem by first locally interpolating the observations to a uniform global analysis grid. The Rossby–Hough functions are orthogonal on this special grid – [in the same way that trigonometric functions are orthogonal on the network (2.5.3)] – and the off-diagonal elements of the Gram matrix vanish. Therefore, (2.6.6) collapsed to a form similar to (2.5.4) and could be easily solved. However, Flattery's procedure does not strictly minimize (2.6.5) because of the preliminary interpolation of the raw observations to the special global grid. Flattery's approach also fits the geopotential and winds separately; so the expansion coefficients are not common, and the resulting objective analyses does not exactly satisfy the imposed dynamic constraint. More recent attempts to fit global atmospheric observations to Hough functions have been made by Halberstam and Tung (1984).

(2.6.7) Multivariate function fitting can be examined further in the very simple one-dimensional context of the previous section. Suppose the analysis domain is $-\pi \leq x \leq \pi$ and the imposed dynamic constraint on the objective analysis is

$$\frac{d\Phi_A}{dx} = f_0 v_A \quad (2.6.7)$$

Assume Φ_A and v_A are expanded as

$$\Phi_A(x) = \sum_{m=-M}^M c_m e^{imx} \quad \text{and} \quad v_A(x) = \sum_{m=-M}^M \frac{im}{f_0} c_m e^{imx} \quad (2.6.8)$$

satisfying (2.6.7).

Consider observations $\Phi_O(x_k)$ and $v_O(x_k)$, $1 \leq k \leq K$, with expected observation error variances $E_\Phi^2 = \langle \varepsilon_\Phi^2(x_k) \rangle$ and $E_v^2 = \langle \varepsilon_v^2(x_k) \rangle$, which are independent of k . The normal equations (from 2.6.6) are

$$\sum_{\mu=-M}^M c_\mu [1 + m\mu\gamma] \sum_{k=1}^K e^{i(\mu-m)x_k} = \sum_{k=1}^K e^{-imx_k} [\Phi_O(x_k) - imf_0\gamma v_O(x_k)] \quad (2.6.9)$$

where $\gamma = E_\Phi^2/f_0^2 E_v^2$. In the derivation of (2.6.9), $w_\Phi(x_k) = 0.5E_\Phi^{-2}$, $w_v(x_k) = 0.5E_v^{-2}$, and the normal equations have been multiplied by E_Φ^2 .

Over the special network (2.5.3), the off-diagonal terms of the Gram matrix vanish and

$$c_m = q_m \sum_{k=1}^K e^{-imx_k} [\Phi_O(x_k) - imf_0\gamma v_O(x_k)] \quad (2.6.10)$$

where $q_m = (1 + m^2\gamma)^{-1}$. At an analysis gridpoint x_i , $\Phi_A(x_i)$ and $v_A(x_i)$ can be written as

$$\begin{aligned} \Phi_A(x_i) &= \sum_{k=1}^K W_{\Phi\Phi}(x_k - x_i) \Phi_O(x_k) + \sum_{k=1}^K W_{\Phi v}(x_k - x_i) v_O(x_k) \\ v_A(x_i) &= \sum_{k=1}^K W_{v\Phi}(x_k - x_i) \Phi_O(x_k) + \sum_{k=1}^K W_{vv}(x_k - x_i) v_O(x_k) \end{aligned} \quad (2.6.11)$$

where

$$\begin{aligned} W_{\Phi\Phi}(x_k - x_i) &= \frac{\Delta x}{\pi} \left[\frac{1}{2} + \sum_{m=1}^M q_m \cos m(x_k - x_i) \right] \\ W_{\Phi v}(x_k - x_i) &= -\frac{\Delta x \gamma f_0}{\pi} \sum_{m=1}^M m q_m \sin m(x_k - x_i) \\ W_{v\Phi}(x_k - x_i) &= \frac{\Delta x}{f_0 \pi} \sum_{m=1}^M m q_m \sin m(x_k - x_i) \\ W_{vv}(x_k - x_i) &= \frac{\Delta x \gamma}{\pi} \sum_{m=1}^M m^2 q_m \cos m(x_k - x_i) \end{aligned}$$

are the multivariate a posteriori weight functions corresponding to (2.5.6) in the univariate case. Δx is defined in (2.5.3) and γ has dimensions of length squared.

Figure 2.5(a) shows $W_{\Phi\Phi}(x_k - x_i)/\Delta x$ and (b) shows $W_{\Phi v}(x_k - x_i)/f_0 \Delta x$ as functions of $x_k - x_i$, in the same format as Figure 2.3. All plots are for $M = 4$ (9 degrees of freedom). In (a) and (b), the three curves correspond to $\gamma = 0, 0.2$, and 100 . $\gamma = 0$

(2.6.8)

servation
of k . The

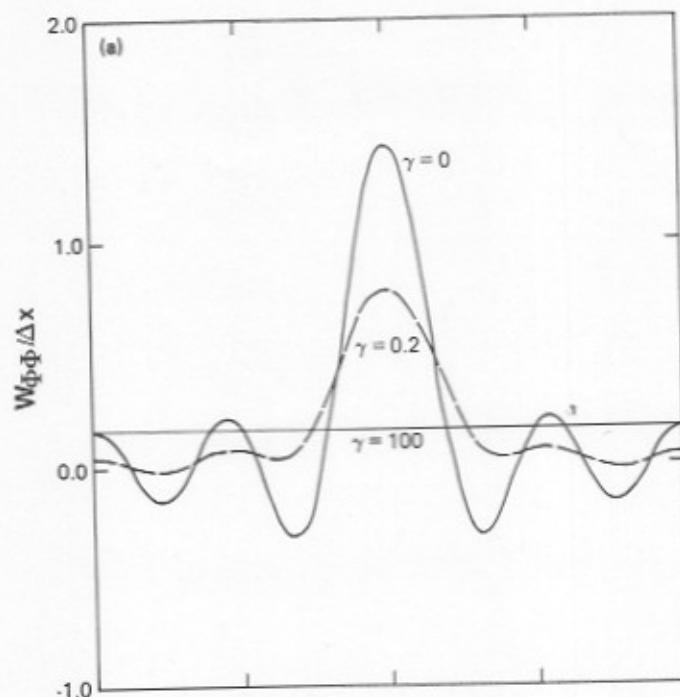
(2.6.9)

$0.5E_v^{-2}$,

x vanish

(2.6.10)

written as



(2.6.11)

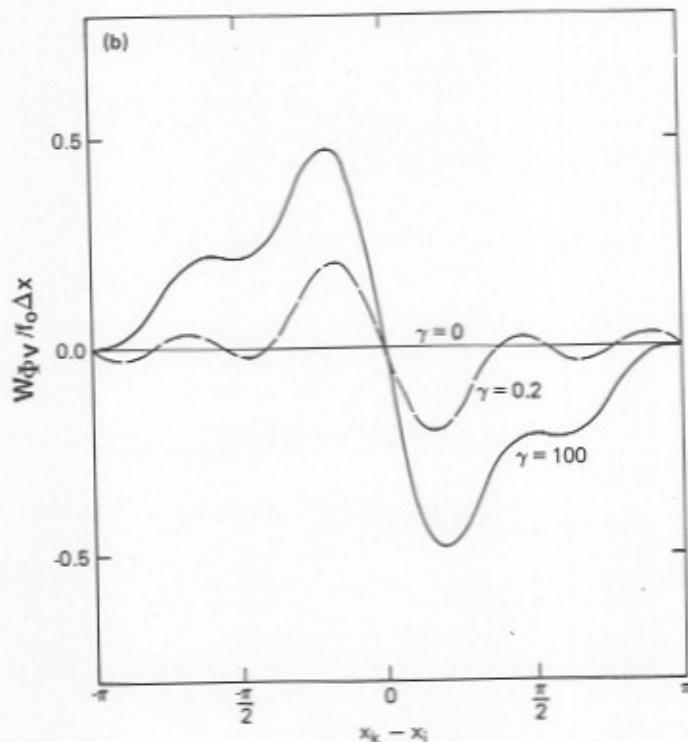


Figure 2.5 The a posteriori weights: (a) $W_{\Phi\Phi}/\Delta x$ and (b) $W_{\Phi v}/f_0 \Delta x$ as a function of $x_k - x_i$ for three values of γ in the multivariate case on the equally spaced network.

implies highly inaccurate or nonexistent wind observations, and $\gamma = 100$ implies highly inaccurate or nonexistent geopotential observations.

In Figure 2.5(a), we see that as γ is increased, less and less weight is given to the geopotential observations $\Phi_O(x_k)$ in the analysis of the geopotential. Moreover, the weight given to $\Phi_O(x_k)$ becomes increasingly large scale in nature (similar to the weights for small values of M in Figure 2.3). In the limit $\gamma \rightarrow \infty$, $W_{\Phi\Phi}(x_k - x_i)/\Delta x$ approaches a constant, indicating that only the constant coefficient c_0 is derived from the geopotential observations. Thus, the values of c_m for large $|m|$ must be derived from the wind observations.

This is consistent with $W_{\Phi v}(x_k - x_i)/f_0 \Delta x$ plotted in (b). Note that when $x_k = x_i$, $W_{\Phi v}(x_k - x_i) = 0$ for all γ . This indicates that a wind observation that happens to coincide with the analysis gridpoint is given no weight. Thus, in the analysis of $\Phi_A(x_i)$, the largest weights are given to wind observations that are somewhat displaced from the analysis gridpoint. In Figure 2.5(b), it is apparent that negative/positive values of $v_O(x_k)$ for x_k greater/less than x_i will increase $\Phi_O(x_i)$. This is in accord with intuition.

$W_{v\Phi}(x_k - x_i)$ and $W_{vv}(x_k - x_i)$ are not plotted, but their behavior can be inferred from Figure 2.5. The a posteriori weights derived in this section can be compared with a posteriori weights derived for other algorithms such as multivariate statistical interpolation in Chapter 5.

2.7 Constrained minimization: penalty functions

In recent years, function-fitting algorithms have been rarely used in atmospheric objective analysis. An important exception, however, is constrained minimization or spline-fitting algorithms, which have been actively investigated by mathematicians and statisticians for some years and are now discussed in the meteorological literature as well. In these techniques, quadratic forms such as (2.2.9) are minimized subject to imposed mathematical constraints. These algorithms generally contain several user-supplied parameters that should be chosen on the basis of physical characteristics (the spectrum, in particular) of the state variable to be analyzed. Consequently, the imposed formal mathematical constraints are implicitly physical constraints as well.

The technique is first illustrated using one-dimensional trigonometric functions on the domain $x_a \leq x \leq x_b$. The analyzed field $f_A(x)$ is represented as in (2.4.2),

$$f_A(x) = \sum_{m=-M}^M c_m e^{imx} \quad (2.7.1)$$

Define observations $f_O(x_k)$, $1 \leq k \leq K$, with expected observation errors $\langle \varepsilon_O^2(x_k) \rangle$ which are normally distributed and spatially uncorrelated. Define the following quadratic form:

$$I(\gamma, p) = I_1(f_A) + \gamma J_p(f_A) \quad (2.7.2)$$

ighly
the
the
)/ Δx
rom
ived
 x_i ,
s to
(x_i),
om
lues
n.
red
red
ical
ric
or
ins
tre
to
er-
ics
he
on
1)
>
g
)

where

$$I_1(f_A) = \sum_{k=1}^K w_k [f_A(x_k) - f_O(x_k)]^2, \quad J_p(f_A) = \int_{x_a}^{x_b} \left[\frac{d^p}{dx^p} f_A(x) \right]^2 dx$$

and $w_k = 0.5 / \langle \varepsilon_O^2(x_k) \rangle$. Also, $f_A(x_k)$ is given by (2.7.1), and γ and p are user-specified positive real constants (p integer).

$I_1(f_A)$ is a measure of the closeness with which $f_A(x)$ fits the observations $f_O(x)$ and is often referred to as the *cost function*. $J_p(f_A)$ is an imposed mathematical constraint known as a *penalty function*. As will be demonstrated shortly, $J_p(f_A)$ is a measure of the smoothness of the objective analysis $f_A(x)$.

Now choose $x_a = -\pi$ and $x_b = \pi$, and minimize (2.7.2) with respect to the coefficients c_m in (2.7.1), which yields the normal equations

$$\sum_{\mu=-M}^M c_\mu \sum_{k=1}^K w_k e^{i(\mu-m)x_k} + 2\pi\gamma m^{2p} c_m = \sum_{k=1}^K w_k f_O(x_k) e^{-imx_k} \quad (2.7.3)$$

using (2.4.6).

The Gram matrix in (2.7.3) is similar to that of (2.4.4) except that it has the additional term $2\pi\gamma m^{2p}$ on the main diagonal. This extra term makes the Gram matrix more diagonally dominant. In the nonconstrained case ($\gamma = 0$), the Gram matrix can be ill-conditioned if two of the basis functions (and, thus, two of the rows) are similar. The addition of a term $2\pi\gamma m^{2p}$ ($\gamma > 0$) to the main diagonal of the Gram matrix makes these two rows less similar, increases the smallest eigenvalue λ_s [see (2.3.16)], and improves the conditioning of the matrix.

Now consider the observation network (2.5.3) and all w_k equal to w . The normal equations (2.7.3) for this special network can be written as

$$c_m = \frac{1}{1 + \alpha m^{2p}} \frac{1}{K} \sum_{k=1}^K f_O(x_k) e^{-imx_k}, \quad -M \leq m \leq M \quad (2.7.4)$$

and $\alpha = 2\pi\gamma/Kw$.

Equation (2.7.4) becomes equal to (2.5.4) when $\gamma = 0$. The expansion coefficient c_m is reduced by a factor of $(1 + \alpha m^{2p})^{-1}$ with respect to the value in (2.5.4). The addition of the term $J_p(f_A)$ acts like a high-wavenumber filter (also known as a low-pass filter) in reducing the amplitude of the large m (small spatial scale) components of the analysis $f_A(x)$. The amount of filtering is controlled by the specification of γ and p ; γ controls the trade-off between the penalty function $J_p(f_A)$, which is a measure of the smoothness of the analysis, and the cost function $I_1(f_A)$, which is a measure of the fidelity of the analysis to the observations; p controls the spectrum of $f_A(x)$, and large values of p imply more effective filtering.

Figure 2.6 illustrates the effect of a penalty function for one-dimensional trigonometric functions over the periodic domain $-\pi \leq x \leq \pi$. The diagram is in the same format as Figure 2.2, with 10 observations indicated by solid dots. Equation (2.7.3) is applied to these observations with $M = 3$ (7 degrees of freedom), $p = 2$, and $w_k = 1$,

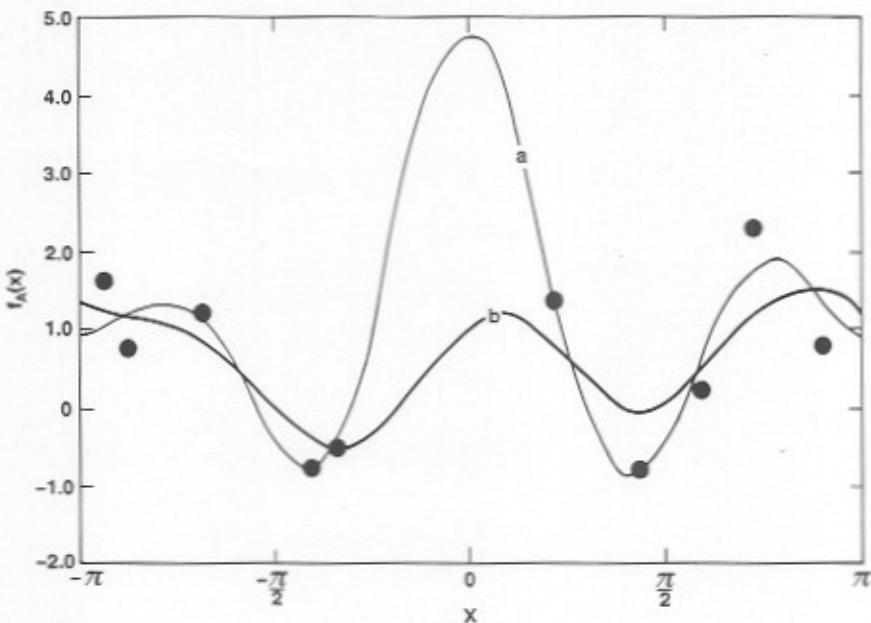


Figure 2.6 The effect of a penalty function on trigonometric fitting. Curve a is unconstrained ($\gamma = 0$), and curve b is constrained ($\gamma = 0.01$).

$1 \leq k \leq 10$. Two curves are shown: $\gamma = 0$ (curve a) and $\gamma = 0.01$ (curve b). As expected from Section 2.4, there is enormous overshooting in the data void near $x = 0$ for the unconstrained case ($\gamma = 0$). The addition of the penalty function has a relative minor effect in the data-rich areas but is very effective in reducing the overshooting in the data void.

One of the desirable properties of any objective analysis algorithm, as noted in Section 1.7, is the filtering of noise (primarily small-scale and nonphysical noise) from the objective analysis. It can be seen that the constrained minimization algorithm (2.7.2) is an effective small-scale filter. Equation (2.7.4) demonstrates the filtering properties (or spectral response) of the algorithm in (2.7.2) in a simple manner. In Chapters 3–5, the spectral responses of other objective analysis algorithms will be discussed in more detail.

In Appendix E, Equation (E9) suggests that cubic-spline functions have a minimum curvature property; that is, they minimize

$$\int_{x_a}^{x_b} [f_A''(x)]^2 dx \quad (2.7.5)$$

This observation was the basis of an important constrained minimization algorithm developed by Reinsch (1967). Consider the minimization of $I(\gamma, p)$ in (2.7.2) with $p = 2$, over the periodic one-dimensional domain, $x_a \leq x \leq x_b$:

$$I(\gamma, 2) = \sum_{k=1}^K w_k [f_A(x_k) - f_O(x_k)]^2 + \gamma J_2(f_A) \quad (2.7.6)$$

where $J_2(f_A) = \int_{x_a}^{x_b} [f_A''(x)]^2 dx$.

Denote $f''_A(x_k) = S_k$. Define $f''_A(x)$ by Equation (E3), and evaluate the penalty function of (2.7.6). After integration by parts, the result is

$$J_2(f_A) = \frac{1}{6} \sum_{k=1}^K \Delta_k [S_{k-1}^2 + S_k^2 + (S_{k-1} + S_k)^2] = \underline{S}^T \underline{T} \underline{S} \quad (2.7.7)$$

where $\Delta_k = x_k - x_{k-1}$, \underline{S} is the column vector of length K of unknown values of the second derivative S_k of the objective analysis evaluated at the observation stations. \underline{T} is the symmetric, positive definite, tri-diagonal matrix defined in Equation (E8). The positive definiteness of \underline{T} can be seen from (2.3.13) and (2.7.7).

Equation (2.7.6) can be written in matrix form:

$$I(\gamma, 2) = 0.5[\underline{f}_A - \underline{f}_O]^T \underline{Q}^{-1} [\underline{f}_O - \underline{f}_A] + \gamma \underline{S}^T \underline{T} \underline{S} \quad (2.7.8)$$

where \underline{f}_A and \underline{f}_O are column vectors of length K of values $f_A(x_k)$ and $f_O(x_k)$. The observation error $\varepsilon_O(x_k)$, $1 \leq k \leq K$, is assumed to be spatially uncorrelated with expected variance $\langle \varepsilon_O^2(x_k) \rangle$. \underline{Q} is the diagonal $K \times K$ observation error covariance matrix, and the a priori weights of (2.7.6) are $w_k = 0.5/\langle \varepsilon_O^2(x_k) \rangle$. In (2.7.8), both \underline{S} and \underline{f}_A are unknown at this point.

In the spline interpolation discussed in Appendix E, the approximation for $f_A(x)$ passed exactly through the observations, so $\underline{f}_A = \underline{f}_O$. This is not true in the case of (2.7.8). However, the relation (E8) between \underline{S} and \underline{f}_A is still valid because it is a consequence of requiring continuity in the first and second derivatives of $f_A(x)$. Consequently,

$$\underline{T} \underline{S} = \underline{Q}^T \underline{f}_A \quad (2.7.9)$$

where \underline{Q} is defined in Appendix E. It might be noted that \underline{Q} is actually a square symmetric matrix in the periodic case and, hence, the transpose notation is redundant. However, in the nonperiodic case, \underline{Q} is a rectangular matrix, so the transpose notation is retained. Introducing (2.7.9) into (2.7.8) gives

$$I(\gamma, 2) = 0.5\{[\underline{f}_A - \underline{f}_O]^T \underline{Q}^{-1} [\underline{f}_A - \underline{f}_O] + \underline{f}_A^T \underline{N}^{-1} \underline{f}_A\} \quad (2.7.10)$$

where $\underline{N}^{-1} = 2\gamma \underline{Q} \underline{T}^{-1} \underline{Q}^T$.

Minimization with respect to \underline{f}_A following Section 2.2 yields

$$\underline{Q}^{-1} [\underline{f}_A - \underline{f}_O] + \underline{N}^{-1} \underline{f}_A = 0 \quad \text{or} \quad \underline{f}_A = \underline{N} [\underline{N} + \underline{Q}]^{-1} \underline{f}_O \quad (2.7.11)$$

Because \underline{T} is a symmetric, positive definite matrix, its eigenvalues are real and positive. This implies \underline{T}^{-1} is positive definite because its eigenvalues are the reciprocals of the eigenvalues of \underline{T} . Consequently, $\underline{Q} \underline{T}^{-1} \underline{Q}^T$ and \underline{N} are symmetric, positive definite (and consequently nonsingular). The functional form (2.7.11) might be compared with that of (2.2.15).

More practically, we can write (2.7.11) as

$$[\underline{T} + 2\gamma \underline{Q}^T \underline{Q} \underline{Q}] \underline{S} = \underline{Q}^T \underline{f}_O \quad (2.7.12)$$

Generally, (2.7.12) is solved for \underline{S} , and the objective analysis at a gridpoint $f_A(x_i)$ is

obtained from (E3). The matrix on the left-hand side of (2.7.12) is a sparse matrix with nonzero elements only on the main diagonal and the two subdiagonals on either side of it. It can be inverted simply and efficiently.

The Reinsch algorithm (2.7.8–12) can be generalized to the two-dimensional case following Duchon (1976) and Wahba and Wendelberger (1981). Define a two-dimensional analogue to (2.7.2):

$$I(\gamma, p) = \sum_{k=1}^K [f_A(x_k, y_k) - f_O(x_k, y_k)]^2 + \gamma J_p(f_A) \quad (2.7.13)$$

where

$$J_p(f_A) = \int_x \int_y \sum_{q=0}^p \binom{p}{q} \left[\frac{\partial^p f_A}{\partial x^q \partial y^{p-q}} \right]^2 dx dy$$

is the penalty function. Here $\binom{p}{q}$ defines the binomial coefficients. When $p = 2$, the penalty function becomes

$$J_2(f_A) = \int_x \int_y \left[\left(\frac{\partial^2 f_A}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 f_A}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f_A}{\partial y^2} \right)^2 \right] dx dy \quad (2.7.14)$$

Minimization of (2.7.13) with $p = 2$, leads to thin-plate smoothing splines, the natural generalization of the Reinsch algorithm. The spectral filtering characteristics of (2.7.13) are given by a generalization of (2.7.4), $(1 + \alpha k^{2p})^{-1}$, where k is a two-dimensional wavenumber. Equation (2.7.13) can be generalized to three or more dimensions. As discussed by Wahba and Wendelberger (1981), if d is the number of dimensions, then $2p - d > 0$. Thus, for the two-dimensional case, $p \geq 2$. Wahba (1982) has also extended (2.7.13) to the spherical case. In (2.7.13), both γ and p must be specified. Methods for their optimal specification are discussed in Craven and Wahba (1979). Constrained minimization will also be approached from the point of view of weak variational constraints and the Euler–Lagrange equations in Chapter 8.

Spline-fitting algorithms are ultimately based on a minimization principle, as are the statistical interpolation algorithms of Chapters 4 and 5. The objective analysis technique to be discussed in the next chapter is not.

Exercises

- 2.1 The column vector $f_A - f_O$ (the difference between the analyzed and observed values at the observation locations) is called the residual. Show that the residual is orthogonal to each of the basis functions $h_m(r_k)$, $0 \leq m \leq M$, over the network r_k , $1 \leq k \leq K$, with respect to the a priori weights.
- 2.2 Write a quadratic form corresponding to (2.3.2) for the case of spatially correlated observation error. Show that minimizing this expression leads to (2.3.6).
- 2.3 Suppose the column vectors of observed, analyzed, and true values f_O , f_A , and f_T defined at the observation locations satisfy

$$f_O = f_T + \varepsilon_O, \quad f_A = Hc, \quad \text{and} \quad f_T = H\tilde{c}$$

where H coefficients covarian the colou respectively
 (a) Show (b) Show gives

- 2.4 Consider observational two basic (a) Two (b) Three Using the for the t analysis less than
- 2.5 Consider uncorrelated at x_1 and for the c are real
- 2.6 Consider Assume $E_O^2(x_1) =$ Construct and generate
- 2.7 For the (2.5.6) s (a) (b) (c)

where \underline{H} is the matrix defined in (2.3.4), $\underline{\epsilon}_c$ and $\underline{\epsilon}_A$ are the analyzed and true expansion coefficients, and $\underline{\epsilon}_O$ is the column vector of observation errors with observation error covariance \underline{Q} . (See, for example, Figure 2.2a.) Define $\underline{\epsilon}_e = \underline{\epsilon}_c - \underline{\epsilon}_A$ and $\underline{\epsilon}_A = \underline{f}_A - \underline{f}_T$ to be the column vectors of coefficient errors and analysis errors (at the observation locations), respectively.

- (a) Show that if $\langle \underline{\epsilon}_O \rangle = 0$, then $\langle \underline{\epsilon}_e \rangle = 0$ and $\langle \underline{\epsilon}_A \rangle = 0$.
 (b) Show that the coefficient error covariance and analysis error covariance matrices are given by

$$\langle \underline{\epsilon}_e \underline{\epsilon}_e^T \rangle = [\underline{H}^T \underline{Q}^{-1} \underline{H}]^{-1} \quad \text{and} \quad \langle \underline{\epsilon}_A \underline{\epsilon}_A^T \rangle = \underline{H} [\underline{H}^T \underline{Q}^{-1} \underline{H}]^{-1} \underline{H}^T$$

- 2.4 Consider an observation network with uncorrelated observation errors with expected observation error variance equal to E_O^2 at the observation locations. Suppose there are two basis functions $h_0(x) = 1$ and $h_1(x) = x$. Consider two cases:

- (a) Two observation locations at x_1 and x_2
 (b) Three observation locations at x_1 , x_2 , and x_3

Using the expressions derived in Exercise 2.3, find the coefficient error covariance matrix for the two cases. What happens when all observations coincide? Show that the expected analysis error variances at the observation locations are equal to E_O^2 in case a and are less than or equal to E_O^2 in case b.

- 2.5 Consider an observation network with two observation stations x_1 and x_2 with uncorrelated observation error whose expected observation error variance is equal to 1 at x_1 and x_2 . Assume the basis functions are the same as in Exercise 2.4. Find an expression for the eigenvalues of the Gram matrix $\underline{G} = \underline{H}^T \underline{H}$ in this case. Show that the eigenvalues are real and nonnegative and that the condition number becomes infinite as $x_2 \rightarrow x_1$.

- 2.6 Consider the observation network with three observation stations at x_1 , x_2 , and x_3 . Assume the observation errors are uncorrelated with expected observation error variance $E_O^2(x_1) = E_O^2(x_3) = 1$ and $E_O^2(x_2) = \frac{1}{2}$. Assume basis functions $h_0(x) = 1$ and $h_1(x) = x$. Construct a new set of orthogonal basis functions using the Gram-Schmidt procedure and generate a Gram matrix for the new set.

- 2.7 For the equally spaced network (2.5.3) show that the a posteriori weights given in Equation (2.5.6) satisfy the following:

(a)
$$\sum_{k=1}^K W_{ik} = 1$$

(b)
$$W_{ik} \leq W_{ii} \quad \text{and} \quad 0 \leq W_{ii} \leq 1$$

(c)
$$W_{ik} = \delta_{ik} \quad \text{for a fully determined fit}$$