COMPUTING PRESCRIPTIONS

Editors: Isabel Beichl, isabel.beichl@nist.gov

FITTING NATURE'S BASIC FUNCTIONS PART I: POLYNOMIALS AND LINEAR LEAST SQUARES

By Bert W. Rust

THE PROBLEM OF FITTING A MATHEMATICAL MODEL

$$y(t) = \phi(t, \alpha), \tag{1}$$

which depends on an n-vector

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T \tag{2}$$

of unknown parameters, to a measured data set

$$\{(t_i, y_i), i = 1, 2, \ldots, m\} \tag{3}$$

is ubiquitous in science and engineering. This column is the first installment of a series that will demonstrate modern techniques for fitting combinations of basic mathematical functions to measured real-world data.

When $\phi(t, \alpha)$ depends nonlinearly on one or more of the α_j, the problem is considerably more difficult than the case where it depends linearly on all of them. In both cases, the underlying statistical model has the form

$$y_i = \phi(t_i, \alpha^*) + \in_i, \qquad i = 1, 2, \ldots, m, \tag{4}$$

where α^* is the true, but unknown, parameter vector, and the $\in_i$ are unknown random errors that we usually assume to be normally distributed. We attribute all these errors to the y_i, with the t_i being known either exactly or at least more precisely than the y_i. The dominant source of the $\in_i$ can be either measurement errors or an inherent component of random variation in the process generating the y_i.

For both linear and nonlinear fits, we find the "best" estimate $\hat{\alpha}$ by minimizing the *objective function*

$$\mathcal{L}(\alpha) = \sum_{i=1}^{m} \left[y_i - \phi(t_i, \alpha) \right]^2, \tag{5}$$

which is readily recognized to be the sum of squared residuals formed by subtracting the model prediction at $t = t_i$ from the corresponding measured y_i. It is not obvious that minimizing the sum of squared residuals gives the best fit, but this principle of least squares has been widely accepted since Gauss first enunciated it. For fitting linear models, we will rigorously demonstrate that the least-squares estimate is the unbiased estimate with minimum variance.

Two competing criteria govern the choice of the model $\phi(t, \alpha)$. The first is to make the distribution of the residuals as random as possible. The second is to keep the number of unknown parameters as low as possible. We can always fit any set of m points (t_i, y_i) exactly if we choose the model to be a polynomial of degree $m - 1$. However, such a choice would not satisfy either of the two criteria, and, in most cases, it would produce unrealistic wiggles in the spaces between the data points. Modeling's basic goal is to find a parsimonious, physically plausible representation of the data that produces residuals with a distribution similar to the one assumed for the $\in_i$.

Fitting a straight line

Finding the straight line

$$\phi(t, \alpha) = \alpha_1 + \alpha_2 t \tag{6}$$

that best represents a measured data set is the prototypical fitting problem. The objective function is just

$$\mathcal{L}(\alpha) = \sum_{i=1}^{m} \left[y_i - \alpha_1 - \alpha_2 t \right]^2, \tag{7}$$

and setting

$$\frac{\partial \mathcal{L}}{\partial \alpha_1} = 0, \quad \frac{\partial \mathcal{L}}{\partial \alpha_2} = 0, \tag{8}$$

gives the system of linear equations

$$m\alpha_1 + \alpha_2 \sum_{i=1}^{m} t_i = \sum_{i=1}^{m} y_i \tag{9}$$

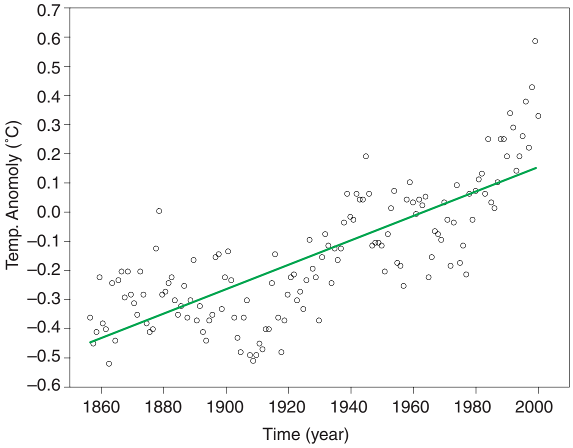

Figure 1. A straight-line fit to yearly average global temperature data.

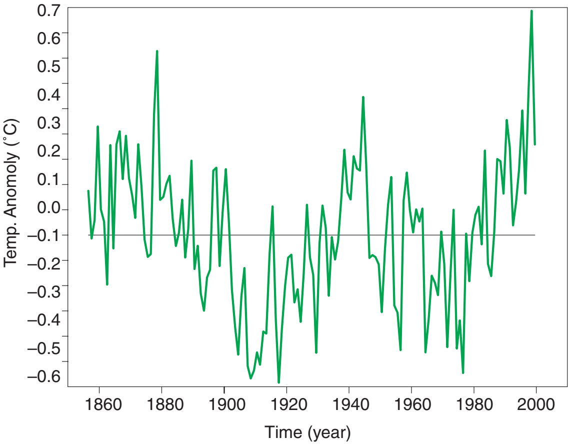

Figure 2. Residuals for the straight-line fit to yearly average global temperature data.

$$\alpha_1 \sum_{i=1}^{m} t_i + \alpha_2 \sum_{i=1}^{m} t_i^2 = \sum_{i=1}^{m} t_i y_i \,, \tag{10}$$

which we can easily solve to give the optimal estimates $\hat{\alpha}_1$ and $\hat{\alpha}_2$.

As an example, let's consider the record of yearly average global temperature anomalies plotted as circles in Figure 1. We obtain these anomalies from the actual average temperatures by subtracting the global average (14.0° C) for the years 1961 to 1990, a procedure that shifts the zero point but does not affect temperature scaling. You can obtain these data, which Phil Jones and his colleagues at the University of East Anglia compiled, from http://cdiac.esd.ornl.gov/trends/temp/jonescru/data.html.[1]

The straight line in Figure 1 is the least-squares fit of the model

$$\phi(t, \alpha) = \alpha_1 + \alpha_2(t - t_0), \tag{11}$$

where $t_0 = 1856.0$. This shift of the zero point for the time scale makes the estimate for α_1 the model prediction for the temperature anomaly at the beginning of the measured record. The least-squares estimates for the parameters are

$$\begin{aligned} \hat{\alpha}_1 &= -0.460 \pm .023 \quad &{}^\circ\mathrm{C} \\ \hat{\alpha}_2 &= (4.25 \pm .28) \times 10^{-3} \quad &{}^\circ\mathrm{C/yr}, \end{aligned} \tag{12}$$

where we calculated the indicated one-standard-deviation uncertainties from the assumption that the unknown errors ϵ_i were independently, identically distributed with a normal distribution having zero mean and unknown variance σ^2. I'll give more details on how we estimate σ^2 and the uncertainties in $\hat{\alpha}_1$ and $\hat{\alpha}_2$ in the next installment of this series.

The global warming indicated by $\hat{\alpha}_2$, although statistically significant, is not alarming. If correct, it would raise temperatures by only 0.085 °C in the next 20 years. But Figure 1 indicates that the straight-line model inadequately represents the data. The plot of the residuals in Figure 2 strengthens this impression, since the undulations around the zero line are clearly not random variations.

Linear least squares

If the straight line model is inadequate, then what model should we choose? The residual plot suggests the possibility of a cycle with a period somewhere in the range of 50 to 75 years. Adding a sinusoidal term to the model could accommodate such a variation. This would give three new unknown parameters corresponding to the sinusoid's amplitude, period, and phase, but the model would then depend nonlinearly on some of the parameters. So, let's defer this option until the third installment, where we will discuss nonlinear least squares.

Another possibility, which can be treated with linear least squares, is to try a higher order polynomial. The residuals in Figure 2 exhibit an apparent local maximum somewhere in the interval [1860, 1880], followed by a local minimum somewhere in [1900, 1920], followed by another local maximum somewhere in [1930, 1950], and finally another local minimum somewhere in [1960, 1980]. These local optima appear also in the data in Figure 1, although perhaps not as clearly as in the residual plot. The lowest degree polynomial that can accommodate such a pattern of variation is one of order 5, so let's try a model of the form

$$y(t) = \phi(t, \alpha) = \sum_{\nu=1}^{6} \alpha_\nu (t - t_0)^{\nu-1}, \tag{13}$$

where, again, $t_0 = 1856.0$. Even though the model is a fifth degree polynomial, it is linear in the unknown parameters α_ν, which can therefore be estimated by linear least squares.

To simplify the notation in explaining how to formulate and solve the fitting problem, let's make the formal change of variables

$$\tau = t - t_0, \qquad t_0 = 1856.0. \tag{14}$$

Writing the model at the points $t_1, t_2, \ldots, t_m$ in vector–matrix form gives

$$\begin{bmatrix} y(\tau_1) \\ y(\tau_2) \\ \cdot \\ \cdot \\ \cdot \\ y(\tau_m) \end{bmatrix} = \begin{bmatrix} 1 & \tau_1 & \tau_1^2 & \ldots & \tau_1^5 \\ 1 & \tau_2 & \tau_2^2 & \ldots & \tau_2^5 \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ 1 & \tau_m & \tau_m^2 & \ldots & \tau_m^5 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \end{bmatrix} \tag{15}$$

or in shorter form,

$$\mathbf{y}(t) = \Phi(\mathbf{t})\alpha, \tag{16}$$

where $\Phi(\mathbf{t})$ is the $m \times 6$ matrix, which depends only on the vector

$$\mathbf{t} = (t_1, t_2, \ldots, t_m)^T, \tag{17}$$

and $\mathbf{y}(\mathbf{t})$ is the m-vector of model predictions. Defining a measurement vector

$$\mathbf{y} = (y_1, y_2, \ldots, y_m)^T \tag{18}$$

and a residual vector

$$\mathbf{r}(\alpha) = \mathbf{y} - \Phi(\mathbf{t})\alpha \tag{19}$$

with elements

$$r_i = [\mathbf{y} - \Phi(\mathbf{t})\alpha]_i, \qquad i = 1, 2, \ldots, m, \tag{20}$$

let us write the objective function in Equation 5 as

$$\mathcal{L}(\alpha) = \sum_{i=1}^{m} \left[\mathbf{y} - \Phi(\mathbf{t})\alpha\right]_i^2 \tag{21}$$

$$= \left[\mathbf{y} - \Phi(\mathbf{t})\alpha\right]^T \left[\mathbf{y} - \Phi(\mathbf{t})\alpha\right] \quad , \tag{22}$$

which we can expand to give

$$\mathcal{L}(\alpha) = \mathbf{y}^T\mathbf{y} - 2\alpha^T\Phi^T(\mathbf{t})\mathbf{y} + \alpha^T\Phi^T(\mathbf{t})\Phi(\mathbf{t})\alpha. \tag{23}$$

This function is not restricted to the present example, which has $n = 6$, but it is valid for any linear least-squares problem. To assure a unique minimum, however, we must require that $n \le m$ and that the columns of $\Phi(\mathbf{t})$ comprise a linearly independent set of m-vectors.

Geometrically, the objective function defines an $(n + 1)$-dimensional, quadratic hypersurface, sometimes called the *response surface*, whose level curves correspond to concentric n-dimensional ellipsoids in the α-space. It has a unique global minimum that we can find by differentiating $\mathcal{L}(\alpha)$ with respect to α and equating the result to the zero vector,

$$\frac{\partial \mathcal{L}}{\partial \alpha} = -2\Phi^T(\mathbf{t})\mathbf{y} + 2\Phi^T(\mathbf{t})\Phi(\mathbf{t})\alpha = \mathbf{0}. \tag{24}$$

Thus, the minimizing α must satisfy the $n \times n$ system of linear equations

$$\Phi^T(\mathbf{t})\Phi(\mathbf{t})\alpha = \Phi^T(\mathbf{t})\mathbf{y}, \tag{25}$$

which are often called the *normal equations*. Because the columns of $\Phi(\mathbf{t})$ are linearly independent, the matrix product on the left is nonsingular, so the unique solution is

$$\hat{\alpha} = \left[\Phi^T(\mathbf{t})\Phi(\mathbf{t})\right]^{-1}\Phi^T(\mathbf{t})\mathbf{y}. \tag{26}$$

Until the early 1960s, linear least-squares estimates were usually calculated by forming the matrix product $\Phi^T(\mathbf{t})\Phi(\mathbf{t})$ and inverting it. But if the problem is ill-conditioned (if relatively small perturbations in the data produce relatively large perturbations in the solution), then we can get more numerically stable algorithms by computing an orthogonal factorization of the form

$$\Phi(\mathbf{t}) = \mathbf{Q}\begin{bmatrix} \mathbf{R} \\ \mathbf{O} \end{bmatrix}, \tag{27}$$

where $\mathbf{Q}$ is an $m \times m$ orthogonal matrix ($\mathbf{Q}^T\mathbf{Q} = \mathbf{I} = \mathbf{Q}\mathbf{Q}^T$), $\mathbf{R}$ is an $n \times n$ upper triangular matrix, and $\mathbf{O}$ is an $(m - n) \times n$ matrix of zeroes. By substituting this factorization into Equation 25, we can easily verify that $\hat{\alpha}$ satisfies the $n \times n$ upper- triangular system

$$\mathbf{R}\alpha = \mathbf{Q}_1\mathbf{y}, \tag{28}$$

where $\mathbf{Q}_1$ is the $m \times n$ matrix formed by the first n columns of $\mathbf{Q}$. This is the basis for the least-squares programs in the Lapack, Linpack, and Matlab collections, which, in the numerical analysis community, are considered to be the gold standards.[2–4] (More details on the QR factorization and its advantages for linear least

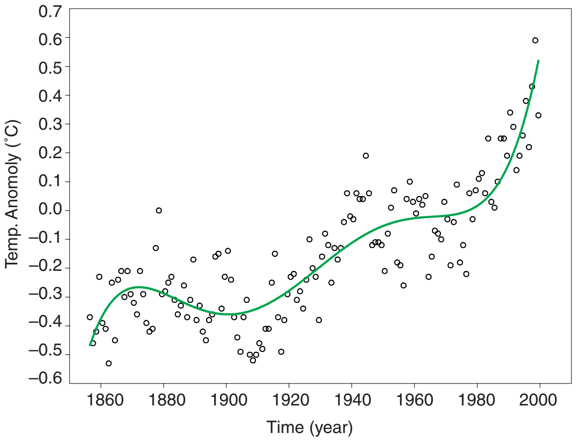

Figure 3. The fifth-degree polynomial fit to yearly average global temperature data.

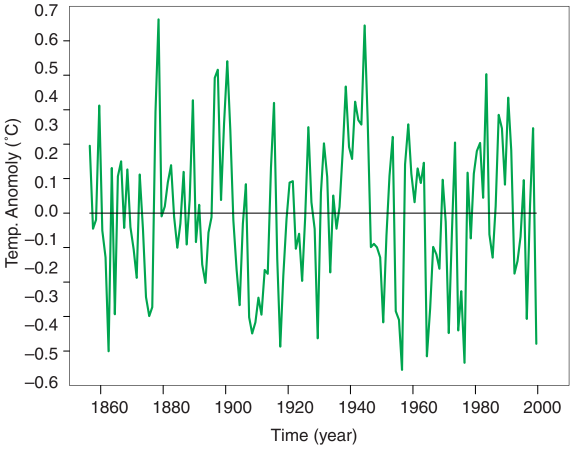

Figure 4. Residuals for the fifth degree polynomial fit to yearly average global temperature data.

squares appear in Stewart's classic text.[5]) These advantages become crucial only for problems in which the columns of $\Phi(\mathbf{t})$ are almost linearly dependent. If the calculations are done in double precision, then the older method will usually work quite well. It has the advantage of producing the matrix $[\Phi^T(\mathbf{t})\Phi(\mathbf{t})]^{-1}$, which we need to calculate the uncertainties in the estimates $\hat{\alpha}_j$.

No matter which algorithm we use to compute the linear least-squares estimate, two points about it are important. The first is that we can compute the $\hat{\alpha}_j$, with the maximum accuracy the data allow, in a finite sequence of calculations. The second is that no prior knowledge or estimates of the values of the α_j are required in making those calculations. Nonlinear least-squares problems do not share these advantages.

The linear least-squares estimates for the fifth order polynomial fit to the global annual temperature record are

$$\begin{aligned}
\hat{\alpha}_1 &= -0.484 \pm .055 && {}^\circ\mathrm{C} \\
\hat{\alpha}_2 &= 0.0336 \pm .0078 && {}^\circ\mathrm{C}/\mathrm{yr} \\
\hat{\alpha}_3 &= (-1.72 \pm .33)\times 10^{-3} && {}^\circ\mathrm{C}/\mathrm{yr}^2 \\
\hat{\alpha}_4 &= (3.35 \pm .59)\times 10^{-5} && {}^\circ\mathrm{C}/\mathrm{yr}^3 \\
\hat{\alpha}_5 &= (-2.69 \pm .45)\times 10^{-7} && {}^\circ\mathrm{C}/\mathrm{yr}^4 \\
\hat{\alpha}_6 &= (7.6 \pm 1.2)\times 10^{-10} && {}^\circ\mathrm{C}/\mathrm{yr}^5\ .
\end{aligned} \tag{29}$$

The most uncertain parameter is $\hat{\alpha}_2$, for which

$$\frac{\hat{\alpha}_2}{\sigma(\hat{\alpha}_2)} = \frac{0.0336}{0.0078} = 4.3\,. \tag{30}$$

So, if the random errors in the data are independently normally distributed as assumed, then all the parameters are statistically significant.

The fit is plotted as the solid curve in Figure 3. It obviously tracks the data better than the straight line in Figure 1, and the corresponding residuals, given in Figure 4, are both smaller and more random looking than those in Figure 2. The improvement in the tracking is especially evident in the last 20 years of the record, where 17 of the data points fall above the straight-line fit. In 1983, the old record high anomaly of 0.19° C, set in 1944, was exceeded. In the following 16 years, this new high of 0.25° C was matched or exceeded nine times. New record highs were established in four of those years. The current record is 0.59° C, set in 1998. If the temperatures in the next 20 years should follow this same fifth-order polynomial, the results would be catastrophic. The model predicts a temperature anomaly of 0.63° C for 2001 and 3.19° C for 2021. But we shouldn't take this prediction too seriously because it is well known that polynomial fits almost always give unrealistic extrapolations, even when they fit the data quite well. Indeed, the fit in Figure 3 illustrates this point if we consider extrapolating the fit backward to 1836.

The question remains of just how well the fit in Figure 3 actually represents the data. The residuals in Figure 4, although more random than those in Figure 2, still display systematic, nonrandom variations, including a weak indication of the 50- to 80-year quasicycle that the fifth-degree polynomial was invoked to explain. So the order 5 polynomial might not be a totally adequate representation of the data, but this does not mean that we should try even higher-order polynomials. In fact, we have not yet established that the improvement in fit obtained thus far is sufficient to justify the addition of the four additional parameters to the model. Let's consider that matter in the next installment of this series.

The best linear unbiased estimate

This section gives a statistical justification for the principle of least squares and, in the process, emphasizes the role the random errors ϵ_i play in the fitting procedure. To simplify the notation, let $\Phi \equiv \Phi(\mathbf{t})$. Then the most general statistical assumptions for a linear model are that

$$\mathbf{y} = \Phi\alpha^* + \in, \qquad E(\in) = \mathbf{0}, \qquad E(\in\in^T) = \Sigma^2, \tag{31}$$

where α^* is the "true" parameter vector, $\in$ is the m-vector of unknown random errors, E is the expectation operator, and Σ^2 is the symmetric, positive definite variance matrix. If we consider $\mathbf{y}$ as a random vector, then we can also write the statistical assumptions as

$$E(\mathbf{y}) = \Phi\alpha^*, \tag{32}$$

$$E\left(\left[\mathbf{y} - \Phi\alpha^*\right]\left[y - \Phi\alpha^*\right]^T\right) = \Sigma^2 . \tag{33}$$

Until now, we have assumed that

$$\Sigma^2 = \sigma^2\mathbf{I}_m, \tag{34}$$

where σ^2 is an unknown variance common to all of the errors, which we also assume to be uncorrelated with one another. For many problems the errors are uncorrelated, but the variances do not all have the same value. In such cases,

$$\Sigma^2 = \mathbf{diag}\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2\right), \tag{35}$$

where we might or might not know the values of the σ_i^2.

The assumptions about $\in$ in Equation 31 are all that we need to guarantee that the least-squares estimate is the *best linear unbiased estimate* of α^*. To see what this means, write the separate elements of α^* as

$$\alpha_i^* = \mathbf{e}_i^T\alpha^*, \qquad i = 1, 2, \ldots, m, \tag{36}$$

where $\mathbf{e}_i$ is the unit vector whose ith element is 1, meaning the ith column of $\mathbf{I}_m$. A linear estimate for α_i^* is a linear combination $\mathbf{u}_i^T\mathbf{y}$ of the elements of $\mathbf{y}$, where $\mathbf{u}_i$ is a vector chosen to give the estimate the desired properties. The first desired property is that the estimate be unbiased. This means that $\mathbf{u}_i$ must satisfy

$$E\left(\mathbf{u}_i^T\mathbf{y}\right) = \mathbf{e}_i^T\alpha^*, \tag{37}$$

which, by Equation 32, will be satisfied if

$$\mathbf{u}_i^T\Phi = \mathbf{e}_i^T . \tag{38}$$

Many vectors $\mathbf{u}_i$ might satisfy this condition, and all of them give unbiased estimates $\mathbf{u}_i^T\mathbf{y}$. The best linear unbiased estimate is the one with minimum variance. The variance of the estimate is

$$V\left(\mathbf{u}_i^T\mathbf{y}\right) = E\left[\left(\mathbf{u}_i^T\mathbf{y} - \mathbf{e}_i^T\alpha^*\right)\left(\mathbf{u}_i^T\mathbf{y} - \mathbf{e}_i^T\alpha^*\right)^T\right] , \tag{39}$$

which, by Equations 38 and 33, becomes

$$V\left(\mathbf{u}_i^T\mathbf{y}\right) = \mathbf{u}_i^T\Sigma^2\mathbf{u}_i. \tag{40}$$

We can thus write the problem to be solved as

$$V\left(\hat{\mathbf{u}}_i^T\mathbf{y}\right) = \min_{\mathbf{u}_i}\left\{\mathbf{u}_i^T\Sigma^2\mathbf{u}_i \mid \mathbf{u}_i^T\Phi = \mathbf{e}_i^T\right\}. \tag{41}$$

Applying the method of Lagrange multipliers to this constrained minimization problem gives

$$\hat{\mathbf{u}}_i^T = \mathbf{e}_i^T\left[\Phi^T\Sigma^{-2}\Phi\right]^{-1}\Phi^T\Sigma^{-2}, \tag{42}$$

so the minimum variance unbiased estimate is

$$\hat{\alpha}_i = \hat{\mathbf{u}}_i^T\mathbf{y} = \mathbf{e}_i^T\left[\Phi^T\Sigma^{-2}\Phi\right]^{-1}\Phi^T\Sigma^{-2}\mathbf{y} . \tag{43}$$

Stacking all n of these estimates into a single n-vector gives

$$\hat{\alpha} = \left[\Phi^T\Sigma^{-2}\Phi\right]^{-1}\Phi^T\Sigma^{-2}\mathbf{y} , \tag{44}$$

which is analogous to the least-squares estimate in Equation 26. For the latter estimate, we implicitly assumed that the error variance matrix was given by Equation 34. When this matrix is substituted into Equation 44, the σ^2 factors cancel out, so we can compute the least-squares estimate even when we don't know the value of σ^2.

In the more general case, the estimate in Equation 44 is the solution to a weighted least-squares problem where the weights are obtained from the matrix Σ^2. If the errors are uncorrelated, as in Equation 35, then it is easy to compute the matrix

$$\Sigma^{-1} = \mathbf{diag}\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \ldots, \frac{1}{\sigma_m}\right) \tag{45}$$

and use it to rescale the statistical model in Equation 31:

$$\begin{aligned} &\Sigma^{-1}\mathbf{y} = \Sigma^{-1}\Phi\alpha + \Sigma^{-1}\in, \\ &E(\Sigma^{-1}\in) = \mathbf{0}, \\ &E(\Sigma^{-1}\in\in^T\Sigma^{-1}) = \mathbf{I}_m. \end{aligned} \tag{46}$$

This scaled model corresponds to a least-squares problem with a weighted objective function

$$\mathcal{L}(\alpha) = [\mathbf{y} - \Phi\alpha]^T \Sigma^{-2} [\mathbf{y} - \Phi\alpha] \quad (47)$$

$$= \sum_{i=1}^{m} \frac{1}{\sigma_i^2} [\mathbf{y} - \Phi\alpha]_i^2 , \quad (48)$$

whose minimum value occurs at the estimate in Equation 44. If Σ^2 is not diagonal, it is still positive definite, so it has a computable Cholesky factorization

$$\Sigma^2 = \mathbf{L}\mathbf{L}^T, \quad (49)$$

where $\mathbf{L}$ is a lower triangular matrix that we can easily invert. Scaling Equation 31 with $\mathbf{L}^{-1}$ produces an analogous weighted least-squares problem.

This section concludes the first installment in this series of articles. I've given only the briefest review of linear estimation theory. You can find more details on this important subject in Alexander Mood and Franklin Graybill's classic book and Graybill's more recent encyclopedic text on the subject.[6,7] These books are also good sources on the statistical analysis and interpretation of least-squares estimates, a subject I will discuss in more detail in the second installment in this series.

References

1. P.D. Jones et al., "Global and Hemispheric Temperature Anomalies: Land and Marine Instrumental Records," *Trends: A Compendium of Data on Global Change*, Carbon Dioxide Information Analysis Center, Oak Ridge Nat'l Laboratory, Oak Ridge, Tenn., 2000.
2. E. Anderson et al., *LAPACK Users' Guide*, Soc. for Industrial and Applied Mathematics, Philadelphia, 1992.
3. J.J. Dongarra et al., *LINPACK Users' Guide*, Soc. for Industrial and Applied Mathematics, Philadelphia, 1979.
4. *MATLAB: The Language of Technical Computing*, The MathWorks, Natick, Mass., 2000.
5. G.W. Stewart, *Introduction to Matrix Computations*, Academic Press, New York, 1973.
6. A.M. Mood and F.A. Graybill, *Introduction to the Theory of Statistics*, McGraw-Hill, New York, 1963.
7. F.A. Graybill, *Theory and Application of the Linear Model*, Duxbury Press, North Scituate, Mass., 1976.

Bert W. Rust is a research mathematician at the National Institute of Standards and Technology. His research interests include ill-posed problems, time-series modeling, nonlinear regression, and observational cosmology. He received a BS in engineering physics and an MS in mathematics from the University of Tennessee and his PhD in astronomy from the University of Illinois. Contact him at the Nat'l Inst. of Standards and Technology, 100 Bureau Dr., Stop 8910, Gaithersburg, MD 20899-8910; bwr@cam.nist.gov.

How to Reach *CiSE*

Writers

For detailed information on submitting articles, write to cise@computer.org or visit http://computer.org/cise/edguide.htm.

Letters to the Editors

Send letters to

Jenny Ferrero, Contact Editor
jferrero@computer.org

Please provide an email address or daytime phone number with your letter.

On the Web

Access http://computer.org/cise or http://ojps.aip.org/cise for information about *CiSE*.

Subscription Change of Address (IEEE/CS)

Send change-of-address requests for magazine subscriptions to address.change@ieee.org. Be sure to specify *CiSE*.

Subscription Change of Address (AIP)

Send general subscription and refund inquiries to subs@aip.org.

Subscribe

Visit http://ojps.aip.org/cise/subscrib.html or http://computer.org/subscribe.

Missing or Damaged Copies

If you are missing an issue or you received a damaged copy, contact membership@computer.org.

Reprints of Articles

For price information or to order reprints, send email to cise@computer.org or fax +1 714 821 4010.

Reprint Permission

To obtain permission to reprint an article, contact William Hagen, IEEE Copyrights and Trademarks Manager, at whagen@ieee.org.