LEAST SQUARES ADJUSTMENT
A MODERN APPROACH

by
PETER MEISSL

Part D: SPECIAL TOPICS

MITTEILUNGEN
der geodätischen Institute der Technischen Universität Graz
Folge 43

Graz, 1982

Herausgeber:
Geodätische Institute der Technischen Universität Graz

Redaktion für diese Folge:
Abteilung für Mathematische Geodäsie und Geoinformatik
des Institutes für Theoretische Geodäsie

Mit freundlicher Genehmigung der Geodätischen Institute der Technischen Universität Graz wurde diese Folge am Institut für Geodäsie und Geoinformation der Universität Bonn eingescannt.

Druck und Herstellung:
Druck- und Kopierzentrum der Technischen Universität Graz

Adresse:
Technische Universität Graz
Rechbauerstraße 12
A-8010 Graz, Österreich.

Preface

For his lectures at the Tongji University in Shanghai and at other institutions in China in November - December 1981, Peter Meissl prepared a set of lecture notes on contemporary least-squares adjustment and applications. Subsequently he worked on correcting and expanding them, but this was interrupted by his tragic death on May 22, 1982. (For Peter Meissl's life and work, the reader is referred to his biography by Franz Allmer, Mitteilungen der geodätischen Institute der Technischen Universität Graz, Folge 44, 1983.)

In view of the unique importance of this work, the Institute of Theoretical Geodesy decided to edit the manuscript posthumously and to publish the book in the series of the Geodetic Institutes of the Technical University, Graz, although Peter Meissl himself would certainly have included additional topics such as inner adjustment theory, expanded others such as the theory of large networks, and polished the manuscript much more before being satisfied with its publication.

The finishing of the book is due to Peter Meissl's closest associates: Dr. Norbert Bartelme, Dr. Helmut Fuchs, Dr. Bernhard Hofmann-Wellenhof, Dipl.-Ing. Wolf-Dieter Schuh and Dipl.-Ing. Manfred Wieser. In addition to being responsible for the careful editing of the manuscript, they also prepared the printing text using the word processing facilities of the computer WANG 2200 MVP.

A glance at the table of contents shows that this book is a thoroughly modern text on least-squares adjustment. In the contemporary spirit, the usual linear algebra is treated in the context of general linear spaces, which makes possible an easy transition to Hilbert space important for advanced topics. Also modern is the division into an algebraic and geometric approach (without statistics) and a stochastic approach, including statistical tests. Applications to Doppler observations, large networks, geodetic data bases, and splines essentially increase the practical usefulness. Although the book develops adjustment theory in a systematic and self-contained way, it will be best appreciated by readers who already have some elementary previous knowledge of adjustment computations.

The book needs no recommendation. Both students and research workers will find it indispensable. It is a fitting memorial of a great scientist.

Helmut Moritz

TABLE OF CONTENTS

A.5. Projectors

1. Decomposition of a vector space into a direct sum of subspaces
2. Orthocomplementary subspaces
3. Theorem by Pythagoras
4. Matrix representation of orthogonal projectors
5. Projections of functionals

A.6. Least squares adjustment

1. Projecting the vector of observations
2. Inhomogeneous form of least squares adjustment
3. Fundamental rectangular triangle of least squares adjustment
4. Least squares adjustment by projecting functionals

A.7. Partitioned matrices

1. Definitions
2. Computational rules
3. Block diagonality
4. Block Gauss elimination
5. Theoretical background of partitioned matrices

A.8. Isometric mappings between inner product spaces

1. Definitions
2. Preservation of inner products
3. Matrix representation
4. Examples of isometric mappings
5. Canonical transformation of an adjustment problem

A.9. Partial reduction

1. Partitioning the set of parameters
2. Partial reduction of the normal equations
3. Orthogonal decomposition of the parameter space
4. Partially reduced observation equations
5. Alternative derivation of the partially reduced observation equations

A.10. Adjustment phased with respect to observations

1. Formulation of the problem
2. Addition of normal equations
3. Updating the solution of the previous phase
4. Geometrical insight
5. Pre-elimination of group-internal unknowns
6. Helmert blocking

A.11. Complementary extremum principles in least squares adjustment

1. Basic geometric principle
2. Reformulation for linear manifolds
3. Adjustment by minimizing the norm of the residuals
4. Adjustment by minimizing variances

A.12. Generalized inverses

1. Range space and null space of a linear operator
2. g-inverse
3. Reflexive generalized inverse
4. Generalized inverse with least squares property
5. Generalized inverse with minimum norm property
6. Minimum norm least squares inverse
7. Pseudo inverse

A.13. Adjustment of rank-deficient systems

1. Formulation of the problem
2. Solution via generalized inverse of A
3. A minimum norm property of the covariance matrix of adjusted parameters
4. Solution via singular normal equations
5. Calculation of the ℓm-inverse
6. Application to free network adjustment

B. STOCHASTIC APPROACH TOWARD LEAST SQUARES ADJUSTMENT

B.1. Probabilities

1. Relative frequencies
2. Probability space
3. Examples
4. Calculus of probabilities

B.2. Random variables

1. One-dimensional random variables
2. Probability density function
3. n-dimensional random variables
4. Functions of random variables
5. Marginal distribution
6. Stochastic independence

B.3. Expectation, variances and covariances

1. Expectation of a one-dimensional random variable
2. Variance of a one-dimensional random variable
3. Various kinds of observation errors
4. Simple computational rules for expectation and variance
5. The case of higher dimensional random variables
6. Covariance matrix
7. Propagation of expectations and covariances
8. Important special cases
9. Zero correlation and stochastic independence

B.4. Gauss-Markoff model of least squares adjustment

1. Stochastic model
2. Unbiased estimates
3. Best linear unbiased estimation
4. Error calculus

B.5. Applications of the error propagation law

1. Triangle with three measured sides
2. First fundamental problem in the plane
3. Error ellipses
4. Polar survey with redundancy
5. Area calculated from polar survey
6. Conventionally adjusted regular traverse
7. Rigorously adjusted regular traverse
8. Systematic errors in a regular traverse

C. CONFIDENCE REGIONS AND TESTS OF LINEAR HYPOTHESES

C.1. Probability distributions used in statistical tests

1. One-dimensional Gauss distribution (normal distribution)
2. Multi-dimensional Gauss distribution (normal distribution)
3. Chi-squared distribution (χ^2-distribution)
4. Student's distribution (t-distribution)
5. Fisher's distribution (F-distribution)

C.2. Canonical transformation

1. Preliminaries
2. Making the functionals a part of the parameters
3. Orthogonal decomposition of the space L_A
4. Orthogonal decomposition of L into L_A and L_B
5. Orthonormalizing the bases of the subspaces

C.3. Distribution of various quantities resulting from least squares adjustment

1. Joint distribution of BLUE's and residuals
2. Distribution of the "weighted sum of residuals"
3. Expressions in $\Phi\tilde{x}$ and v having χ^2- or t-distribution
4. Expressions in $\tilde{x}$ and v having t-distribution

C.4. Confidence regions

1. Confidence intervals for one-dimensional Gauss variables
2. Application to the Gauss-Markoff model with known unit weight error
3. Studentization
4. Confidence regions for σ^2
5. Ellipsoidal confidence regions for sets of linear estimates

C.5. Tests of linear hypotheses

1. Linear hypotheses
2. Tests of variances
3. A simple example
4. A sophisticated example

D. SPECIAL TOPICS

D.1. Adjustment of Doppler observations

1. Transit system
2. Observing a difference in light travel time
3. Frequency shift
4. Technique of cycle counting
5. Parameters accounting for receiver imperfections
6. Transformations into an earth-fixed frame
7. Parameters accounting for orbit corrections
8. Linearization of the observation equations
9. Single station adjustment
10. Multi-station adjustment

D.2. Geodetic data bases

1. Storage media
2. Requirements for geodetic data bases
3. The data base of NGS

D.3. Cholesky's algorithm applied to normal equations of geodetic networks

1. Cholesky's algorithm for a general symmetric positive definite system
2. Partial reduction by Cholesky's algorithm
3. Geodetic normal equations
4. Geodetic interpretation of the partially Cholesky-reduced system
5. Problem of station ordering

D.4. One-dimensional cubic spline interpolation

1. Introduction
2. Parameterizing a cubic polynomial
3. Conditions at the inner nodes
4. Boundary conditions
5. Tridiagonal linear system
6. Modification of the periodic case
7. Interpolation of curves in the plane
8. Splines viewed as a vector space
9. The locality of splines

D.5. Two-dimensional spline interpolation

1. Introduction
2. Bicubic polynomials
3. Hermite bicubic interpolation
4. Bicubic splines

D.6. Geometry of exact spline interpolation

1. Formulation of the problem
2. Definition of splines
3. Existence and uniqueness of splines
4. Minimum properties of splines
5. Other examples
6. Prediction as a special case of spline interpolation
7. Noise-free collocation with trend parameters

D.7. Approximation with splines

1. Introduction
2. Approximation in one dimension
3. Basis splines with local support
4. Two dimensions

D. SPECIAL TOPICS

1. Adjustment of Doppler observations.

1.1. The Transit system.

This short subsection cannot replace any solid background information on the Transit Doppler system. For more information the reader may consult e.g. D.E. Wells (1974). (See the reference at the end of this chapter.)

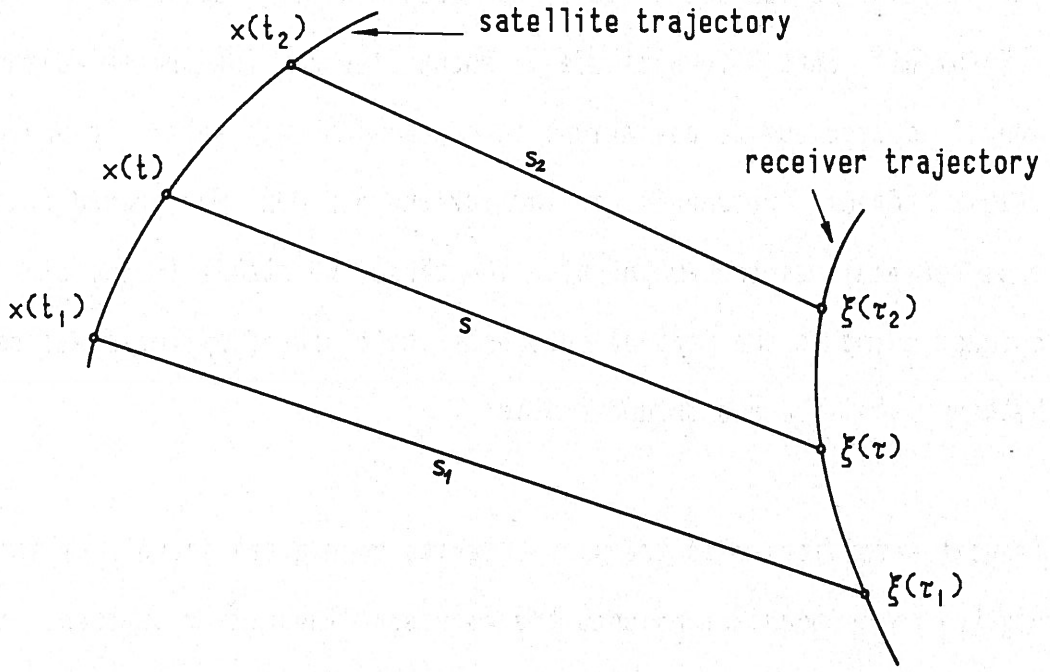
The Navy-Navigation satellite system uses 5 satellites in polar orbits (as of 1979/01/01). Satellite altitude is about 1100 km. The orbital planes are not equally spaced due to deviations in precession. Any satellite broadcasts at two stable harmonic frequencies of 150 MHz and 400 MHz. The ground station Doppler receiver measures the amount by which these two stable frequencies have been changed owing to the Doppler frequency shift caused by the relative velocity between satellite and ground station.

Transit satellites also transmit a series of digital signals by impressing digital phase modulations onto the carriers. The timing of these phase modulations is controlled by the satellite time standard so that they can be used as timing signals. The signals include parameters describing the satellite orbit. (Broadcast ephemerides.)

The two different frequencies allow a correction for errors introduced by the ionosphere. The influence of the troposphere is accounted for by a mathematical model. We do not discuss any relevant details.

There are 4 tracking stations in the United States responsible for monitoring the satellite orbits. About 20 cooperating stations distributed all over the globe contribute observations which are used to calculate and circulate (in a somewhat restricted way) improved a-posteriori ephemerides, called "precise ephemerides".

1.2. Observing a difference in light travel time.



Satellite and receiver are moving in inertial space. A geocentric coordinate system with non-rotating axes can be considered as inertial to a sufficient degree of accuracy. A signal emitted from the satellite at time t is received at the ground-station at time

$$\tau = t + \frac{s}{c}$$

$\tau(t)$...arrival time of signal emitted at time t

$s(t)$...distance traveled by light signal emitted at t and received at τ

c...speed of light.

We occasionally also denote (somewhat inconsistently):

$s(\tau)$...distance traveled by light signal received at time τ. The signal was emitted at time $t = \tau - \frac{s}{c}$.

Considering two signals emitted at t_1, t_2, we get

$$(\tau_2 - \tau_1) = (t_2 - t_1) + \frac{1}{c}(s_2 - s_1).$$

This equation immediately implies the following two relations between differentials:

$$d\tau = dt + \frac{1}{c}ds$$

This can be used in the following two ways:

$$\frac{d\tau}{dt} = 1 + \frac{1}{c}\frac{ds}{dt}$$

$$\frac{dt}{d\tau} = 1 - \frac{1}{c}\frac{ds}{d\tau}$$

1.3. The frequency shift.

Assume now that N oscillation periodes are emitted between t_1 and t_2. The satellite frequency is then

$$f_s = \frac{N}{t_2 - t_1}$$

The same N oscillation periods are received between τ_1 and τ_2. The averaged receiver frequency is then

$$f_r = \frac{N}{\tau_2 - \tau_1}$$

While f_s is constant, f_r is time varying. We see

$$\frac{f_r}{f_s} = \frac{t_2 - t_1}{\tau_2 - \tau_1}$$

In the limit

$$\frac{f_r}{f_s} = \frac{dt}{d\tau} = 1 - \frac{1}{c}\frac{ds}{d\tau} = \left[\frac{d\tau}{dt}\right]^{-1} = \left[1 + \frac{1}{c}\frac{ds}{dt}\right]^{-1}$$

Note that $\frac{ds}{dt}$, $\frac{ds}{d\tau}$ are not relative velocities of satellite and receiver. They are time rates of change of the light travel distance of a signal emitted at t and received at $\tau = t + \frac{s}{c}$.

1.4. Technique of cycle counting.

At the receiver a frequency

$$f_g = f_s + \Delta f$$

is generated. Superposition with the received signal allows to observe the slower cycles of the beat frequency

$$f_b = f_s + \Delta f - f_r$$

There are two basic alternatives

(1) The count is gated (i.e. initiated and terminated) by satellite time marks. Beats between $\tau_1 = t_1 + \frac{s_1}{c}$ and $\tau_2 = t_2 + \frac{s_2}{c}$ are counted.

(2) The count is gated by receiver time marks. Beats between τ_1 and τ_2 are counted, whereby $\tau_2 - \tau_1$ is a fixed time interval.

The most common receivers are Magnavox, JMR and Marconi. Magnavox and Marconi permit both modes. JMR uses (1), however a comparison between satellite time and receiver time takes place.

Corresponding to modes (2) and (1) we get two versions of the subsequent equation

$$D = \int_{\tau_1}^{\tau_2} f_b \, d\tau = \int_{\tau_1}^{\tau_2} (f_s + \Delta f) \, d\tau - \int_{\tau_1}^{\tau_2} f_r \, d\tau =$$

$$= (\tau_2 - \tau_1)(f_s + \Delta f) - \int_{\tau_1}^{\tau_2} f_s \left(1 - \frac{1}{c}\frac{ds}{d\tau}\right) d\tau =$$

$$\boxed{D = (\tau_2 - \tau_1)\,\Delta f + \frac{f_s}{c}(s_2 - s_1)} \qquad (2)$$

$$= \left(t_2 - t_1 + \frac{s_2 - s_1}{c}\right)\Delta f + \frac{f_s}{c}(s_2 - s_1) =$$

$$\boxed{D = (t_2 - t_1)\,\Delta f + \frac{f_g}{c}(s_2 - s_1)} \qquad (1)$$

The fly-by of a satellite allowing uninterrupted observation of its frequency is called a "pass". The duration of a pass is about 20 minutes. During a pass many counts may be observed. The typical duration of a count is between 4.6 seconds and 2 minutes. (This is $\tau_2 - \tau_1$; it is fixed in case of mode (2), and (slightly) variable in case of mode (1).)

1.5. Parameters accounting for receiver imperfections.

The parameters depend on the type of equipment. In particular they depend on the two alternative ways to gate the Doppler counts.

In case of mode (1), i.e. satellite gated counts, the socalled delay of time mark reception is important. Time marks are realized by phase modulations of the carrier frequency. They are processed by different circuits and delayed much stronger than the carrier signal itself. There is a constant delay communicated

by the manufacturer. It is about 500-1000 μs. Superimposed is a variable part of $\pm$30 μs. It is usually modelled by a pass internal parameter.

The delay of time mark reception causes the integration to start at $\tau_1 + \Delta$ and to end at $\tau_2 + \Delta$. Without delay, the starting and ending times would be τ_1 and τ_2. The effect on the observation equation (1) is visible only if arguments of s_1 and s_2 are revealed. We have

$$D = (t_2 + \Delta - t_1 - \Delta)\,\Delta f + \frac{f_g}{c}\,\{s(t_2 + \Delta) - s(t_1 + \Delta)\}$$

Not taking Δ into account would mean that wrong satellite positions are used, namely positions at t_1, t_2 instead of positions at $t_1 + \Delta$, $t_2 + \Delta$.

If a receiver clock is used to gate the counts, a receiver clock offset has a similar effect.

Another receiver error is given if the receivers reference frequency is imperfect. Instead of its nominal value $f_g = f_s + \Delta f$ the receiver frequency may be given by the following expression:

$$f_g + \delta f_g + \dot{\delta f_g}\tau$$

The error splits into a constant part and a drift rate. Both may be introduced as parameters into the observation equation. Reference frequency errors effect the above equations (1),(2).

Let us see what happens in case of equation (1), i.e. in case of satellite gated counts. Assume $\delta\dot{f}_g = 0$. The beat frequency is

$$f_b = f_g + \delta f_g - f_r$$

Since τ_1 and τ_2 are not falsified if gating is prompted by satellite time marks, we have

$$D = \int_{\tau_1}^{\tau_2} f_b \, d\tau$$

This leads us to equation (1) with f_g replaced by $f_g + \delta f_g$:

$$D = (t_2 - t_1)\,(f_g - f_s + \delta f_g) + \tfrac{1}{c}(f_g + \delta f_g)(s_2 - s_1) \qquad (1')$$

1.6. Transformation into an earth-fixed frame.

The rotation of the earth is described by the following rotation matrix

$$U(\tau) = \begin{bmatrix} \cos \omega t & -\sin \omega t & 0 \\ \sin \omega t & \cos \omega t & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

We write

$$\xi(\tau) = U(\tau)\bar{\xi}$$

$$x(t) = U(t)\,\bar{x}(t)$$

$\xi(\tau)$...station position in inertial frame

$\bar{\xi}$...station position in earth-fixed frame

$x(t)$.....satellite position in inertial frame

$\bar{x}(t)$...satellite position in earth-fixed frame.

We have

$$\begin{aligned} s(\tau) &= \|x(\tau - \tfrac{s}{c})\ \xi(\tau)\| = \\ &= \|U(\tau - \tfrac{s}{c})\ \bar{x}\ (\tau - \tfrac{s}{c}) - U(\tau)\ \bar{\xi}\| = \\ &= \|U(-\tfrac{s}{c})\ \bar{x}\ (\tau - \tfrac{s}{c}) - \bar{\xi}\| = \\ &= \|\bar{x}(\tau - \tfrac{s}{c}) - U(\tfrac{s}{c})\ \bar{\xi}\| \end{aligned}$$

(Mind $U(t_1 + t_2) = U(t_1)U(t_2)$, $U(-t) = U^{-1}(t) = U^{T}(t)$)

The above equation could also be written as

$$s(t) = \|\bar{x}(t) - U(\tfrac{s}{c})\ \bar{\xi}\|$$

(Recall the somewhat inconsistent notation $s(t) = s(\tau)$).

1.7. Parameters accounting for orbit corrections.

The following information on satellite orbits is available:

Broadcast ephemerides. They are transmitted by the satellites at 2 minutes intervals. They are based on a 36 hours period of observations at 4 tracking stations in the USA.

They are injected (uploaded, transmitted to satellites) at 12 hours intervals. Their errors may amount to 20-30 meters.

Precise ephemerides. They are post-computed by NWL-DMA. Positions at 1 min. intervals are communicated to users. The precise ephemerides are based on 20 tracking stations distributed over the globe. Their errors are estimated at 2-5 meters.

The orbit is represented in the computer by means of Chebyshev polynomials of degree 7 to 9. (One could imagine that spline interpolation would do somewhat better.) Corrective parameters to the orbits are assumed. Typical are 3 parameters accounting for deviation along track, across track and out of (orbital) plane. The parameters are not allowed to vary freely. They are viewed as pseudo observation with value zero and a prespecified variance. This corresponds to a mixed adjustment model, combining elements of conventional least squares adjustment with techniques of prediction and collocation.

1.8. Linearization of the observation equations.

It is apparent from the previous discussion that the Doppler observation equation may appear in various different shapes. In order to illustrate the principle, we take the nonlinear equation in the form (1') corresponding to satellite gated counts:

$$D = (t_2-t_1)(f_g - f_s + \delta f_g) + \frac{1}{c}(f_g + \delta f_g)[s(t_2+\Delta) - s(t_1+\Delta)]$$

It is seen that we assume a receiver delay Δ and a receiver frequency bias δf_g. We also assume orbital parameters. We denote the along-track, across-track and out-of-plane errors by a, b, c, respectively.

In agreement with section 1.6 we represent

$$s(t+\Delta) = \|\bar{x}(t+\Delta) - U(\tfrac{s}{c})\bar{\xi}\|$$

We assume that $\bar{X}(t)$ is the ephemeris satellite locus in the earth fixed frame. We also introduce unit vectors A(t), B(t), C(t) pointing into the tangential, (i.e along-track), the across-track and into the out-of-plane direction of the orbit. We introduce v(t), the scalar satellite velocity. The vectors A(t), B(t), C(t), and the scalar v(t) may be derived from the ephemeris. The satellite position at time t + Δ is then given by

$$\bar{x}(t+\Delta) = \bar{X}(t+\Delta) + A(t)a + B(t)b + C(t)c$$

Thus we have

$$s(t+\Delta) = \|\bar{X}(t+\Delta) + A(t)a + B(t)b + C(t)c - U(\tfrac{s}{c})\bar{\xi}\|$$

The quantities Δ, a, b, c, $\frac{s}{c}$ are small. Although s is unknown, a known approximate value may be used in $\frac{s}{c}$. We also represent the station coordinate vector as

$$\bar{\xi} = \bar{\xi}^{(0)} + \Delta\bar{\xi}.$$

Here $\bar{\xi}^{(0)}$ are approximate known values. Now comes the familiar linearization procedure. One gets to a sufficient degree of accuracy:

$$U(\tfrac{s}{c}) = \begin{bmatrix} 1 & -\omega\frac{s}{c} & 0 \\ \omega\frac{s}{c} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Hence

$$U(\tfrac{s}{c})\bar{\xi} = \begin{bmatrix} \bar{\xi}_1^{(0)} + \Delta\bar{\xi}_1 - \bar{\xi}_2^{(0)}\omega\frac{s}{c} \\ \bar{\xi}_2^{(0)} + \Delta\bar{\xi}_2 + \bar{\xi}_1^{(0)}\omega\frac{s}{c} \\ \bar{\xi}_3^{(0)} + \Delta\bar{\xi}_3 \end{bmatrix}$$

Introducing the direction cosines $\alpha_1(t)$, $\alpha_2(t)$, $\alpha_3(t)$ from $\bar{\xi}^{(0)}$ to $\bar{X}(t)$ one gets

$$s(t+\Delta) = \|\bar{X}(t) - \bar{\xi}^{(0)}\| +$$
$$+ \alpha_1(t)\ \{A_1(t)(v(t)\Delta+a) + B_1(t)b + C_1(t)c - \Delta\bar{\xi}_1 + \bar{\xi}_2^{(0)}\omega_c^s\}$$
$$+ \alpha_2(t)\ \{A_2(t)(v(t)\Delta+a) + B_2(t)b + C_2(t)c - \Delta\bar{\xi}_2 - \bar{\xi}_1^{(0)}\omega_c^s\}$$
$$+ \alpha_3(t)\ \{A_3(t)(v(t)\Delta+a) + B_3(t)b + C_3(t)c - \Delta\bar{\xi}_3\}$$

We abbreviate this as

$$s(t+\Delta) = \|\bar{X}(t) - \bar{\xi}^{(0)}\| + k_0(t) + k_\Delta(t)\Delta + k_a(t)a + k_b(t)b + k_c(t)c +$$
$$+ k_{\bar{\xi}_1}\ \Delta\bar{\xi}_1 + k_{\bar{\xi}_2}\ \Delta\bar{\xi}_2 + k_{\bar{\xi}_3}\ \Delta\bar{\xi}_3$$

Here we have put

$$k_0(t) = (\alpha_1(t)\bar{\xi}_2^{(0)} - \alpha_2(t)\bar{\xi}_1^{(0)})\ \omega_c^s$$

As mentioned above, in the small corrective term ω_c^s an approximate known value for s may be inserted.

We have thus linearized the worst term occurring in the Doppler observation equation.

We use the abbreviation

$$r(t) = \|\bar{X}(t) - \bar{\xi}^{(0)}\|$$

and obtain :

$$
\begin{aligned}
D = {} & (t_2 - t_1)(f_g - f_s) + \frac{f_g}{c}[r(t_2) - r(t_1) + k_0(t_2) - k_0(t_1)] \\
& + \{(t_2 - t_1) + \tfrac{1}{c}[r(t_2) - r(t_1)]\}\, \delta f_g \\
& + \frac{f_g}{c}\, \{(k_\Delta(t_2) - k_\Delta(t_1))\, \Delta \\
& \quad + (k_a(t_2) - k_a(t_1))\, a \\
& \quad + (k_b(t_2) - k_b(t_1))\, b \\
& \quad + (k_c(t_2) - k_c(t_1))\, c \\
& \quad + (k_{\bar{\xi}_1}(t_2) - k_{\bar{\xi}_1}(t_1))\, \Delta\bar{\xi}_1 \\
& \quad + (k_{\bar{\xi}_2}(t_2) - k_{\bar{\xi}_2}(t_1))\, \Delta\bar{\xi}_2 \\
& \quad + (k_{\bar{\xi}_3}(t_2) - k_{\bar{\xi}_3}(t_1))\, \Delta\bar{\xi}_3\}
\end{aligned}
$$

Remark: Because the effect of an along track error and a time delay are nearly equal, the two parameters a and Δ are practically inseparable. The normal equations are near singular unless a-priori pseudo observations for either a or Δ, or both, are faked. The pseudo observations have a value of zero, and a certain weight is given to them.

The typical structure of a Doppler observation equation is

$$
l + v_D = \varphi^T \Delta e + \psi^T \Delta o + \chi^T \Delta\bar{\xi}
$$

Here Δe refers to Δf_g, Δ, and, in general, to parameters resulting from receiver imperfections. Δo refers to a, b, c and, in general, to parameters resulting from orbit corrections. $\Delta\bar{\xi}$ refers to station coordinate increments. l stands for

the difference of D and the constant terms in the above equation. Newly introduced is v_D, the correction to D.

1.9. Single station adjustment.

At a single station a large number of satellite passes is observed. Any pass gives many observations, i.e. integrated Doppler counts D over short periods of time. (Typically, a pass lasts 20 minutes, while the periods for Doppler counts are 4.6 seconds to 2 minutes.)

A suitable adjustment model is one of phased observations. Any pass gives rise to a phase. Common parameters are the station coordinates. Pass-internal parameters are all others, i.e. parameters due to short periodic receiver imperfections and parameters for orbit corrections. Sometimes also meteorological parameters are included.

Precise ephemerides are almost obligatory in order to get meaningful results for single station positioning.

1.10. Multi-station adjustment.

If a satellite is co-observed during a pass from several stations, meaningful results can be obtained also on the basis of broadcast ephemerides. The reason is that orbit uncertainties affect all station locations in nearly the same way, causing, so to speak, a common translation and rotation of the co-observing stations. This mode of observations is sometimes called translocation. It may give reasonably good relative position vectors.

For a group of co-observing stations a phased approach may be used again. Common parameters are the station coordinates. Pass-internal parameters are the orbital corrections as before, the corrective parameters for the receivers, however, multiply. There are as many sets of receiver parameters as passes are observed by individual receivers.

A set of partially reduced normal equations is obtained for a group of co-observing stations. Other groups of stations may be treated similarly. There may be and should be overlaps between the groups. The partially reduced normals of the groups are added and a set of normals for all participating stations is finally obtained and solved.

References and bibliography.

BROWN, D.C. (1976): Doppler positioning by the short arc method. Paper presented at Satellite Doppler positioning International Symposium, Las Cruces, N.M, Oct.1976.

CHEN, J.Y. (1981): Geodetic Datum and Doppler Positioning. Dissertation, Technical University Graz.

JENKINS, R.E.; B.D. Merrit; D.R. Messent; J.R. Lucas (1979): Refinement of positioning software. (DOPPLR). Proceedings of 2nd International Symposium on Satellite Doppler positioning, Austin, Texas, Jan. 1979.

KOUBA, J. (1979): Improvements of Canadian geodetic Doppler programs. Proceedings of 2nd International Symposium on Satellite Doppler positioning, Austin, Texas, Jan. 1979.

SEEBER, G. (1980): Satelliten-Dopplerverfahren. In Pelzer (ed.) Geodaetische Netze in Landes- und Ingenieurvermessung. Konrad Wittwer, Stuttgart, p. 145-162. (In German).

WELLS, D.E. (1974): Doppler Satellite control. Technical Report No. 29, Dep. of Surveying Engineering, Univ. of New Brunswick, Fredericton, N.B., Canada.

2. Geodetic data bases.

2.1. Storage media.

The central memory of a computer allows very rapid processing of data. Searches through tables, matching of data, computations, can be done very fast once the data are in central memory. Data can be accessed there in fractions of 10^{-6} seconds. Central memory is expensive, hence its size is limited. Microcomputers or desktop computers may offer about $32 * 10^3$ to $128 * 10^3$ bytes of central memory. One byte contains 8 bits of information and is usually taken to encode one digit, one alphabetic character or one special symbol. On microcomputers we may have four times as much. On large computers, $4 * 10^6$ to $8 * 10^6$ bytes of central memory may be available.

None of these numbers is sufficient to store the data associated with a large network, nor is it desirable to do this. Data should be in central memory if they are being processed. Otherwise they reside on secondary storage such as disks or tapes. Tapes allow the sequential storage of data. A few multiples of 10^7 bytes may be put on a single reel of tape. Data on tapes can only be accessed sequentially. This makes tapes useless for many applications. However, sequential reading of data from a tape into central memory is fast, typically at a rate of a few 10^5 bytes per second. Tapes offer a very cheap way to store information.

Disks are most suitable for data bases where data must be accessed instantly. The amount of data which can be stored on the disks of one single disk drive is comparable to that of a tape. ($3 * 10^8$ bytes may be stored on some disk drives.)

One may imagine that data are stored on disks in sections. Sequential reading of a section is fast. However, locating the beginning of a new section may require one millisecond to 1/10 of a second.

It is not necessary to acquire a much deeper understanding of computer hardware. The few pieces of information given above shall serve to create a feeling for the difficulties encountered during the design of a data base.

2.2. Requirements for geodetic data bases.

The requirements depend on the type of application. We consider two applications, namely

(1) a data base for the automatically recorded observations of a field project

(2) a data base for a large national network

Let us first discuss the common features of these two types of data bases. In both cases, information on points and on observations is stored. These data are stored in a structured way. There will be indexes serving the rapid access of certain data, e.g. point coordinates, according to specified key values, e.g. point numbers. There will also be cross references (pointers) between the data, allowing e.g. quick access to all observations taken at a certain point. In a different situation, it may be desirable to find all observations whose forepoint (i.e. target point) is the station under consideration.

All situations requiring a linkage of the data must be foreseen when the data base is designed. If unexpected situations arise later on, there is frequently no other way than performing an exhausting (and expensive) search through the entire data base.

The most convenient type of linkage of data is that one of contiguous locations on segments of external storage. Recall that segments of disk storage can be loaded into central memory very quickly. One can imagine that such segments hold points which are located in a certain area. If one point is needed in that area, it is very likely that other points in the vicinity are needed, too. One can also imagine that all observations taken at a certain station are located in one physical segment of external storage.

Unfortunately, it is not possible to realize all types of linkages between data in the convenient way of placing them together. If observations for one standpoint are located together, then the observations having one and the same target point are scattered over various segments.

Let us now briefly elaborate on the differences between the two types of data bases. The purpose of data base (1) is the calculation of coordinates from redundant observations. Quick access to all types of data is necessary. On the other hand, the total amount of data is not large, usually less than 10^6 bytes. Such data can be accommodated on minidisks or diskettes of small desktop computers. Many computer manufacturers offer general purpose data base software that can be used. A self-made solution can be faster and more economical. Such a

system is described in Bartelme, Hofmann-Wellenhof, and Meissl (1981).

Not all data stored on a data base of type (2), i.e. one for a large national network, must be accessed instantly. It may be convenient to get point coordinates quickly. It is also useful to get quick information about the connectivity of points and about the availability of data in a certain region. Hence only a small portion of data should be stored on disk. The mass of observational data, of station descriptions and of historical data can be put onto magnetic tape. It will be sufficient to answer requests for such data within a few days. In the subsequent section we briefly describe the data base that has been established at the U.S. National Geodetic Survey.

2.3. The data base of NGS. [cf. Schwarz (1975)].

As part of the activity related to the readjustment of the North American network, NGS is placing all its horizontal positions, observations and descriptions into a data base. The storage of data is station-oriented. For any station, the following types of information are stored:

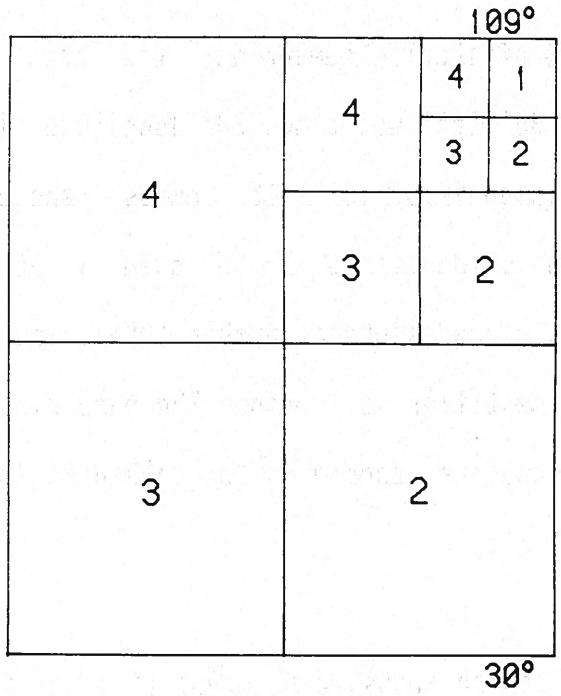
Position, and, if needed, elevation
Terrestrial observations taken at this station
Descriptions
Doppler observations
Astronomical observations
Cross references
Historical data

For one station, all these data occupy contiguous portions on magnetic tape.

For the purpose of station numbering, the area of the U.S. is subdivided into quadrangles of 30' latitude times 30' longitude. In areas of dense control, a further subdivision into 15' * 15' quads, and even into 7.5' * 7.5' quads, may occur. Any quad is identified by a quad identifier (QID). Within a quad, any station carries a quad sequence number (QSN). Thus the concatenation of QID and QSN uniquely identifies a station. The numbering system automatically implies a grouping of stations according to geographical regions. The details are explained by fig. 2.1.

The data base can be accessed by means of a query language. This language offers menues to the user, allowing him to quickly access the index part of the data base, and to submit batch jobs for detailed investigations. Figures 2.2, 2.3 illustrate the separation into an index part and a part with the detailed records.

Indexing of stations



φ λ

QID = 30 109 1 1 1

30′ indicator

15′ indicator

$7\frac{1}{2}$′ indicator

QSN = quad station number
(within 30′, 15′, or $7\frac{1}{2}$′ quads)

Fig. 2.1.

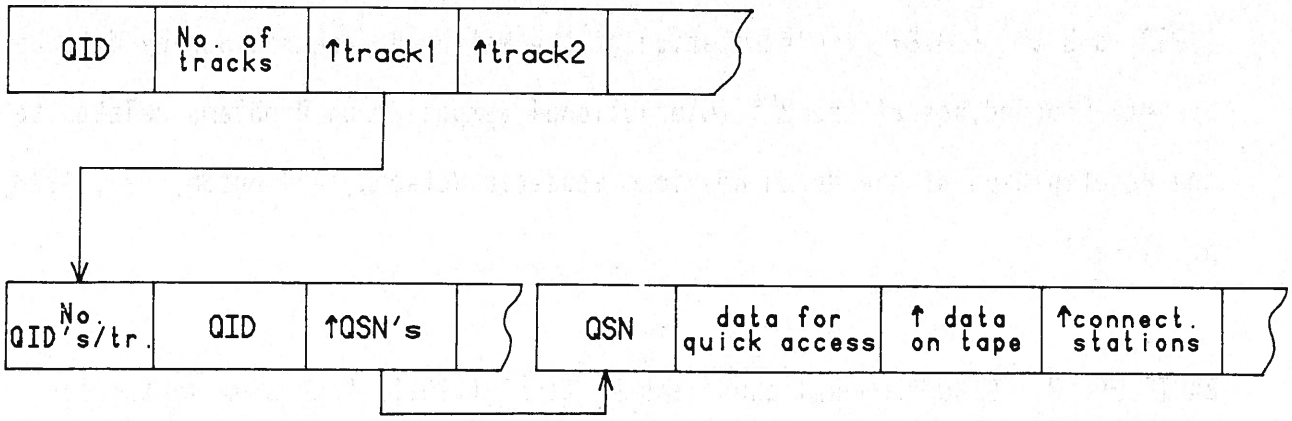
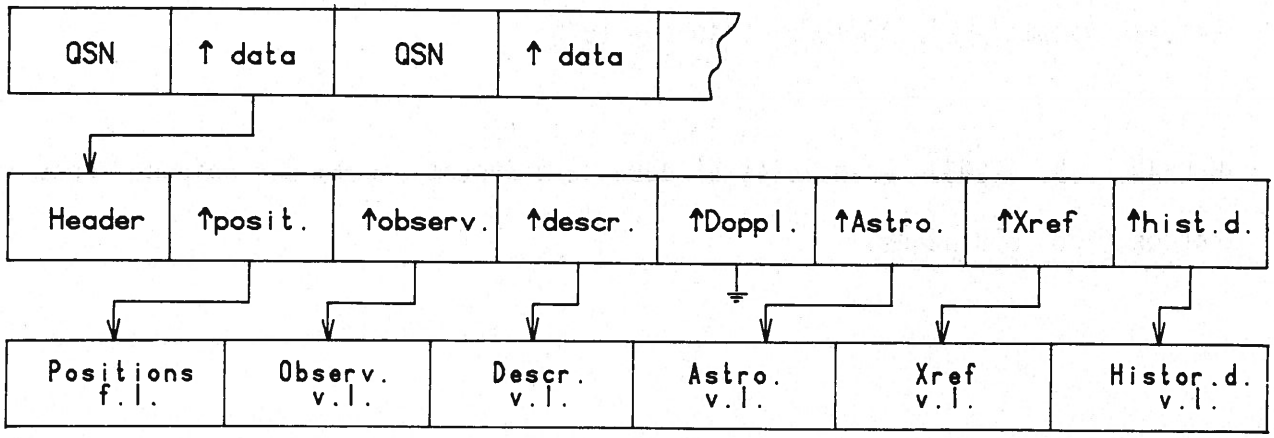


Fig. 2.2. Data stored on disk.



f.l. ... fixed length
v.l. ... variable length

Fig. 2.3. Data stored on tape.

REFERENCES.

ALGER, D.E.Jr. (1978): Implementation of the National Geodetic Survey data base system. Proceedings of the 2nd International Symposium on Problems related to the Redefinition of the North American Geodetic Network, Arlington, Va., 1978, pp.247-257.

BARTELME, N., B.Hofmann-Wellenhof and P.Meissl (1981): A program system for interactive processing of automatically recorded geodetic measurement data. (English translation of an article written in German). ZfV, Jg. 107, Heft 4, pp.144-154.

SCHWARZ, C.R. (1975): The geodetic data base of NGS. Paper presented at XVI General Assembly of IAG/IUGG, Grenoble 1975.

ULLMAN, J.D. (1980): Principles of Data Base Systems. Computer Science Press, Potomac, Md. 379 pages.

3. Cholesky's algorithm applied to the normal equations of geodetic networks.

3.1 Cholesky's algorithm for a general symmetric positive definite system.

Suppose that the system is written in matrix form as

$$A\,x = b$$

Cholesky's algorithm relies on a decomposition of the positive definite matrix A as

$$A = R^T R$$

where R is an upper triangular matrix. During the first or so-called "triangular decomposition phase" of the algorithm, the system is, in effect, multiplied by $(R^T)^{-1}$. The result is the following triangular system:

$$R\,x = s$$

with

$$s = (R^T)^{-1} b$$

During the second or "back-substitution phase" of Cholesky's algorithm, the triangular system is solved for x recursively, starting with the last component of x and proceeding to the first.

The details of Cholesky's algorithm can be best described by switching to indices notation. The original system then reads

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, \ldots, n$$

The triangularized system is

$$\sum_{j=i}^{n} r_{ij}x_j = s_i, \quad i = 1, \ldots, n$$

which is calculated from the original system by

$$\left.\begin{aligned} r_{ii} &= \Big(a_{ii} - \sum_{k=1}^{i-1} r_{ki}^2 \Big)^{1/2} \\ r_{ij} &= \Big(a_{ij} - \sum_{k=1}^{i-1} r_{ki}r_{kj} \Big)/r_{ii}, \quad j = i+1, \ldots, n \\ s_i &= \Big(b_i - \sum_{k=1}^{i-1} r_{ki}s_k \Big)/r_{ii} \end{aligned}\right\} \quad i = 1, \ldots, n$$

During the back substitution phase, the triangular system is solved by

$$x_i = \Big(s_i - \sum_{j=i+1}^{n} r_{ij}x_j \Big)/r_{ii}, \quad i = n, \ldots, 1$$

3.2 Partial reduction by Cholesky's algorithm.

Split the original system as

$$A_{11}x_1 + A_{12}x_2 = b_1$$
$$A_{21}x_1 + A_{22}x_2 = b_2$$

Split R accordingly:

$$R = \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix}$$

From the equation $R^T R = A$ we deduce the following identities

$$R_{11}^T R_{11} = A_{11}$$

$$R_{11}^T R_{12} = A_{12}$$

$$R_{12}^T R_{12} + R_{22}^T R_{22} = A_{22}$$

Multiply the first set of the original normals by $(R_{11}^T)^{-1}$ and then eliminate the unknowns x_1 from the second set by subtracting proper multiples of the equations of the first set, the multiplying matrix factor being $A_{21}R_{11}^{-1}=R_{12}^T$. The resulting system is

$$R_{11}x_1 + R_{12}x_2 = s_1$$

$$A_{22}^{(p)}x_2 = b_2^{(p)}$$

The second set of these equations is called the "partially reduced" set of normal equations. Explicit and equivalent expressions for the quantities involved are

$$A_{22}^{(p)} = A_{22} - R_{12}^T R_{12} = R_{22}^T R_{22} = A_{22} - A_{21}A_{11}^{-1}A_{12}$$

$$b_2^{(p)} = b_2 - R_{12}^T s_1 = b_2 - A_{21}A_{11}^{-1}b_1$$

These expressions are easily checked by the identities exhibited above. The last expression in any of the two lines reveals that the reduced normal equations do not depend in any way on the peculiarities of Cholesky's algorithm. In fact, any method of elimination that removes the unknowns x_1 from the second set by subtracting proper multiples of the first set must uniquely arrive at the partially reduced normals exhibited above.

In indices notation, the partial Cholesky reduction is

$$\left.\begin{aligned} r_{ii} &= \Big(a_{ii} - \sum_{k=1}^{i-1} r_{ki}^2 \Big)^{1/2} \\ r_{ij} &= \Big(a_{ij} - \sum_{k=1}^{i-1} r_{ki}r_{kj} \Big)/r_{ii}, \quad j = i+1, \ldots, n \\ s_i &= \Big(b_i - \sum_{k=1}^{i-1} r_{ki}s_k \Big)/r_{ii} \end{aligned}\right\} \; i = 1, \ldots, p$$

$$\left.\begin{aligned} a_{ij}^{(p)} &= a_{ij} - \sum_{k=1}^{p} r_{ki}r_{kj}, \quad j = i+1, \ldots, n \\ b_i^{(p)} &= b_i - \sum_{k=1}^{p} r_{ki}s_k \end{aligned}\right\} \; i = p+1, \ldots, n$$

Cholesky's algorithm can be organized in many different ways. The programs used by the NGS, writen by R.H. Hanson and based on earlier work of Poder and

Tscherning (1973) at the Danish Geodetic Institut, execute Cholesky's algorithm in the following manner;

```
FOR j=1 TO n+1
  FOR i=1 TO MIN(n,j)
    SUM = 0
    FOR k=1 TO MIN(p,i-1)
      SUM = SUM + A(k,i)*A(k,j)
    NEXT k
    A(i,j) = A(i,j) - SUM
    IF (i≤MIN(j-1,p)) A(i,j) = A(i,j)/A(i,i)
  NEXT i
  IF (j≤p) A(j,j) = SQRT(A(j,j))
NEXT j
```

In this algorithm, the A(i,j) are place holders. They denote storage locations for a number of quantities. In detail,

* the original coefficients a_{ij} are stored at A(i,j);
  the original coefficients b_i are stored at A(i,n+1);
* the $a_{ij}^{(p)}$ are stored at A(i,j), the $b_i^{(p)}$ at A(i,n+1);
* the r_{ij} are stored at A(i,j), the s_i at A(i,n+1).

It should be stressed that the above algorithm is still a simplification of the actual NGS programs. First, these programs make use of a more complicated data

structure which allows storage and retrieval of coefficients A(i,j) columnwise to and from mass storage (disks). Second, the programs allow for exploiting the sparsity of the normal equations to some extent. The normals have many zero coefficients. If the equations are ordered in a sensitive way, many of the zeros are retained throughout the reduction. This results in a great saving of computer storage und computation time. The NGS programs store only a section of each column, excluding coefficients that will never become nonzero during the execution of the algorithm. We shall come back to the problem of ordering in section 3.5.

Remark: We briefly mention another way to execute Cholesky's algorithm, which amounts to a series of partial reductions for p proceeding from 1 to n. In this fashion Cholesky's algorithm becomes very similar to Gauss' algorithm. Denote $a_{ij}^{(0)}=a_{ij}$ and $b_i^{(0)}=b_i$. We then have

$$
\left.
\begin{aligned}
r_{pp} &= \left(a_{pp}^{(p-1)} \right)^{1/2} \\
r_{pj} &= a_{pj}^{(p-1)}/r_{pp}, \quad j = p+1, \ldots, n \\
s_p &= b_p^{(p-1)}/r_{pp} \\
a_{ij}^{(p)} &= a_{ij}^{(p-1)} - r_{pi}r_{pj} \qquad \left. \begin{aligned} & i = p+1, \ldots, n \\ & j = i, \ldots, n \end{aligned} \right\} \\
b_i^{(p)} &= b_i^{(p-1)} - r_{pi}s_i
\end{aligned}
\right\} \quad p = 1, \ldots, n
$$

If the algorithm stops at any p<n, a partially reduced system results. It adds insight into the problem of equation ordering, discussed later in this section, that any equation is modified by either dividing it by the square root of the diagonal element or by subtracting proper multiples of preceding equations.

Remark. Common to all versions of Cholesky's algorithm is that they operate only on the portion above and including the main diagonal of the matrix A, as well as on the right-hand side. Hence only the upper triangular portion of the matrix A needs to be stored in computer memory. Substantially more storage is saved if the sparse structure of A is exploited, which is typical for matrices associated with network problems.

3.3 Geodetic normal equations.

Our system of normal equations results from a geodetic ground control network. Adjustment is done on a spheroidal rotational ellipsoid. We assume that the reader is familiar with the principles of network adjustment. Our outline will mainly serve to point out peculiarities and to specify the terminology and notation used in the sequel.

The network will be adjusted by variation of parameters. The parameters, or unknowns, are the ellipsoidal coordinates of the stations (points,nodes). Any station has two parameters, namely ellipsoidal latitude and longitude. The so-called orientation unknowns of direction bundles will be eliminated before the normal equations are assembled and will not appear in the final set of

equations.

Approximate coordinates must be known a priori. Denote these coordinates by the vector $p^{(0)}$. The observations ℓ, comprising distances, azimuths, bundles of directions, and Doppler posiotions, will not fit the approximate coordinates. There will be discrepancies $\Delta\ell$, i.e. only the set of observations $\ell-\Delta\ell$ will fit the approximate coordinates. An adjustment applies corrections v to the observations, so that they become the corrected observations $\ell+v$. It also applies shifts Δp to the approximate coordinates so that they become the adjusted coordinates $p=p^{(0)}+\Delta p$. The functional relation between the corrected observations and the adjusted coordinates is (after elimination of the orientation unknowns) in linearized form written as:

$$\Delta\ell + v = B\,\Delta p$$

Weights are prescribed for the individual observations. They are arranged along the diagonal of the weight matrix P which has zero off-diagonal coefficients. Gauss' minimum principles, i.e.,

$$v^T P v = \text{Minimum}$$

is used to uniquely determine v and Δp satisfying the side constraints $\Delta\ell+v=B\Delta p$. The extremum problem leads to the normal equations

$$B^T P B\Delta p = B^T P\Delta\ell$$

which for brevity are written as

$$A\,x = b$$

Note that the unknowns x are actually small shifts leading from the approximate coordinates to the adjusted coordinates.

An important feature of geodetic network adjustment is the local nature of the observations. Any observation involves only a small number of stations which are located close together. For distance and direction observations, direct visibility between two stations must be given. This limits the spacings between stations connected by such a line of vision to 30 km or less in most cases. The normal equation matrix will have only nonzero off-diagonal elements $a_{ij} \neq 0$, if i,j refer either to the two coordinates of one station or to coordinates of two stations connected by a measurement. Such a connection is established either by a direction, a distance, or an azimuth between the two stations, or is due to the preelimination of the orientation unknowns in case of a directional co-observation of the two stations from a third station. The Doppler position observations refer to the two coordinates of one station and will not cause any a_{ij}, $i \neq j$, to be nonzero. While the network covers a large portion of a continent and extends over several thousands of kilometers, there will only be nonzero coefficients a_{ij} if the involved stations are not farther apart than 60 km (in most cases).

Remark. In the literature on numerical linear algebra it is frequently argued that formation and solution of a normal equation system is not a good procedure for doing a least squares adjustment. Instead one should go along with the observation equation system, subjecting it to orthogonalization, singular value decomposition, or other procedures. The argument is based on the condition number of a matrix. The condition number of the normal equation matrix is inferior to that of the observation equations. This is certainly true. On the other hand, it has been proven that storage requirement and computational labor is much less for a geodetic network if it is adjusted by the direct solution of a normal equation system as compared to any other procedure. Refer to the discussion in Avila et al. (1978,p.16). Singular value decomposition or orthogonalization appears to be very efficient for moderately large linear systems that are very ill-conditioned. In the case of very large sparse geodetic network systems which are not extremely ill-conditioned, storage requirement and computational labor are the decisive criteria for selecting a solution method. The observation equation matrix for the U.S. network is of size 3,000,000 x 350,000. To my knowledge no technique is known that preserves sparsity during orthogonalization or singular value decomposition as efficiently as that method which applies direct elimination to the normal equation system, as will be shown later in this chapter.

3.4 Geodetic interpretation of the partial Cholesky-reduced system.

The geodetic meaning of the quantities appearing in the system that has undergone a partial reduction by Cholesky's method is perhaps best understood in

terms of a parameter transformation. The original normals are written as

$$A_{11}x_1 + A_{12}x_2 = b_1$$
$$A_{21}x_1 + A_{22}x_2 = b_2$$

and consider a parameter transformation which changes x_1 into y_1 leaving x_2 unchanged:

$$y_1 = R_{11}x_1 + R_{12}x_2$$
$$x_2 = \qquad\qquad x_2$$

The inverse transformation is

$$x_1 = R_{11}^{-1}y_1 - R_{11}^{-1}R_{12}x_2$$
$$x_2 = \qquad\qquad\qquad x_2$$

The normal equations for the new parameters are

$$y_1 \qquad\qquad = s_1$$
$$A_{22}^{(p)}x_2 = b_2^{(p)}$$

If we substitute for y_1, we get

$$R_{11}x_1 + R_{12}x_2 = s_1$$
$$A_{22}^{(p)}x_2 = b_2^{(p)}$$

This is precisely what we get after partial Cholesky reduction. We see that hidden behind these equations is the system of normal equations involving y_1, x_2. This system completely decomposes into two separate systems for y_1 and x_2. It follows that the adjusted values for y_1, x_2 will be uncorrelated. The covariance matrix for x_2 will be

$$\Sigma(x_2) = (A_{22}^{(p)})^{-1}$$

Let us go back to the original normal equations:

$$A\,x = b$$

If a certain subset of the components of x are forced to fixed values, which amounts to fixing the corresponding coordinates at the values $p^{(0)}+x$, then the normal equations for the remaining unknowns are obtained as follows: Noting that any equation belongs to a certain coordinate, disregard all equations belonging to the fixed components. In the remaining equations, insert the prescribed values for the x's to be fixed, and move these terms toward the right. The desired system results. Note that the same procedure may be applied to the partially reduced Cholesky system

$$R_{11}x_1 + R_{12}x_2 = s_1$$
$$A_{22}^{(p)}x_2 = b_2^{(p)}$$

provided that the fixing is restricted to coordinates out of set x_2. This observation allows us to give the coefficients r_{ij}, s_i, $a_{ij}^{(p)}$, $b_i^{(p)}$ the following geodetic interpretation.

* $a_{ii}^{(p)}$, $i>p$, is the reciprocal of the variance of coordinate i, povided that the coordinates k, $p<k\leq n$, $k\neq i$ are fixed, while the coordinates k, $1\leq k\leq p$, as well as coordinate i itself, are allowed to vary freely.

* $-a_{ij}^{(p)}/a_{ii}^{(p)}$, $i,j>p$, $i\neq j$ is the shift, with respect to the adjusted position, suffered by coordinate i if coordinate j is displaced by one unit from the adjusted position, and if coordinates k, $p<k\leq n$, $k\neq i,j$ are fixed to their adjusted position, while coordinates k, $1\leq k\leq p$ as well as coordinate i itself, are allowed to vary freely.

* $b_i^{(p)}/a_{ii}^{(p)}$, $i>p$ is the shift, with respect to the approximate position, suffered by coordinate i if coordinates k, $p<k\leq n$, $k\neq i$ are fixed to their approximate positions, while coordinates k, $1\leq k\leq p$, as well as coordinate i itself, are allowed to vary freely.

* r_{ii}, $i\leq p$, is the standard deviation of coordinate i, if coordinates k, $i<k\leq n$, are fixed, while coordinates k, $1\leq k\leq i$ are allowed to vary freely.

* $-r_{ij}/r_{ii}$, $i\leq p$, $j>i$ is the shift, with respect to the adjusted position, suffered by coordinate i, provided that coordinate j is displaced by one unit from its adjusted position, that coordinates k, $i<k\leq n$, $k\neq j$ are fixed to their adjusted positions while coordinates k, $1\leq k\leq i$ are allowed to vary freely.

* s_i/r_{ii}, $i\leq p$ is the shift, with respect to the approximate position, suffered by coordinate i, provided that coordinate k, $i<k\leq n$ are fixed to their approximate positions, while coordinates k, $1\leq k\leq i$ can vary freely.

The last three statements require an additional argument because coordinates k, $k \leq p$ are also held fixed, while earlier we said that fixing is restricted to the second set of unknowns, i.e., those with $k > p$.

The three last statements should be clear if we set $i=p$, because then only coordinates $k>p$ are fixed. On the other hand, the r_{ij}'s are no longer subject to any change, as p moves on from i to higher values. Hence the argument also applies for $i<p$.

Remark. (Elastostatic interpretation of normal equations before and after partial reduction.) To the structural engineer the normal equations $Ax=b$ appear as equilibrium equations of an elastic system. The matrix A is called the stiffness matrix, x are coordinate shifts of the nodes, and b are external forces acting at the nodes. The coefficients of the stiffness matrix have the following physical meaning: Suppose that the system is in equilibrium with $x=0$, $b=0$. Displace coordinate j by one unit from its equilibrium position, keeping all other coordinates fixed to their equilibrium position. An elastic force will then be acting on coordinate i. This force is precisely a_{ij}. This holds also for $i=j$. The partially reduced normals $A_{22}^{(p)}x_2=b_2^{(p)}$ refer to a so-called statically reduced system. $A_{22}^{(p)}$ is still a stiffness matrix. $a_{ij}^{(p)}$, $p<i$, $j \leq n$ is the force acting in coordinate i when coordinate j is displaced by one unit from its equilibrium position, when coordinates k, $p<k \leq n$ are fixed, while coordinates k, $1 \leq k \leq n$ are allowed to adjust freely. The right-hand coefficients $b_i^{(p)}$ have the meaning of forces. The original $b_i=b_i^{(0)}$ are nodal forces due to inconsistencies

in the network. As nodes are freed during elimination, different forces $b_i^{(p)}$ must be applied to the remaining nodes such that the equilibrium position of the remaining nodes remains the same. The forces of the eliminated nodes must be transported to the uneliminated ones. Occasionally it is also advantageous to consider external forces. If the vector b is chosen as the j-th column of the unit matrix, the solution x of the system becomes the j-th column f_{*j} of the inverse F of the stiffness matrix A. Hence f_{ij} is the shift of coordinate i if a unit force is applied to coordinate j. Thereby it is assumed that prior to application of the unit force a free equilibrium state had been reached. In particular, f_{ii} is the shift of coordinate i with respect to its adjusted position, if (after adjustment) a unit force is applied to coordinate i. A more lucid interpretation of the variance f_{ii} of the adjusted coordinate i can hardly be given. The elastostatic interpretation is thus somewhat simpler and of great physical significance. I personally prefer to think in terms of elastostatics, where the $a_{ij}^{(p)}$, $b_i^{(p)}$ themselves have a most simple interpretation, whereas in geodetic reasoning the ratios $a_{ij}^{(p)}/a_{ii}^{(p)}$, $b_i^{(p)}/a_{ii}^{(p)}$ are most easily understood. However, since this publication is addressed to the geodesist, elastostatic language will very rarely be used in the sequel. For further details the reader is referred to Rubinstein and Rosen (1970).

Remark. (On the near vanishing of row sums.) Another property of geodetic normal equations is concerned with the row sums

$$\sum_{j=1}^{n} a_{ij}^{(p)}$$

of the original as well as the partially reduced normals. If i is a coordinate whose station - call it P - is involved only in relative measurements, i.e. in measurements other than absolute positioning by Doppler, then the above row sum nearly vanishes for any p. The row sum vanishes precisely if the network is plane. On the ellipsoid it vanishes only approximately. The proof, for the plane network, goes back to the observational equations $Bx=\Delta l+v$. All observational equations involving station P can be thought of as being formulated in terms of differences of coordinate increments. This implies that the row sums pertaining to the station P vanish. The property of station P's vanishing row sums carries over from the observational equations matrix B to the original normal equation matrix $A=B^TPB$. Note that station P's normal equations can be formed by considering only the observations that involve this station. If station P is involved in a Doppler measurement, the row sum of equation i will not vanish, even if the network is plane. However, since the Doppler observations have weights much smaller than those of the relativ measurements (directions, distances, azimuths), the row sum will be appreciably smaller than the larger coefficients in the i-th row of A. Hence, we conclude that all row sums of the normals are small. The remark at the end of section 3.2 tells us that Cholesky's algorithm is a successions of subtractions of multiples of rows from others. Hence the property of near vanishing of row sums is retained throughout reduction and carries over to the partially reduced normal equation matrix $A_{22}^{(p)}$.

3.5 Problem of station ordering.

Coordinate i is associated with row and column i of the normal equations. Ordering the coordinates in a different way leads to a system of normal equations with rows and columns simultaneously permuted, i.e., with diagonal elements permuted rows and columns arranged accordingly. Mathematically, the two systems are equivalent, numerically they are not. Widely recognized in recent literature are the great differences in storage requirement and computation time that result from different orderings and when algorithms are used that take into account the sparseness of A.

In geodetic networks, nonzero off-diagonal elements result from observations between stations rather than between coordinates. The problem of ordering the unknowns becomes a problem of ordering the stations. The two coordinates of one station will always be placed together.

We will refrain from giving a thorough discussion of ordering schemes currently in fashion. We shall briefly review three ordering strategies. The first serves as an introduction to the problem, the other two will be relevant to the readjustment of the U.S. network.

3.5.1 Ordering for small bandwidth.

A supposed geodetic network is depicted in figure 3.1. The solid lines indicate directions observed at both end points. Additional distances and azimuths (measured along some of the solid lines) as well as some Doppler positional

observations may be available. Recall that two stations are connected by nonzero off-diagonal coefficients in the normal equations if there is a direction-, distance-, or azimuth-observation between these two points, or if the two points are directionally co-observed from a third station. In this way, station 1 is connected to stations 2,3,5,6,8,9. Station 8 is connected to 1,2,5,6,9,10,12, 13,14,17,18,19. For any station i we can specify the highest numbered station s_i connected to station i. Thus $s_1=9$, $s_8=19$. We may calculate the number

$$w = 2 * MAX(s_i-i+1)$$

which is called the bandwidth of the system. The factor 2 has been introduced to account for the fact that we have two coordinates per station. In our above example we would have $w=2(s_8-8+1)=24$.

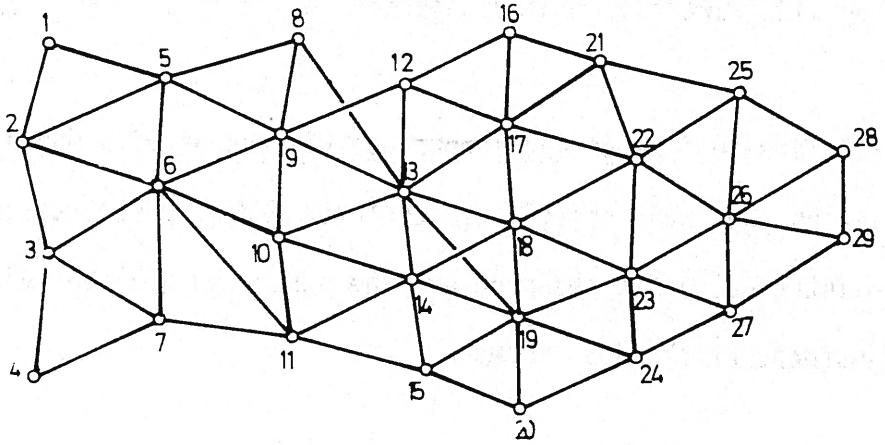


Figure 3.1. - Sample network.

It turns out that the normal equation matrix A will have nonzero coefficients restricted to a band of width w as indicated in figure 3.2. Note that w counts

only lines of coefficients above and including the main diagonal. The coefficients below the main diagonal are never used.

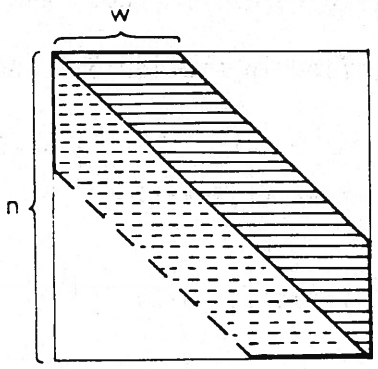


Figure 3.2. - Banded normal equations.

In general, the band will not be completely filled with nonzero coefficients of A. It will also contain some zeroes. The important thing to note, however, is that nonzero coefficients r_{ij}, $a_{ij}^{(p)}$, arising during (partial) Cholesky reduction, are also confined within the band. Some of these will appear at places where A also had nonzero coefficients, and others will take the place of original zeroes. The latter ones are called "fill-in" coefficients.

The proof that fill-in is confined to the band is most easily derived from the next to the last remark in section 3.2. There we saw that any row of any of the Cholesky reduced states results by subtracting multiples of preceding rows from it (and by dividing the row by a factor, if $i \leq p$). However, preceding rows k, $k<i$ can never have nonzero coefficients to the right of the rightmost eligible position for a nonzero coefficient of a row i.

A consequence of the banded structure of A is that any one of the inner products, i.e., the sums of products appearing in Cholesky's algorithm, will have, at most, w-1 nonzero terms. In fact, the first version of the full Cholesky algorithm specified in section 3.1 can be respecified as follows:

$$\left.\begin{aligned} r_{ii} &= \Big(a_{ii} - \sum_{k=MAX(1,i-w+1)}^{i-1} r_{ki}^2 \Big)^{1/2} \\ r_{ij} &= \Big(a_{ij} - \sum_{k=MAX(1,j-w+1)}^{i-1} r_{ki} r_{kj} \Big) / r_{ii} \\ & \quad j = i+1, \ \ldots \ , MIN(n, i+w-1) \\ s_i &= \Big(b_i - \sum_{k=MAX(1,i-w+1)}^{i-1} r_{ki} s_k \Big) / r_{ii} \end{aligned}\right\} \quad i=1, \ \ldots \ ,n$$

and

$$x_i = \Big(s_i - \sum_{j=i+1}^{MIN(n,i+w-1)} r_{ij} x_j \Big) / r_{ii} \qquad i=n, \ \ldots \ ,1$$

On the one hand, a computer program for this algorithm would be more complicated; on the other hand, for w<<n, it would be much faster. It would save much storage if the coefficients within the band were stored in a compacted way, for example, as the columns of an array of size w*n.

A different numbering of the stations would generally result in a different bandwidth w. One could try to minimize w over all possible permutations; however, this is not economical. There are computer algorithms that find near optimal orderings in a short time. Frequently, a good ordering is found by inspection. If a network is elongated, as in the example above, then numbering along the lines that cross the network at the shorter distances often leads to a good ordering. I believe the ordering specified in the figure 3.2 is near optimal.

3.5.2 Ordering for small profile.

As pointed out in section 3.5.1, the saving of computer time and storage comes from the reduced number of product accumulations in Cholesky's algorithm. It is plausible, therefore, that methods have been designed which aim at minimizing the number of product accumulations in the first place rather than doing this indirectly by minimizing the bandwidth. One such method, which is still a compromise between simplicity and efficiency, is ordering for small profile.

The profile of a symmetric matrix includes all elements of a column that are located between the topmost nonzero element and the main diagonal, inclusively. Hence, an element a_{ij}, $i \leq j$ is within the profile, if there is an element $a_{kj} \neq 0$ for a certain $k \leq i$. A typical profile is shown in figure 3.3.

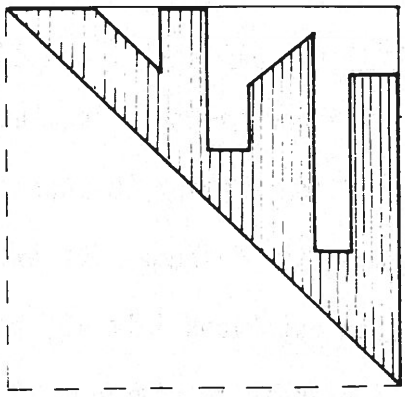


Figure 3.3. - Profiled normal equations.

The profile may include zero coefficients. Again it is important to note that fill-in is restricted to the profile. The proof relies on a similar argument as previously given for the banded structure. Subtracting a multiple of a row from a subsequent row will never cause any nonzero entry outside the profile. The

Cholesky factorization $A=R^TR$ will result in a matrix R which has nonzero coefficients only within the profile. R, being upper triangular, will have zeroes below the main diagonal, whereas A will have coefficients implied by the symmetry there.

NGS computer programs which are currently being used to adjust moderately small networks (up to about 2,500 stations) rely on ordering for a small profile. The ordering algorithm, designed and described by Snay (1976), is heuristic and does not yield a minimal profile in the strict sence. It will, however, establish a fairly small profile in a short time. As will be clear later on, the algorithm will also contribute to the adjustment of the entire U.S. network.

3.5.3 Identifying nonzero coefficients for a certain reduction state.

Before we proceed to still another ordering technique, we pause briefly and reflect on the problem of identifying the nonzero coefficients of A associated with a certain reduction state. Assume, for example, that the partial Cholesky reduction has "eliminated" stations 1 to 12, also marked by black circles in figure 3.4. White circles indicate stations 13 to 29 that participate in the partially reduced system $A_{22}^{(p)}x_2=b_2^{(p)}$. The network is the same as that one in section 3.5, except that the station numbering now conforms with a changed sequence of elimination steps. From section 3.4, dealing with the geodetic interpretation of a Cholesky-reduced system, we infer that the pattern of nonzero coefficients after partial Cholesky reduction up to station p=12, inclusively is shown in figure 3.5.

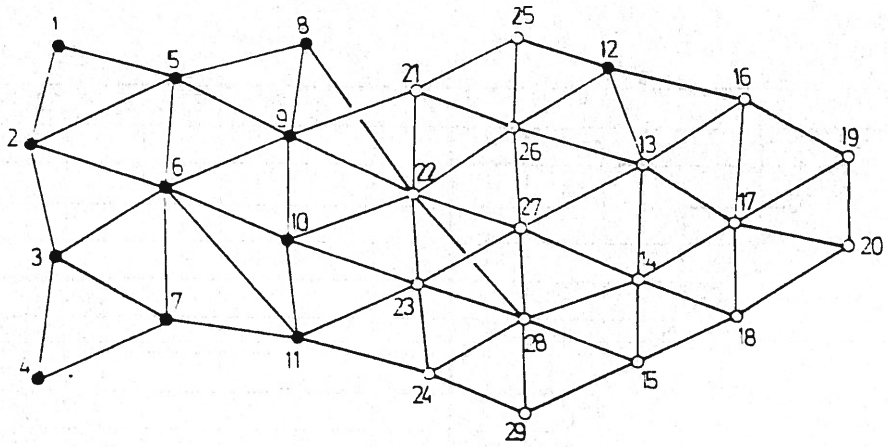


Figure 3.4. - Sample network with stations 1 through 12 eliminated from normal equations.

The numbering of rows and columns of the matrix A in figure 3.5 refers to nodes rather than to coordinates. Hence the individual entries represent actually 2x2 matrices. Heavily shaded entries represent nonzero elements of the original normals. Lightly shaded areas indicate the fill-in which occurs during partial Cholesky reduction up to and including station p=12. Let us give the appropriate argument for a few entries.

* Entry (14,21). The shading indicates fill-in. Why are nodes 14 and 21 connected at this time? According to section 3.4 (cf., the explanation of the expression $-a_{ij}^{(p)}/a_{ii}^{(p)}$ there), we assume that nodes 1 to 12(=p) are free, as well as node 14. We assume the other nodes fixed to their adjusted position, except for node 21, which is displaced from its adjusted position. The displacement of node 21 will cause the direction bundles at the neigboring nodes 25,26 to rotate. As a consequence, the free node 12 will move away from its

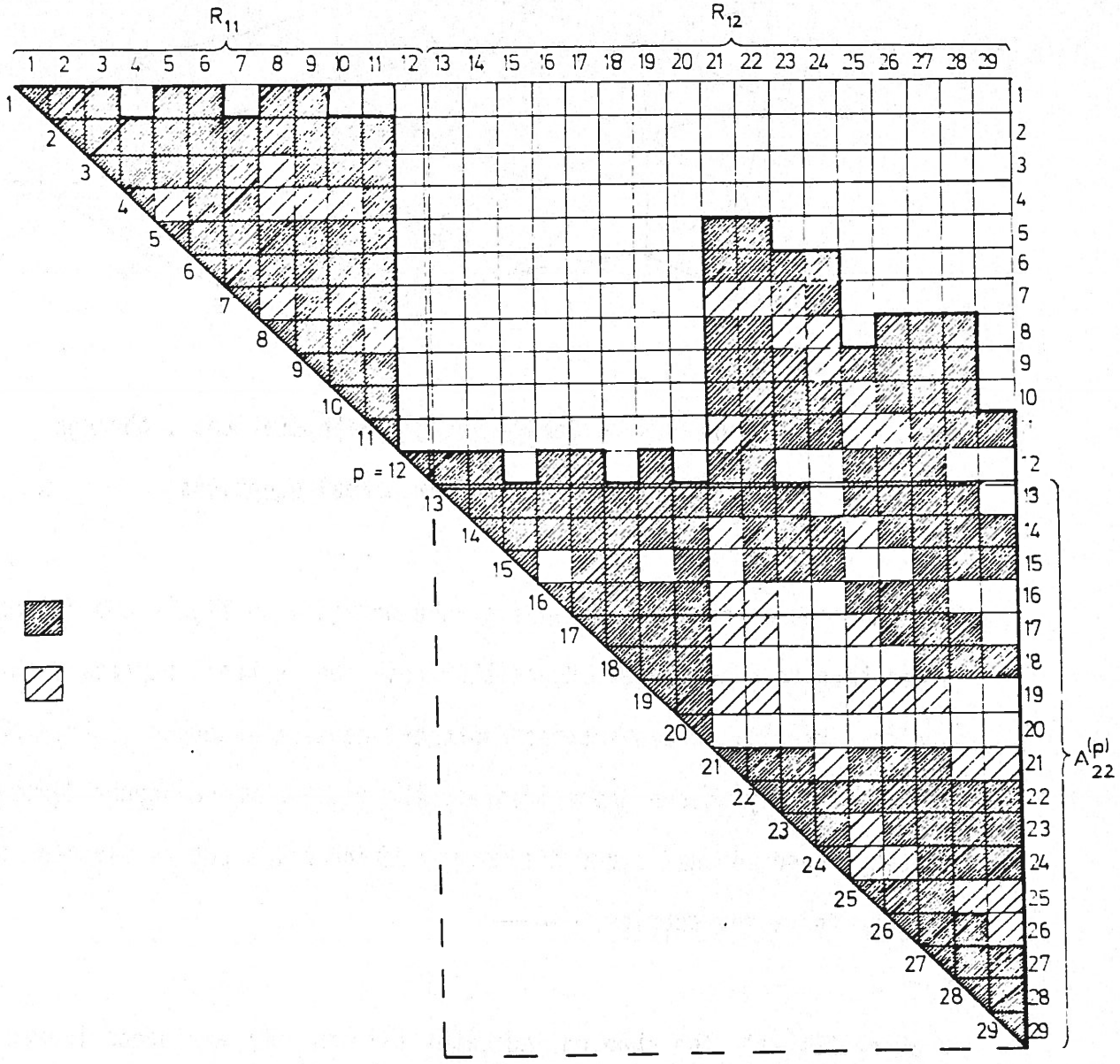


Figure 3.5. - Structure of normal equation when stations 1 through 12 are eliminated.

adjusted position, causing in turn the bundle in 13 to be displaced rotationally. This bundle finally will displace station 14. Hence $a_{14,21}^{(p)}$ will be nonzero, as was to be shown. (The possibility that the resulting movement of 14

is the zero movement is neglected here, as it is in all treatises of sparse matrices.)

* Entry (3,8). The shading indicates fill-in again. This time we refer to the rule for $-r_{ij}/r_{ii}$ given in section 3.4. We pretend that only nodes 1,2,3 have been eliminated, i.e., we temporarily assume p=3. We further assume nodes 4 to 29 fixed to their adjusted positions, except for node 8 which is displaced. This causes the bundle in 5 to deviate from its adjusted position, which in turn displaces nodes 1,2. The displacement of 1 and 2 will finally displace node 3. Hence $r_{3,8}$ must be nonzero, in general.

* Entry (10,13). We may put p=10. Displacing node 13 causes movements of the bundles connected to node 13. No movement takes place to the left of the barrier formed by the double line of nodes 21 to 29. Hence the coefficient must be zero. In fact, coefficients (i,j), $i \leq 11$, $12 \leq j \leq 20$, must be zero. We see that a barrier of a double line of nodes crossing the network can effectively keep down the fill-in. This observation leads us to the ordering scheme considered in the next subsection.

3.5.4 Nested dissection.

We have just seen that by appropriately ordering the stations we may establish barriers which divide the network into parts such that the interior stations of one part will never become connected to interior stations of another part. The numerical analyst George (1973) fully exploited this idea. He calls his ordering

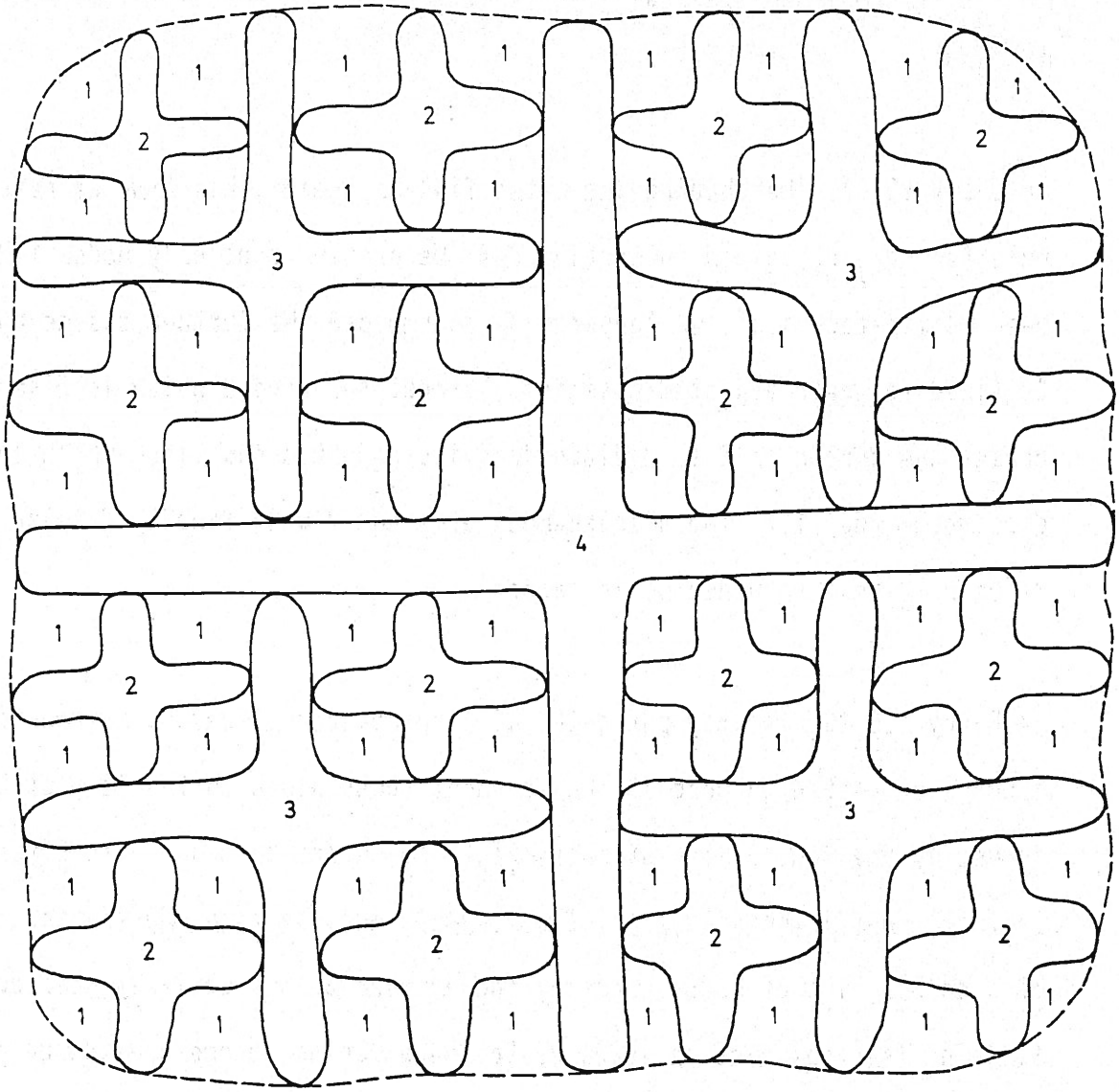


Figure 3.6. - Nested dissection

scheme "nested dissection". As we shall see later, this is anticipated to some extent by what is known among geodesists as "Helmert blocking".

Figure 3.6 exemplifies the idea of nested dissection. The individual stations are not shown here. Instead, we see subsets of stations carrying labels 1 to 4.

We imagine that these labels are attached to all nodes of a particular subset. Nodes carrying label 1 are eliminated first. The sequence in which this is done is not of much importance as long as the number of stations in one connected subset is small. Should this number be larger, we may imagine that an ordering for small profile is done in each individual subset. At the next step we eliminate nodes labeled 2, then 3, and finally 4.

Let us now take a look at the connections a certain node labeled i may encounter to nodes that come later in the ordering sequence. Such nodes carry either the label i or a label j>i. Connections to label i nodes are possible only if the other node is in the same connected label i subset. This is true, because all other label i subsets are separated by barrier subsets of higher labels. Connections of a label i node to nodes of higher labels are only possible if the higher label nodes are located at a barrier surrounding the subset of node i.

Any node will be connected to only a few nodes that come later in the ordering sequence. This is particularly true at the lower levels. It follows that matrix A will be quite sparse, although the pattern of zeroes is now rather complicated.

In order to see the power of nested dissection, we imagine a fairly homogeneous network of n stations covering a region which is shaped somewhat like a square. George (1973) shows that the number of nonzero coefficients (original A plus fill-in) is bounded by

$$const_1 \; n \log n$$

If, in contrast to this, we subject the network to ordering for small bandwidth, we can bound the nonzeroes only by

$$const_2 \; n^{3/2}$$

Also ordering for small profile could not achieve anything much better. Assuming an efficient storage scheme, the storage requirement grows roughly proportional to the number of nonzeroes. However, the factor of proportionality is different from method to method. Nested dissection, in particular, has a more complicated pattern of zeroes that necessitates the storage of additional pointers to keep track of the nonzero elements.

Despite the different proportionality factors and also the difference between $const_1$ and $const_2$ in the above formulas, it becomes clear that asymptotically, i.e., as n grows on and on, nested dissection is superior. In fact, as $n \to \infty$, the ratio of storage requirement for nested dissection and bandwidth tends to zero as const log n / $n^{1/2}$. In this context it is interesting to note that no ordering scheme can improve upon nested dissection asymtotically by more than a constant factor.

We have argued that the number of nonzeroes is directly related to storage requirement. It is also indirectly related to the amount of computational labor. Let us take a look at the number of product accumulations necessary for the

triangular decomposition of A. As it turns out, these product accumulations account for most of the computation time needed to solve the normal equations by Cholesky's method. George (1973) shows that this number is bounded by

$$\text{const}_3\ n^{3/2}$$

if nested dissection is done. Bandwidth ordering, on the other hand, requires

$$\text{const}_4\ n^2$$

for a homogeneous network of the type mentioned. Again the asymptotic superiority of nested dissection becomes evident.

We conclude this subsection with a few remarks.

Remark. Asymptotic superiority of a method does not necessarily mean superiority for moderately small networks. As already indicated, the exploitation of a complicated pattern of zeroes can cause an overhead of storage and computation time. In addition to nonzero coefficients, overhead storage is needed for addressing information which must be stored and for storing a more complicated program.

Remark. Faced with a given network, the subdivision of nodes into categories of different labels is not always immediate. The network will not always be rectangularly shaped, and it will not always be possible to identify a number of

first level sets equal to a power of 4. In pratice, it will be necessary to compromise. Occasionally, the connected subsets of stations of the same label will deviate in number and shape from the ideal case shown in figure 3.6.

Remark. To avoid pitfalls, one must be sure that the barriers dividing the network, as indicated in figure 3.6, are virtually impenetrable. For the types of networks considered, i.e., those involving bundles of directions, distances, azimuths, and absolute positions, the following rule applies. From and to a node of label i there may be lines of vision only to and from; (1) nodes of an adjacent lower label set, (2) nodes of label i which are in the same label i subset, (3) nodes of higher labeled adjacent sets. Otherwise one will try to keep the barriers as thin as possible. Roughly one will arrive at barrier sets composed of double rows of points, as already encountered in the example of figure 3.4. However, there will be exceptions, particularly in the presence of very long lines of vision.

3.5.5 Helmert blocking.

Let us briefly review the basic idea of Helmert blocking for the small network shown in figure 3.4. We reproduce the network in figure 3.7. The dashed line separates two blocks. The nodes marked by simple circles are interior to the relevant block. The nodes marked by double circles are junction nodes, forming a barrier between the two blocks. The normal equations are assembled separately for each block:

$$\text{Block 1:} \begin{bmatrix} A_{11} & B_{13} \\ B_{31} & B_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_1 \\ b_3 \end{bmatrix}$$

$$\text{Block 2:} \begin{bmatrix} A_{22} & C_{23} \\ C_{32} & C_{33} \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_2 \\ c_3 \end{bmatrix}$$

Here x_1, x_2 denote the coordinates of stations interior to blocks 1, 2, and x_3 denotes the junction station coordinates. Observations between interior stations of block 1 contribute to block 1 equations. Observations between stations interior to block 1 and junction stations also contribute to it. A similar statement can be made for block 2. Observations between junction stations contribute to the block in which the instrument was positioned. In this context note that the dashed line attributes uniquely a block to any station.

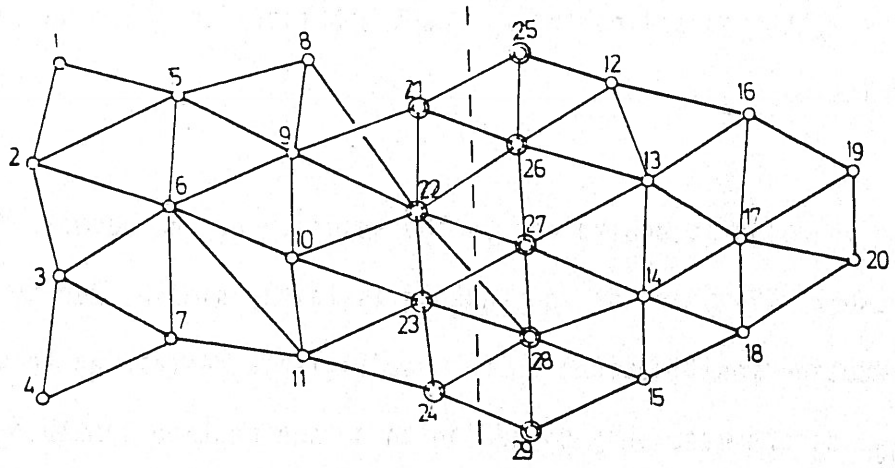


Figure 3.7. - Sample network decomposed into two Helmert blocks.

Adding the two systems of normal equations would result in the conventional normals for the entire network. However, elimination starts for each block

separately. The unknowns x_1, x_2 are eliminated from the two systems by partial Cholesky reduction:

$$\text{Block 1:} \begin{bmatrix} R_{11} & R_{13} \\ & B_{33}^{(p)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} s_1 \\ b_3^{(p)} \end{bmatrix}$$

$$\text{Block 2:} \begin{bmatrix} Q_{22} & Q_{23} \\ & C_{33}^{(q)} \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} t_2 \\ c_3^{(q)} \end{bmatrix}$$

The two partially reduced systems for the unknowns x_3 are taken out and added:

$$\left\{ B_{33}^{(p)} + C_{33}^{(q)} \right\} x_3 = \left\{ b_3^{(p)} + c_3^{(q)} \right\}$$

This system is solved for x_3. Back substitution into the two above systems yields x_1, x_2.

The solution is equivalent to the solution of the normals for the entire network. The proof of equivalence is fairly simple. During the partial Cholesky reduction modifications to the coefficients pertaining to x_3, i.e., to B_{33}, b_3, C_{33}, c_3 are made only by adding to or subtracting something from them. Because the quantities added or subtracted are same as they would be if the entire system were partially reduced, it is irrelevant whether the equations for x_3 are added before or after the partial reduction.

A larger network will be partioned into more than two blocks. A hierarchy of blocks can be established that is similar to the nested dissection procedure. In

fact, one can view figure 3.6 as a Helmert blocking scheme. There are as many first-level blocks as there are sets labeled 1, i.e., the number is 64. The normal equations are formed for each first level block separately. Higher labeled nodes situated in adjacent barrier sets take part in the normal equations as junction nodes. The dashed lines separating the first-level blocks have to be imagined as bisecting the barrier sets between the sets labeled 1. All observations must be used in forming the normals, and any observation must be used only once.

The interior nodes are eliminated from the first-level blocks. The partially reduced normals for the junction nodes of four adjacent earlier first-level blocks are added to form the normals of a second-level block. In such a second-level block the nodes labeled 2 now play the role of interior nodes. The junction nodes have labels higher than two. There are 16 second-level blocks. The number of blocks has been reduced by the factor of one-fourth. The interior nodes are eliminated from the second-level blocks, etc. Finally at the fourth and last level we deal with a system for the coordinates of these stations. Back substitution cascades down through the previous levels and successively yields the coordinates of the lower labeled stations.

What is the difference now between Helmert blocking as described here and nested dissection? Not much. In fact, Helmert blocking is slightly more sophisticated because the normals are not fully formed before reduction starts. Instead, the normals are formed separately for each first-level block. After partial reduction at any level, normals of a number of blocks are merged by adding them.

These operation have to be viewed as part of the formation of the normals rather than part of the solution process. George (1973) pointed out that substantial savings are realized in computer time and storage associated with the peculiar way of combining four i-level blocks to form one i+1-level block. Although Helmert blocking has been widely used by geodesists, I do not know of any reference where it has been done by nested dissection. Instead, in most cases, only two levels have been considered. Helmert or his geodetic followers did not appear to anticipate George's logarithmic law.

The U.S. network will be adjusted by the Helmert blocking technique. Partial reduction at the intermediate block level, as well as the reduction of the last level system will be done by Cholesky's method. First-level blocks will be ordered individually for small profile. Higher level blocks will also be ordered to some extent, but ordering becomes less significant as the systems tend to become less and less sparse.

References

AVILA, J., Malloy, B., and Tomlin, J., 1978: Use of the ILLIAC IV for the readjustment of the North American Datum. T.M. 5732, Institute for Advanced Computation, Sunnyvale, Calif.94085, 87 pp.

FORSYTHE, G., and Moler, C.B., 1967: Computer Solution of Linear Algebraic Systems. Prentice-Hall, 148 pp.

GEORGE, A., 1973: Nested dissection of a regular finite element mesh. SIAM Journal of Numerical Analysis, 10(2), 345-363.

MEISSL, P., 1980: A Priori Prediction of Roundoff Error Accumulation in the Solution of a Super-Large Geodetic Normal Equation System. NGS, NOS, Rockville, Md. 20852, 128 pp.

PODER, K., and Tscherning, C.C., 1973: Cholesky's method on a computer. Internal Report No. 8 of the Danish Geodetic Institute, Copenhagen, 22 pp.

RUBINSTEIN, M., and Rosen, R., 1970: Error analysis in structural computation. Journal of the Franklin Institute, 290, 37-48.

SNAY, R.A., 1976: Reducing the profile of sparse symmetric matrices. NOAA Tech. Memorandum NOS NGS-4, Nat. Oceanic and Atmospheric Administration, Rockville, Md.,24 pp.(Available from NTIS, Sprinfield, Va., accession No. PB 258476.)

4. One dimensional cubic spline interpolation.

4.1. Introduction.

We start with the familiar problem of interpolation. We are given a finite number of abscissas $x_0, x_1, \ldots, x_n$, and a corresponding set of function values $y_0, y_1, \ldots, y_n$. The problem is to specify a smooth curve interpolating these data. Without loss of generality we may assume

$$x_0 < x_1 < \ldots < x_n$$

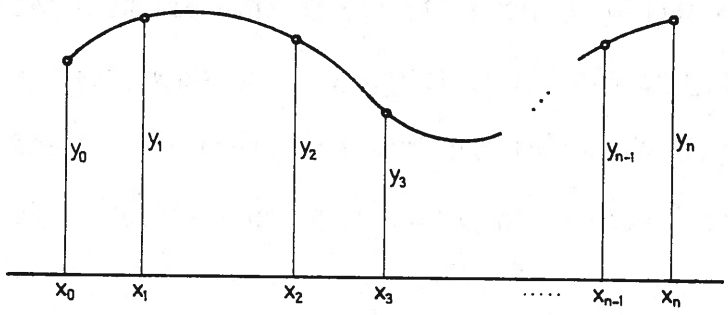


Fig. 4.1.

The abscissas need not be equally spaced.

The classical solution to this problem is polynomial interpolation. The polynomials are either algebraic, i.e.

$$p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

or trigonometric

$$t(x) = \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \ldots$$

The polynomials are fairly easy to set up by means of methods specified by Newton, Lagrange and others. They are also quickly evaluated, in particular if they are algebraic.

A disadvantage, often noted in practice, and strongly supported by theory, is a tendency toward instabilities if the number of specified nodes increases. If the polynomial is algebraic, its degree equals the number of locations minus 1. A polynomial of high degree, which is forced to interpolate a set of specified data, has a tendency to oscillate. It can even be shown, that a sequence of interpolating polynomials resulting from a sequence of more and more dense data, will diverge in "most cases". In practice, polynomials of degree exceeding 5 are rarely used.

In a number of fundamental papers, I.I. Schoenberg proposed a different interpolation scheme which is based on the use of piece-wise polynomials. Instead of using a single polynomial interpolating all data $y_0, y_1, \ldots, y_n$ at $x_0, x_1, \ldots, x_n$, Schoenberg proposes to use different polynomials in the successive intervals

$$[x_i, x_{i+1}], \quad i = 0, \ldots, n-1$$

The polynomials are of low degree. Cubic polynomials have proved themselves to be very useful. At the interval boundaries, i.e. at the abscissas

$x_0, x_1, \ldots, x_n$, the polynomials are forced to attain the prescribed values $y_0, y_1, \ldots, y_n$.

This obvious interpolation requirement makes the resulting function s(x) continuous everywhere. In addition one postulates continuity of a number of derivatives s'(x), s''(x), ... at $x_1, \ldots, x_{n-1}$.

The interpolating function s(x) is called a spline functior.. In the case of cubic polynomials we call it a cubic spline. Cubic splines are required to be continuous together with their first and second derivatives. Cubic spline curves are very smooth. Their name "splines" is derived from elastic rules used by Dutch shipbuilders as an aid to draw smooth curves which are constrained to pass through prespecified points.

4.2. Parameterizing a cubic polynomial.

A cubic polynomial is represented as

$$p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$$

A slightly different representation is obtained if the origin is shifted to $x = x_0$:

$$p(x) = a_0 + a_1 (x - x_0) + a_2 (x - x_0)^2 + a_3 (x - x_0)^3$$

In any case, the polynomial has 4 coefficients which serve to parameterize it,

i.e. to identify it. Other ways to parameterize the polynomial are frequently more useful. Suppose that at two locations x_a, x_b, corresponding function values $y_a = p(x_a)$, $y_b = p(x_b)$ are prescribed. Assume that also the derivatives $y'_a = p'(x_a)$, $y'_b = p'(x_b)$ are prescribed. Our aim is to parameterize the polynomial in terms of these 4 values.

We imagine p(x) to be written as

$$p(x) = a_0 + a_1(x - x_a) + a_2(x - x_a)^2 + a_3(x - x_a)^3$$

We require p(x) to attain the prescribed values at $x = x_a$, x_b. Four equations result:

$$y_a = a_0$$

$$y'_a = a_1$$

$$y_b = a_0 + a_1(x_b - x_a) + a_2(x_b - x_a)^2 + a_3(x_b - x_a)^3$$

$$y'_b = a_1 + 2a_2(x_b - x_a) + 3a_3(x_b - x_a)^2$$

This system of equations expresses the new parameters $y_a, \ldots, y'_b$ in terms of the old ones $a_0, \ldots, a_3$. In order to obtain the expressions of $a_0, \ldots, a_3$ in terms of $y_a, \ldots, y'_b$, we solve the linear system for $a_0, \ldots, a_3$. The result is

$$a_0 = y_a$$

$$a_1 = y'_a$$

$$a_2 = \frac{3(y_b - y_a) - (2y'_a + y'_b)(x_b - x_a)}{(x_b - x_a)^2}$$

$$a_3 = \frac{-2(y_b - y_a) + (y'_a + y'_b)(x_b - x_a)}{(x_b - x_a)^3}$$

For later use we express the second derivatives of the polynomial at $x = x_a, x_b$ in terms of the new parameters $y_a, \ldots, y'_b$. Using $p''(x) = 2a_2 + 6a_3(x - x_a)$, we obtain:

$$p''(x_a) = -\frac{4y'_a}{x_b - x_a} - \frac{2y'_b}{x_b - x_a} + \frac{6(y_b - y_a)}{(x_b - x_a)^2} = y''_a$$

$$p''(x_b) = \frac{2y'_a}{x_b - x_a} + \frac{4y'_b}{x_b - x_a} - \frac{6(y_b - y_a)}{(x_b - x_a)^2} = y''_b$$

Remark: We mention in passing that another useful parameterization of a cubic polynomial relies on the parameters y_a, y_b, y''_a, y''_b. Such a parameterization is occasionally used in the literature. However we prefer the one described earlier.

4.3. Condition at the inner nodes.

Our interpolating cubic spline is now represented as

$$s(x) = p_{i,i+1}(x), \quad x_i \leq x \leq x_{i+1}, \quad i = 0, \ldots, n-1$$

The cubic polynomial $p_{i,i+1}(x)$ refers to the interval $[x_i, x_{i+1}]$. The formulas at the previous section apply if we identify

$$x_i = x_a; \quad x_{i+1} = x_b; \quad y_i = y_a; \quad y_{i+1} = y_b$$

Let the polynomial $p_{i,i+1}(x)$ be represented in terms of coefficients as

$$p_{i,i+1}(x) = a_0^{i,i+1} + a_1^{i,i+1}(x - x_i) +$$

$$+ a_2^{i,i+1}(x - x_i)^2 + a_3^{i,i+1}(x - x_i)^3$$

One could parameterize the whole spline s(x) by the set of parameters $a_0^{i,i+1}, \ldots, a_3^{i,i+1}$, $i = 0, \ldots, n-1$. However this set of parameters is redundant. The interpolation requirement $s(x_i) = y_i$, $i = 0, \ldots, n$, and the continuity requirements for s(x), s'(x), s''(x) at x_i, $i = 1, \ldots, n-1$, impose conditions on the parameters $a_j^{i,i+1}$, $i = 0, \ldots, n-1$, $j = 0, \ldots, 3$.

The polynomial $p_{i,i+1}(x)$ is alternatively parameterized by y_i, y_{i+1}, y'_i, y'_{i+1}. This implies a parameterization of the spline s(x) in terms of y_i, y'_i, $i=0,\ldots,n$. This parameterization automatically guarantees that

(1) s(x) interpolates the values y_i at x_i, $i = 0, \ldots, n$

(2) s(x) has a continuous derivative.

A third condition, namely that

(3) s(x) has continuous second derivatives

must be enforced. It result in the set of n-1 equations:

$$p''_{i-1,i}(x_i) = p''_{i,i+1}(x_i), \qquad i = 1, \ldots, n-1$$

Using the expressions for the second derivatives given in the previous section we obtain after division by 2:

$$\frac{1}{x_i - x_{i-1}} y'_{i-1} + 2\left(\frac{1}{x_i - x_{i-1}} + \frac{1}{x_{i+1} - x_i}\right) y'_i + \frac{1}{x_{i+1} - x_i} y'_{i+1} =$$

$$= \frac{3(y_i - y_{i-1})}{(x_i - x_{i-1})^2} + \frac{3(y_{i+1} - y_i)}{(x_{i+1} - x_i)^2}, \qquad i = 1, \ldots, n-1$$

These are n-1 equations for n+1 unknowns $y'_0, \ldots, y'_n$. Two equations are missing. They will be specified in the next section.

4.4. Boundary conditions.

In the previous section we have seen that a cubic spline is not uniquely determined by the interpolation requirement. Two additional conditions must be imposed. They are usually formulated as boundary conditions. We consider the following 3 types of boundary conditions:

(1) Constrained spline. The values of y'_0 and y'_n are prescribed.

(2) Free spline. The conditions are

$$y''_0 = y''_n = 0$$

Explicitely (using the last equations of section 4.3):

$$\frac{2}{x_1 - x_0} y'_0 + \frac{1}{x_1 - x_0} y'_1 = \frac{3(y_1 - y_0)}{(x_1 - x_0)^2}$$

$$\frac{1}{x_n - x_{n-1}} y'_{n-1} + \frac{2}{x_n - x_{n-1}} y'_n = \frac{3(y_n - y_{n-1})}{(x_n - x_{n-1})^2}$$

(3) The periodic spline. It relies on the periodicity of the data, i.e.

$$y_0 = y_n$$

Continuity of the first derivative implies the assumption

$$y'_0 = y'_n$$

Continuity of the second derivative must be enforced also at x_0, x_n respectively:

$$p''_{0,1}(x_0) = p''_{n-1,n}(x_n)$$

i.e.

$$\frac{1}{x_n - x_{n-1}} y'_{n-1} + 2\left(\frac{1}{x_n - x_{n-1}} + \frac{1}{x_1 - x_0}\right) y'_0 + \frac{1}{x_1 - x_0} y'_1 =$$

$$= \frac{3(y_n - y_{n-1})}{(x_n - x_{n-1})^2} + \frac{3(y_1 - y_0)}{(x_1 - x_0)^2}$$

Other types of boundary conditions are possible, but will not be considered here.

4.5. Tridiagonal linear system.

Boundary conditions (1) and (2) listed in the previous section lead , together with the continuity conditions of s''(x) at the inner nodes, to a linear system which is tridiagonal

$$b_0 y'_0 + c_0 y'_1 = d_0$$

$$a_i y'_{i-1} + b_i y'_i + c_i y'_{i+1} = d_i, \quad i = 1, \ldots, n-1$$

$$a_n y'_{n-1} + b_n y'_n = d_n$$

The system looks like

$$\begin{bmatrix} b_0 & c_0 & & & & & & \\ a_1 & b_1 & c_1 & & & & & \\ & a_2 & b_2 & c_2 & & & & \\ & & a_3 & b_3 & c_3 & & & \\ & & & \ddots & \ddots & \ddots & & \\ & & & & \ddots & \ddots & \ddots & \\ & & & & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & & & & a_n & b_n \end{bmatrix} \begin{bmatrix} y'_0 \\ y'_1 \\ y'_2 \\ y'_3 \\ \vdots \\ \\ \\ y'_{n-1} \\ y'_n \end{bmatrix} = \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ d_3 \\ \vdots \\ \\ \\ d_{n-1} \\ d_n \end{bmatrix}$$

By straightforward Gauss elimination we transform the subdiagonal elements a_i into zeros. This is called the triangular decomposition phase:

Put

$$\bar{b}_0 = b_0 \,, \quad \bar{d}_0 = d_0$$

Compute:

$$\bar{b}_k = b_k - \frac{a_k}{\bar{b}_{k-1}} c_{k-1} \,, \quad \bar{d}_k = d_k - \frac{a_k}{\bar{b}_{k-1}} \bar{d}_{k-1} \,, \quad k = 1, \ldots, n$$

The triangularized system is

$$\begin{bmatrix} \bar{b}_0 & c_0 & & & & & \\ & \bar{b}_1 & c_1 & & & & \\ & & \bar{b}_2 & c_2 & & & \\ & & & \ddots & \ddots & & \\ & & & & \ddots & \ddots & \\ & & & & & \bar{b}_{n-1} & c_{n-1} \\ & & & & & & \bar{b}_n \end{bmatrix} \begin{bmatrix} y'_0 \\ y'_1 \\ y'_2 \\ \vdots \\ \vdots \\ y'_{n-1} \\ y'_n \end{bmatrix} = \begin{bmatrix} \bar{d}_0 \\ \bar{d}_1 \\ \bar{d}_2 \\ \vdots \\ \vdots \\ \bar{d}_{n-1} \\ \bar{d}_n \end{bmatrix}$$

Back-substitution gives the solution:

$$y'_n = \frac{\bar{d}_n}{\bar{b}_n}$$

$$y'_k = \frac{\bar{d}_k - c_k y'_{k+1}}{\bar{b}_k} \quad , \quad k = n-1, n-2, \ldots, 2, 1, 0$$

These calculations can be done very fast in a computer. Only four vectors whose components are a_k, b_k, c_k, d_k, $k = 0, \ldots, n$ need to be stored.

4.6. Modification for the periodic case.

In the periodic case the linear system looks as follows

$$\begin{bmatrix} b_0 & c_0 & & & & & a_0 \\ a_1 & b_1 & c_1 & & & & \\ & a_2 & b_2 & c_2 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & a_{n-2} & b_{n-2} & c_{n-2} \\ c_{n-1} & & & & & a_{n-1} & b_{n-1} \end{bmatrix} \begin{bmatrix} y'_0 \\ y'_1 \\ y'_2 \\ \cdot \\ \cdot \\ \cdot \\ y'_{n-2} \\ y'_{n-1} \end{bmatrix} = \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \cdot \\ \cdot \\ \cdot \\ d_{n-2} \\ d_{n-1} \end{bmatrix}$$

The system is not completely tridiagonal. Hence the formulas for Gauss elimination must be modified. The triangular decomposition phase starts by treating only the first n-2 equations as follows:

Put

$$\bar{b}_0 = b_0; \quad e_0 = \bar{e}_0 = a_0; \quad e_k = 0; \quad k = 2, \ldots, n-3; \quad e_{n-2} = c_{n-2}$$

Thus e_k, $k = 1, \ldots, n-2$ gives the first n-2 elements of the last column of the matrix.

Compute:

$$\bar{b}_k = b_k - \frac{a_k}{\bar{b}_{k-1}} c_{k-1}$$

$$\bar{d}_k = d_k - \frac{a_k}{\bar{b}_{k-1}} \bar{d}_{k-1} \quad , \qquad k = 1, 2, \ldots, n-2$$

$$\bar{e}_k = e_k - \frac{a_k}{\bar{b}_{k-1}} \bar{e}_{k-1}$$

After these operations the system looks as

$$\begin{bmatrix} \bar{b}_0 & c_0 & & & & \bar{e}_0 \\ & \bar{b}_1 & c_1 & & & \bar{e}_1 \\ & & \bar{b}_2 & c_2 & & \bar{e}_2 \\ & & & \ddots & \ddots & \vdots \\ & & & & \ddots & \vdots \\ & & & & \bar{b}_{n-2} & \bar{e}_{n-2} \\ c_{n-1} & & & & a_{n-1} & \bar{b}_{n-1} \end{bmatrix} \begin{bmatrix} y'_0 \\ y'_1 \\ y'_2 \\ \vdots \\ \vdots \\ y'_{n-2} \\ y'_{n-1} \end{bmatrix} = \begin{bmatrix} \bar{d}_0 \\ \bar{d}_1 \\ \bar{d}_2 \\ \vdots \\ \vdots \\ \bar{d}_{n-2} \\ \bar{d}_{n-1} \end{bmatrix}$$

The triangular decomposition phase is completed by the following set of formulas.

Put

$$f_0 = c_{n-1}$$

$$f_k = 0 \qquad k = 2, \ldots, n-3$$

$$f_{n-2} = a_{n-1}$$

Thus $f_0, \ldots, f_{n-1}$ represent the first n-2 elements in the last row. Also put

$$\bar{f}_0 = f_0 .$$

Compute

$$\bar{f}_k = f_k - \frac{f_{k-1}}{\bar{b}_{k-1}} c_{k-1} \quad , \qquad k = 1, \ldots, n-2$$

and:

$$\bar{\bar{b}}_{n-1} = \bar{b}_{n-1} - \sum_{k=1}^{n-1} \frac{f_{k-1}}{\bar{b}_{k-1}} \bar{e}_{k-1}$$

$$\bar{\bar{d}}_{n-1} = \bar{d}_{n-1} - \sum_{k=1}^{n-1} \frac{f_{k-1}}{\bar{b}_{k-1}} \bar{d}_{k-1}$$

The system is now triangular:

$$\begin{bmatrix} \bar{b}_0 & c_0 & & & & \bar{e}_0 \\ & \bar{b}_1 & c_1 & & & \bar{e}_1 \\ & & \bar{b}_2 & c_2 & & \bar{e}_2 \\ & & & \ddots & \ddots & \vdots \\ & & & & \bar{b}_{n-2} & \bar{e}_{n-2} \\ & & & & & \bar{\bar{b}}_{n-1} \end{bmatrix} \begin{bmatrix} y'_0 \\ y'_1 \\ y'_2 \\ \vdots \\ y'_{n-2} \\ y'_{n-1} \end{bmatrix} = \begin{bmatrix} \bar{d}_0 \\ \bar{d}_1 \\ \bar{d}_2 \\ \vdots \\ \bar{d}_{n-2} \\ \bar{\bar{d}}_{n-1} \end{bmatrix}$$

Back-substitution is done by

$$y'_{n-1} = \frac{\bar{\bar{d}}_{n-1}}{\bar{\bar{b}}_{n-1}}$$

$$y'_{n-2} = \frac{\bar{d}_{n-2} - \bar{e}_{n-2} y'_{n-1}}{\bar{b}_{n-2}}$$

$$y'_k = \frac{\bar{d}_k - c_k y'_{k+1} - \bar{e}_k y'_{n-1}}{\bar{b}_k} , \qquad k = n-3, \ldots, 0$$

4.7. Interpolation of curves in the plane.

The parameter representation of a curve in the plane is given by two functions

$$x = x(t) , \quad y = y(t).$$

There is some redundancy in the parameter representation. It defines not only the shape of the curve; it supplies in addition a mapping of an interval of the real line onto the curve. One can require that the parameter t equals s, the arc length along the curve. In our subsequent formulas t will be close to s, but not quite identical.

Let a discrete set of points (x_i, y_i), $i = 0, \ldots, n$ be given. The requirement is to interpolate a smooth curve through those points. Cf. fig. 4.2.

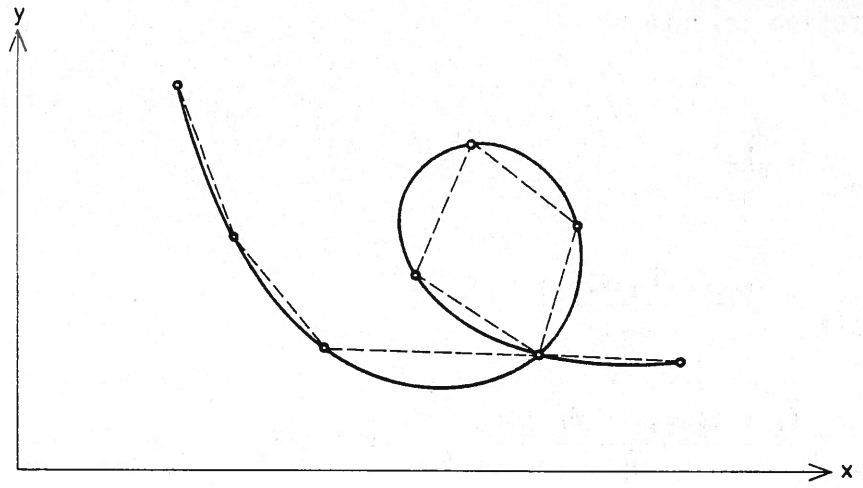


Fig. 4.2.

Note that the curve of fig. 4.2 could not be represented as

$$y = f(x) \quad \text{or} \quad x = g(y).$$

The functions f(x), g(y) would not be single valued.

We consider the polygon of chords also shown in fig. 4.2. Our first choice of the parameter t will be the arc length along this polygon. We arrive at two conventional interpolation problems:

Interpolate

$$x(t_i) = x_i, \quad i = 0, \ldots, n$$

and

$$y(t_i) = y_i, \quad i = 0, \ldots, n$$

for t_i being the length from point 0 to point i measured along the polygon of chords. We use the apparatus developed in the previous section to obtain two spline curves x(t), y(t). The boundary conditions are chosen according to the given situation. A closed curve would require periodic boundary conditions. The resulting curve has continous curvature, (because x(t) and y(t) are twice differentiable functions).

A slight flaw is that t is not the arc length. One can improve upon this by computing the arc length at the points 0, ..., n along the interpolated curve by means of numerical integration. One obtains value $s_0 = 0, s_1, \ldots, s_n$. Replacing $t_0 = 0, t_2, \ldots, t_n$ by these values, one could recompute the spline. The procedure could be iterated a few times. Frequently, however, one is satisfied with t being the arc length along the polygon.

4.8. Splines viewed as a vector space.

Given a fixed partition $x_0 < x_1 < \ldots < x_n$, we consider the set of all spline functions for all possible ordinates $y_0, \ldots, y_n$, and for all possible boundary conditions. One readily verifies that this set forms a vector space. Because n+3 parameters are necessary to uniquely specify a particular spline (a possible choice for those parameters is a set of values $y_0, \ldots, y_n$ together with y'_0, y'_n!), the dimension of the vector space is n+3. We can construct a basis as follows.

Let $a_i(x)$ be the spline fulfilling

$$a_i(x_j) = \delta_{ij} \qquad i,j=0,\ldots,n$$

and

$$a_i'(x_0) = a_i'(x_n) = 0$$

Let $\alpha_0(x)$, $\alpha_n(x)$ be two splines fulfilling

$$\alpha_0(x_j) = \alpha_n(x_j) = 0 \qquad j=0,\ldots,n$$
$$\alpha_0'(x_0) = 1, \quad \alpha_0'(x_n) = 0$$
$$\alpha_n'(x_0) = 0, \quad \alpha_n'(x_n) = 1$$

The splines $a_0(x),\ldots,a_n(x)$, $\alpha_0(x)$, $\alpha_n(x)$ form a basis. In fact the function

$$s(x) = \sum_{i=0}^{n} y_i a_i(x) + y_0'\alpha_0(x) + y_n'\alpha_n(x)$$

gives precisely the spline interpolating $y_0,\ldots,y_n$ and having boundary derivatives $s'(x_0) = y_0'$, $s'(x_n) = y_n'$. Fig. 4.3 shows some of the basis functions, assuming n=8 and equidistant data $x_i = i$, $i=0,\ldots,n$.

Another basis which is practically more important will be discussed in chapter 6.

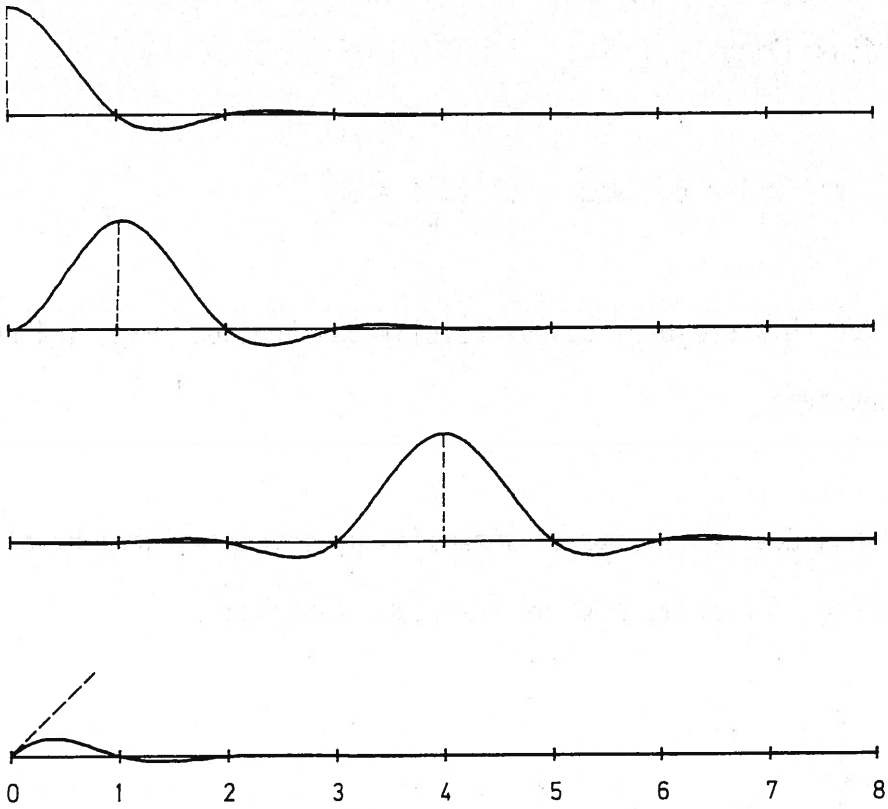


Fig. 4.3

A sample of basis splines for n=8

The splines s(x) fulfilling

$$s'(x_0) = 0 \quad \text{and} \quad s'(x_n) = 0$$

form a subspace of dimension n+1. A basis for this subspace is given by $a_i(x)$, $i=0,\ldots,n$.

Another subspace of dimension n+1 is given by the set of all free splines. i.e. splines fulfilling

$$s''(x_0) = 0 \quad \text{and} \quad s''(x_n) = 0$$

A basis for this subspace can be readily constructed. Fig. 4.4 shows some of the basis functions.

A third subspace, this time of dimension n, is represented by all periodic splines. Fig. 4.5 shows some of its basis functions.

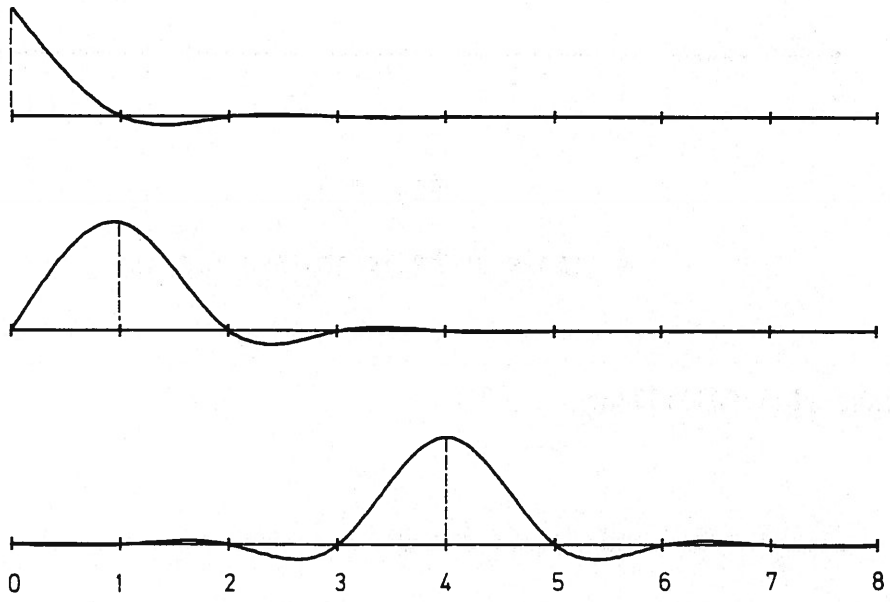


Fig. 4.4

A sample of free basis splines for n=8

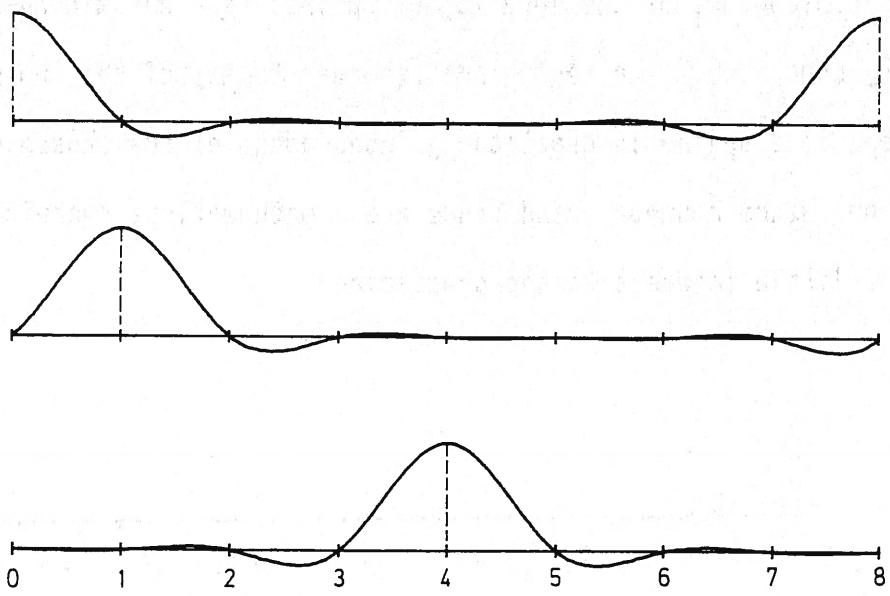


Fig. 4.5

A sample of periodic basis splines for n=8

4.9. The locality of splines.

A look at fig.'s 4.3 - 5 is very instructive. Any basis spline is appreciably different from zero only in the vicinity of the node to which it is associated. The more we go away from this node, the more the amplitudes are dampened. One can show that the dampening is exponential. The practical implication of this phenomenon is very important. The shape of the spline interpolated in a small region is only influenced by data near this region. If data are changed at a location far away the shape of the spline will not be noticeably changed anywhere else. The spline is as smooth as the data in a close vicinity imply. This is not so with polynomial interpolation. Look at fig. 4.6 showing basis functions for polynomial interpolation.

Another consequence of the locality of splines is a certain de-emphasis of the boundary condition. In a region not too near to any of the two boundary nodes the shape of a spline is practically independent of the chosen boundary condition. Hence boundary conditions are a mathematical technicality which is often of little interest to the practician.

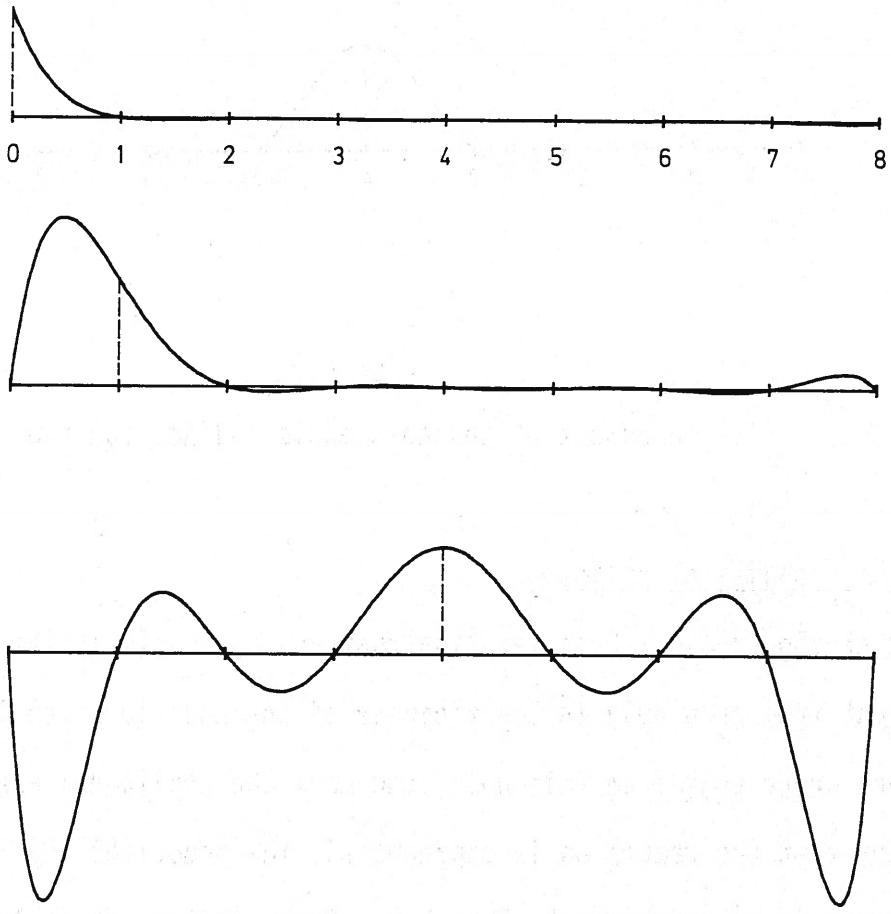


Fig. 4.6

A sample of Lagrange interpolation polynomials for n=8

References

AHLBERG, J.H.; E.N. Nilson, and J.L. Walsh (1967): The theory of splines and their applications. Academic Press, New York and London. XI + 284 pages.

SCHOENBERG, I.J. (1946): Contributions to the problem of approximation of equidistant data by analytic functions. Quart. Appl. Math., vol. 4, pp. 45-99, 112-141.

SPAETH, H. (1973): Spline-Algorithmen zur Konstruktion glatter Kurven und Flaechen. R. Oldenburg Verlag, Muenchen, Wien. 134 Seiten. (In German. Fortran algorithms for computing spline curves are included.)

5. Two-dimensional spline interpolation.

5.1. Introduction.

Assume a Cartesian coordinate system in the plane. Define a grid by means of lines

$$x = x_0, \quad x = x_1, \ \ldots, \quad x = x_m, \quad x_0 < x_1 < \ldots < x_m$$

and

$$y = y_0, \quad y = y_1, \ \ldots, \quad y = y_n, \quad y_0 < y_1 < \ldots < y_n$$

The grid covers the rectangle

$$x_0 \leq x \leq x_m$$
$$y_0 \leq y \leq y_n$$

The intersections of the grid lines, i.e. the points (x_i,y_j), $i = 1, \ldots, m$, $j = 1, \ldots, n$ are called grid points or nodes. Assume that function values

$$z_{ij} = z(x_i,y_j), \quad i = 0, \ldots, m, \quad j = 0, \ldots, n$$

are defined at the grid points. Our purpose is to interpolate these function values by means of a smooth function defined everywhere inside the area covered by the grid.

5.2. Bicubic polynomials.

Focus attention on a particular sub-rectangle

$$x_i \leq x \leq x_{i+1}, \quad y_j \leq y \leq y_{j+1}$$

Consider there a bicubic polynomial

$$p(x,y) = \sum_{k=0}^{3} \sum_{\ell=0}^{3} a_{k\ell}(x-x_i)^k (y-y_j)^\ell$$

The polynomial is currently parameterized by its 16 coefficients $a_{k\ell}$, $k,\ell = 0, \ldots, 3$. We will re-parameterize it in terms of "nodal parameters". These are the values of the function $z(x,y)$ to be interpolated together with some of its derivatives. The nodal parameters are:

$$z(x,y), \quad z_x(x,y), \quad z_y(x,y), \quad z_{xy}(x,y)$$

evaluated at the four corners of the subrectangle. We introduce the notation

$$z_{ij} = z(x_i,y_j) \qquad u_{ij} = z_x(x_i,y_j)$$
$$v_{ij} = z_y(x_i,y_j) \qquad w_{ij} = z_{xy}(x_i,y_j)$$

The 16 nodal parameters are then

$$z_{i+r,j+s}, \quad u_{i+r,j+s}, \quad v_{i+r,j+s}, \quad w_{i+r,j+s}, \quad r, s = 0, 1$$

We introduce auxiliary variables ξ, η by putting

$$x = x_i + (x_{i+1}-x_i)\xi, \qquad y = y_j + (y_{j+1}-y_j)\eta$$

Then the polynomial $p(x,y)$ transforms into

$$\pi(\xi,\eta) = \sum_{k=0}^{3} \sum_{l=0}^{3} \alpha_{kl}\, \xi^k\, \eta^l, \qquad 0 \le \xi,\ \eta \le 1$$

with

$$\alpha_{kl} = a_{kl}(x_{i+1}-x_i)^k\, (y_{j+1}-y_j)^l$$

It also follows that

$$\begin{aligned}
\zeta_{rs} &= \pi(r,s) = p(x_{i+r},y_{j+s}) = z_{i+r,j+s} \\
\varphi_{rs} &= \pi_\xi(r,s) = (x_{i+1}-x_i)\, p_x(x_{i+r},y_{j+s}) = (x_{i+1}-x_i)\, u_{i+r,j+s} \\
\psi_{rs} &= \pi_\eta(r,s) = (y_{j+1}-y_j)\, p_y(x_{i+r},y_{j+s}) = (y_{j+1}-y_j)\, v_{i+r,j+s} \\
\omega_{rs} &= \pi_{\xi\eta}(r,s) = (x_{i+1}-x_i)\,(y_{j+1}-y_j)\, p_{xy}(x_{i+r},y_{j+s}) = \\
&= (x_{i+1}-x_i)\,(y_{j+1}-y_j)\, w_{i+r,j+s}, \qquad r,\ s = 0,\ 1
\end{aligned}$$

The change of variables has transformed our problem into one for a square with unit sides: $0 \le \xi \le 1$, $0 \le \eta \le 1$. Note that

$$\pi(\xi,\eta) = [1 \quad \xi \quad \xi^2 \quad \xi^3] \begin{bmatrix} \alpha_{00} & \alpha_{01} & \alpha_{02} & \alpha_{03} \\ \alpha_{10} & \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{20} & \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{30} & \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \eta \\ \eta^2 \\ \eta^3 \end{bmatrix}$$

Thus it holds that

$$\begin{bmatrix} s_{00} & \psi_{00} & s_{01} & \psi_{01} \\ \varphi_{00} & \omega_{00} & \varphi_{01} & \omega_{01} \\ s_{10} & \psi_{10} & s_{11} & \psi_{11} \\ \varphi_{10} & \omega_{10} & \varphi_{11} & \omega_{11} \end{bmatrix} =$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} \alpha_{00} & \cdots & \alpha_{03} \\ . & & . \\ . & & . \\ \alpha_{30} & \cdots & \alpha_{33} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

We write this as

$$K = H^T A H$$

It follows that

$$A = (H^T)^{-1} K H^{-1}$$

one verifies:

$$H^{-1} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & -3 & 2 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

Consequently the desired matrix of the coefficients α_{kl} is

$$\begin{bmatrix} \alpha_{00} & \alpha_{01} & \alpha_{02} & \alpha_{03} \\ \alpha_{10} & \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{20} & \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{30} & \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} =$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -3 & -2 & 3 & -1 \\ 2 & 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} \varsigma_{00} & \psi_{00} & \varsigma_{01} & \psi_{01} \\ \varphi_{00} & \omega_{00} & \varphi_{01} & \omega_{01} \\ \varsigma_{10} & \psi_{10} & \varsigma_{11} & \psi_{11} \\ \varphi_{10} & \omega_{10} & \varphi_{11} & \omega_{11} \end{bmatrix} \begin{bmatrix} 1 & 0 & -3 & 2 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

From the α_{kl}'s the a_{kl}'s follow by

$$a_{kl} = \alpha_{kl} \; / \; \left\{ (x_{i+1}-x_i)^k \; (y_{j+1}-y_j)^l \right\}$$

Remark: A bicubic polynomial may be parameterized by different nodal values. For example, also the values of

$$z(x,y), \quad z_{xx}(x,y), \quad z_{yy}(x,y), \quad z_{xxyy}(x,y)$$

at the four corners of a subrectangle could be used.

5.3. Hermite bicubic interpolation.

Suppose that $z_{ij} = z(x_i,y_j)$, $u_{ij} = z_x(x_i,y_j)$, $v_{ij} = z_y(x_i,y_j)$, $w_{ij} = z_{xy}(x_i,y_j)$ are available at all $(m+1)(n+1)$ grid points (i,j), $i = 0, \ldots, m$, $j = 0, \ldots, n$. Interpolate a bicubic polynomial

$$p^{ij}(x,y) = \sum_{k,l=0}^{3} a_{kl}^{ij} (x-x_i)^k (y-y_j)^l$$

in any of the $m \cdot n$ rectangles $x_i \leq x \leq x_{i+1}$, $y_j \leq y \leq y_{j+1}$, $i = 0, \ldots, m-1$, $j = 0, \ldots, n-1$. The resulting function

$$h(x,y) = p^{ij}(x,y) \quad \ldots \quad \text{for} \quad x_i \leq x \leq x_{i+1}, \quad y_j \leq y \leq y_{j+1}$$

is defined for the whole domain $x_0 \leq x \leq x_m$, $y_0 \leq y \leq y_n$. It is called a Hermite bicubic interpolation function.

Theorem: $h(x,y)$ is continuous and has continuous derivatives $h_x(x,y)$, $h_y(x,y)$ and $h_{xy}(x,y)$.

Proof: Consider $h(x,y)$ as a function of x, viewing y as a parameter. In $x_i \leq x \leq x_{i+1}$ we have

$$\lim_{y \to y_j + 0} h(x,y) = p^{ij}(x,y_j) = f_+(x), \text{ say}$$
$$\lim_{y \to y_j - 0} h(x,y) = p^{i,j-1}(x,y_j) = f_-(x), \text{ say}$$

The functions $f_+(x)$, $f_-(x)$ are one-dimensional cubic polynomials in x. We have

$$f_+(x_i) = f_-(x_i) = z_{ij}, \quad f'_+(x_i) = f'_-(x_i) = u_{ij}$$
$$f_+(x_{i+1}) = f_-(x_{i+1}) = z_{i+1,j}, \quad f'_+(x_{i+1}) = f'_-(x_{i+1}) = u_{i+1,j}$$

Because a one-dimensional cubic polynomial is uniquely determined by these values we have

$$f_+(x) \equiv f_-(x)$$

It follows that $h(x,y)$, $h_x(x,y)$ are continuous across x-grid lines. (Such grid lines are parallel to the x-axis; they are straight lines of constant $y = y_j$). Likewise it follows that $h(x,y)$, $h_y(x,y)$ are continuous across y-grid lines. Hence the continuity of $h(x,y)$ is already established.

Next we show that $h_x(x,y)$ is continuous across y-grid lines, and also that $h_y(x,y)$ is continuous across x-grid lines. Consider $h_y(x,y)$ as a function of x while y plays the role of a parameter. Call

$$\lim_{y \to y_j + 0} h_y(x,y) = p_x^{i,j}(x,y_j) = g_+(x)$$
$$\lim_{y \to y_j - 0} h_y(x,y) = p_x^{i,j-1}(x,y_j) = g_-(x)$$

The functions $g_+(x)$, $g_-(x)$ are cubic polynomials in x. We have

$$g_+(x_i) = g_-(x_i) = u_{ij}, \qquad g'_+(x_i) = g'_-(x_i) = w_{ij}$$
$$g_+(x_{i+1}) = g_-(x_{i+1}) = u_{i+1,j}, \qquad g'_+(x_{i+1}) = g'_-(x_{i+1}) = w_{i+1,j}$$

Similarly as above, one concludes that

$$g_+(x) \equiv g_-(x)$$

This proves the continuity of h_y across x-grid lines. By symmetry the continuity of h_x across y-grid lines follows too.

Finally, the continuity of $h_{xy}(x,y)$ will be shown. We have shown that $h_y(x,y)$ is continuous. For fixed y, $h_y(x,y)$ is a cubic in x for each subinterval $x_i \leq x \leq x_{i+1}$. Thus $h_{yx}(x,y)$ may be formed. At $y = y_j$, the derivatives $h_{yx}(x_i,y_j) = w_{ij}$ and $h_{yx}(x_{i+1},y_j) = w_{i+1,j}$ are prescribed. One infers that $h_{yx}(x,y)$ is continuous in the strip $x_i \leq x \leq x_{i+1}$ across all the x-grid lines. Interchanging the roles of x, y, noting that $h_{yx} = h_{xy}$, one infers that in the strip $y_j \leq y \leq y_{j+1}$, h_{xy} is continuous across the y-grid lines. Thus h_{xy} is continuous everywhere.

5.4. Bicubic splines.

5.4.1. Definition.

We want interpolating functions that are smoother than the Hermite bicubic interpolators. At least the continuity of all second derivatives will be required.

We take a one-dimensional spline A(x) with nodes $x_0, \ldots, x_m$, and we take

another one B(y) with nodes $y_0, \ldots, y_n$. We form the product

$$s(x,y) = A(x)\, B(y)$$

It is a function having the following continuous derivatives

$$s(x,y),\ s_x(x,y),\ s_y(x,y),\ s_{xx}(x,y),\ s_{xy}(x,y),\ s_{yy}(x,y)$$
$$s_{xxy}(x,y),\ s_{xyy}(x,y),\ s_{xxyy}(x,y)$$

We consider finite sets of one-dimensional splines $A_k(x)$, $k = 1, \ldots, M$ and $B_\ell(y)$, $\ell = 1, \ldots, N$. We form

$$s(x,y) = \sum_{k=1}^{M} \sum_{\ell=1}^{N} A_k(x)\, B_\ell(y)$$

The same statement about the continuity of the above specified derivatives can be made. The set of all functions obtained in this way is a vector space. We call it the space of bicubic splines for the grid $x_0, \ldots, x_m,\ y_0, \ldots, y_n$.

5.4.2. The constrained bicubic spline.

Theorem: Specify

$$z_{ij} = z(x_i, y_j),\quad i = 0, \ldots, m,\quad j = 0, \ldots, n.$$

Also specify

$$u_{0j} = z_x(x_0, y_j),\quad u_{mj} = z_x(x_m, y_j),\quad j = 0, \ldots, n$$
$$v_{i0} = z_y(x_i, y_0),\quad v_{in} = z_y(x_i, y_n),\quad i = 0, \ldots, m$$

Finally specify

$$w_{00} = z_{xy}(x_0,y_0), \quad w_{m0} = z_{xy}(x_m,y_0), \quad w_{0n} = z_{xy}(x_0,y_n),$$
$$w_{mn} = z_{xy}(x_m,y_n)$$

It is asserted that a unique bicubic spline matching these values exists.

Proof: We first prove the existence of the spline. We take the basis $a_i(x)$, $i = 0, \ldots, m$, $\alpha_0(x)$, $\alpha_m(x)$ for the one-dimensional splines A(x). This basis was constructed in section 4.8. We take a similar basis $b_j(y)$, $j = 0, \ldots, n$, $\beta_0(y)$, $\beta_n(y)$ for the splines B(y).

We consider the bicubic spline:

$$\begin{aligned} s(x,y) = & \sum_{i=0}^{m} \sum_{j=0}^{n} z_{ij}\, a_i(x)\, b_j(y) + \\ & + \sum_{j=0}^{n} [u_{0j}\, \alpha_0(x)\, b_j(y) + u_{mj}\, \alpha_m(x)\, b_j(y)] + \\ & + \sum_{i=0}^{m} [v_{i0}\, a_i(x)\, \beta_0(y) + v_{in}\, a_i(x)\, \beta_n(y)] + \\ & + w_{00}\, \alpha_0(x)\, \beta_0(y) + w_{m0}\, \alpha_m(x)\, \beta_0(y) \\ & + w_{0n}\, \alpha_0(x)\, \beta_n(y) + w_{mn}\, \alpha_m(x)\, \beta_n(y) \end{aligned}$$

The bicubic spline s(x,y) completely solves the stated interpolation problem. This proves existence.

We now prove uniqueness. We show that no other function s(x,y) exists which has the following properties.

(1) $s(x,y)$ is a bicubic polynomial in each subrectangle.

(2) $s(x,y)$ has continuous derivatives s, s_x, s_y, ..., s_{xxyy} as specified earlier.

(3) $s(x,y)$ interpolates the data specified in the theorem.

It suffices to show that a function satisfying (1) and (2), having vanishing function values at the nodes and having vanishing boundary conditions, is necessarily the zero function. (In other words, a function satisfying (1) and (2), and interpolating data that are all zero, must be the zero function). Such a function is a one-dimensional bicubic spline along any grid line. Along a grid line, e.g. that one for $y = y_{j_0}$, we have with $f(x) = s(x,y_{j_0})$: $f(x_i) = z_{i,j_0} = 0$ and $f'(x_0) = u_{0,j_0} = f'(x_m) = u_{m,j_0} = 0$. Hence this one-dimensional spline is the zero spline. It follows that all u_{ij} vanish. Likewise one concludes that all v_{ij} are zero. The function $g(x) = s_y(x,y_0)$ is a cubic in each subinterval $x_i \leq x \leq x_{i+1}$. It must be a spline, because the derivatives s_{yx} and s_{yxx} are required to be continuous. One has $g(x_i) = v_{i0} = 0$. Also $g'(x_0) = w_{00} = g'(x_m) = = w_{m0} = 0$. Thus $g(x) = 0$, and consequently $w_{i0} = 0$, $i = 0, \ldots, m$. Similarly one infers that $w_{in} = 0$, $i = 0, \ldots, m$. Now the splines $h_j(x) = s_y(x,y_j)$ can be interpolated. We have $h_j(x_i) = v_{ij_0} = 0$, $h'_j(x_0) = w_{0j} = h'_j(x_m) = w_{mj} = 0$. Thus $h_j(x) = 0$ and $w_{ij} = 0$ for all i, j. A Hermite bicubic function having $z_{ij} = u_{ij} = v_{ij} = w_{ij} = 0$ must be the zero function. This concludes the proof.

During our uniqueness proof we have in effect proved more than we originally wanted. Without any further argument we can state the following theorem.

Theorem: The space of bicubic splines $s(x,y)$ coincides with the space of Hermite bicubic interpolation functions upon which the requirement is imposed that in addition to $s(x,y)$, $s_x(x,y)$, $s_y(x,y)$, $s_{xy}(x,y)$, also the derivatives $s_{xx}(x,y)$, $s_{xy}(x,y)$, $s_{yy}(x,y)$, $s_{xxy}(x,y)$, $s_{xyy}(x,y)$ and $s_{xxyy}(x,y)$ are continuous. The space of bicubic splines over the grid $x_0, \ldots, x_m, y_0, \ldots, y_n$ has dimension $(m+3)(n+3)$. A basis has been exhibited above. It consists of all products $A(x)\,B(y)$ with

$$A(x) \in \{a_0(x), \ldots, a_m(x), \alpha_0(x), \alpha_m(x)\}$$

and

$$B(y) \in \{b_0(y), \ldots, b_n(y), \beta_0(y), \beta_n(y)\}$$

Thus it consists of all products of basis functions for the one-dimensional splines over $x_0, \ldots, x_m$ and $y_0, \ldots, y_n$.

5.4.3. The free bicubic spline.

It has boundary conditions

$$z_{xx}(x_0,y_j) = z_{xx}(x_m,y_j) = 0, \quad j = 0, \ldots, n$$
$$z_{yy}(x_i,y_0) = z_{yy}(x_i,y_n) = 0, \quad i = 0, \ldots, n$$
$$z_{xxyy}(x_0,y_0) = z_{xxyy}(x_m,y_0) = z_{xxyy}(x_0,y_n) = z_{xxyy}(x_m,y_n) = 0$$

A basis is readily obtained by using bases for the one-dimensional free splines

and performing the (tensor) products, as it was done in the case of the constrained spline.

5.4.4. The double periodic spline.

The requirements are that s(x,y) is periodic in x as well as in y. A basis can be specified in an obvious way.

6. Geometry of exact spline interpolation.

6.1. Formulation of the problem.

Let V, W be Hilbert spaces. We consider two linear and continuous mappings.

(1) The operator Λ maps V onto R^n. In case of finite-dimensional V, we let Λ be represented by the matrix A. This matrix has more columns than rows, i.e. V is of higher dimension than n, the dimension of R^n.

(2) The operator Θ maps V into W. The inner product in W shall be represented by the matrix Q. In case of finite-dimensional V, W, we let the matrix B represent the operator Θ. B is then also a matrix having more columns than rows, i.e. V is of higher dimension than W.

We require that the nullspaces of Λ and Θ have only the zero vector in common

$$N(\Lambda) \cap N(\Theta) = 0$$

and pose the following problem.

Given $y \in R^n$, find $x \in V$ such that

(a) $\Lambda(x) = y$ ... the interpolation requirement

(b) $\|\Theta(x)\|_W$ = minimum ... the minimum norm requirement.

Example 1: Let V be the Hilbert space of functions $F(\xi)$, $\alpha \leq \xi \leq \beta$ with inner

product

$$(f_1, f_2) = \int_\alpha^\beta \left\{ f_1(\xi)\, f_2(\xi) + f_1'(\xi)\, f_2'(\xi) + f_1''(\xi)\, f_2''(\xi) \right\} d\xi$$

Let W be the Hilbert space of functions $g(\xi)$, $\alpha \leq \xi \leq \beta$ with inner products

$$(g_1, g_2) = \int_\alpha^\beta g_1(\xi)\, g_2(\xi)\, d\xi$$

Let $\xi_1, \ldots, \xi_n$ be distinct points in the interval $[\alpha,\beta]$. Let Λ be the operator mapping $f \in V$ onto the vector

$$\Lambda(f) = \begin{bmatrix} f(\xi_1) \\ \vdots \\ f(\xi_n) \end{bmatrix} \in R^n$$

Let θ be the operator mapping $f \in V$ onto $f'' \in W$. Thus θ maps a function f onto its second derivative

$$\theta(f) = f'' = \frac{d^2}{d\xi^2} f$$

We then pose the problem

Given

$$y = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} \in R^n$$

find a function $f \in V$ such that

(a) $f(\xi_i) = \eta_i$, $i = 1, \ldots, n$, i.e. f interpolates the prescribed function values η_i at the locations ξ_i

(b) $\|f''\|^2 = \int_\alpha^\beta (f''(x))^2 dx$ = minimum, i.e. f is in a sense the smoothest interpolating function.

Example 2: Instead of a continuous interval of arguments ξ, $\alpha \leq \xi \leq \beta$, we consider a large, but discrete set of equidistant values

$$\alpha = \xi_0 < \xi_1 < \ldots < \xi_m = \beta$$

We denote the step size

$$h = \xi_{i+1} - \xi_i, \quad i = 0, \ldots, m-1$$

We consider a corresponding set of discrete function values

$$f_i = f(\xi_i), \quad i = 0, \ldots, m$$

Out of the set $\{\xi_i\}$ we select a subset of n elements $(\xi_{i_1}, \ldots, \xi_{i_n})$. The $\xi_{i_1}, \ldots, \xi_{i_n}$ need not be equally spaced. At these selected locations we prescribe function values $\eta_1, \ldots, \eta_n$.

We pose the problem:

Given $\eta_1, \ldots, \eta_n$ find a vector $(f_0, f_1, \ldots, f_m)$ such that

(a) $f_{i_j} = \eta_j, \quad j = 1, \ldots, n \quad \ldots$ interpolation requirement

(b) $\frac{1}{h^2} \sum_{i=1}^{m-1} (-f_{i+1} + 2f_i - f_{i-1})^2 = \text{minimum} \quad \ldots$ smoothness requirement.

Obviously this is a discrete version of the problem in example 1. The operator Λ maps $f = (f_0, \ldots, f_m)$ onto the subset of components $\eta_1 = f_{i_1}, \ldots, \eta_n = f_{i_n}$. The operator θ maps the vector f onto the set of second difference quotients

$$\Delta^2 f_i = \frac{1}{h^2} (-f_{i+1} + 2f_i - f_{i-1}) = \frac{\frac{1}{h}(f_{i+1} - f_i) - \frac{1}{h}(f_i - f_{i-1})}{h}$$

6.2. Definition of splines.

In V we consider the nullspace $N = N(\Lambda)$ of the operator Λ. Referring to example 1, N_Λ consists of functions vanishing at the locations of interpolation ξ_i, $i = 1, \ldots, n$. The operator θ maps functions $f \in N(\Lambda)$ onto functions g of a certain subspace $U = U(\Lambda, \theta)$ of W:

$$U = U(\Lambda, \theta) = \{z \in W \mid \text{there is } x \in N(\Lambda) \text{ such that } \theta(x) = z\}$$

Referring to example 1, U is the set of second derivatives of functions vanishing at the locations of interpolation.

We consider the orthocomplement $U^\perp$ of U in W. We consider the pre-images

of elements in $U^{\perp}$ under θ. This is a subspace $S = S(\Lambda,\theta)$ of V.

$$S = S(\Lambda,\theta) = \{x \in V \mid \theta(x) \in U^{\perp}\}$$

S is called the subspace of splines in V. Its elements are called splines. Later it will be shown that they are the solutions of the extremum problem formulated in section 6.1.

Example 1a: (Continuation of example 1)
As we have noticed, $N = N(\Lambda)$ is the set of functions vanishing at the locations of interpolation. $U = U(\Lambda,\theta)$ is the set of second derivatives of such functions. What is the set $U^{\perp}$? What is the set $S = S(\Lambda,\theta)$?

Theorem: For the spaces and operators of example 1 the set S is the set of piecewise cubic polynomials $p(\xi)$ on $\alpha \leq \xi \leq \beta$ having the following properties

$p(\xi)$ is linear in $\alpha \leq \xi \leq \xi_1$, $\xi_n \leq \xi \leq \beta$
$p(\xi)$ is cubic in any of the intervals $\xi_i \leq \xi \leq \xi_{i+1}$, $i = 1, \ldots, n-1$
$p(\xi)$ is continuous together with its first and second derivatives.

Such functions are called cubic splines.

Remark: An equivalent formulation of the above theorem would be: The set $U^{\perp}$ consists of piecewise linear functions vanishing in $\alpha \leq \xi \leq \xi_1$, $\xi_n \leq \xi \leq \beta$, and whose slope is constant in any of the intervals $\xi_i \leq \xi \leq \xi_{i+1}$, $i = 1, \ldots, n-1$.

Proof: It suffices to prove the version given in the preceding remark. We decompose the proof into two steps, showing

(1) Any piecewise linear function $q(\xi)$ with possible discontinuities of $q'(\xi)$ only at ξ_i, $i = 1, \ldots, n$ and vanishing in $\alpha \leqslant \xi \leqslant \xi_1$, $\xi_n \leqslant \xi \leqslant \beta$, is orthogonal to any second derivative $f''(\xi)$ of a function vanishing at ξ_i, $i = 1, \ldots, n$.

To show this compute $(f'',q)_W =$

$$\int_\alpha^\beta f''(\xi)\, q(\xi)\, d\xi = \int_\alpha^{\xi_1} \ldots + \sum_{i=1}^{n-1} \int_{\xi_i}^{\xi_{i+1}} \ldots + \int_{\xi_n}^{\beta} \ldots$$

Apply partial integration twice, obtaining

$$\begin{aligned} & f'(\beta)\, q(\beta-0) - f'(\alpha)\, q(\alpha+0) - f(\beta)\, q'(\beta-0) + f(\alpha)\, q'(\alpha+0) \\ & - \sum_{i=1}^{n-1} \left\{ f(\xi_{i+1})\, q'(\xi_{i+1}-0) - f(\xi_i)\, q'(\xi_i+0) \right\} + \int_\alpha^{\xi_1} f(\xi)\, q''(\xi)\, d\xi \\ & + \sum_{i=1}^{n-1} \int_{\xi_i}^{\xi_{i+1}} f(\xi)\, q''(\xi)\, d\xi + \int_{\xi_n}^{\beta} f(\xi)\, q''(\xi)\, d\xi \end{aligned}$$

All terms vanish because of the properties of $f(\xi)$ and $q(\xi)$.

(2) Let $q(\xi)$ have properties required in (1). Such functions form a linear subspace in W. We show: Any function $g \in W$ orthogonal to this subspace can be viewed as the second derivative $g(\xi) = f''(\xi)$ of an $f \in V$, vanishing

at ξ_i, $i = 1, \ldots, n$. Put

$$f(\xi) = \int_{\alpha}^{\xi} d\eta \int_{\alpha}^{\eta} g(\theta)\, d\theta + c\,x + d$$

Then

$$g(\xi) = f''(\xi)$$

Adjust the constants of integration c, d such that

$$f(\xi_1) = f(\xi_2) = 0.$$

From

$$\int_{\alpha}^{\beta} g(\xi)\, q(\xi)\, d\xi = \int_{\alpha}^{\beta} f''(\xi)\, q(\xi)\, d\xi = 0$$

conclude by two fold partial integration (cf. the above formula under (1)) that:

$$\sum_{i=1}^{n-1} [f(\xi_{i+1})\, q'(\xi_{i+1}-0) - f(\xi_i)\, q'(\xi_i+0)] = 0$$

Choose piecewise linear functions $q_i(\xi)$, $i = 2, \ldots, n-1$ such that

$$q'_i(\xi_i-0) = \frac{1}{\xi_i - \xi_{i-1}} = \varkappa_i, \text{ say}$$

$$q'_i(\xi_i+0) = -\frac{1}{\xi_{i+1} - \xi_i} = \lambda_i, \text{ say}$$

Figure 6.1 shows such a function. The derivatives of $q_i(\xi)$ outside of

$\xi_{i-1} \leq x \leq \xi_{i+1}$ are required to vanish.

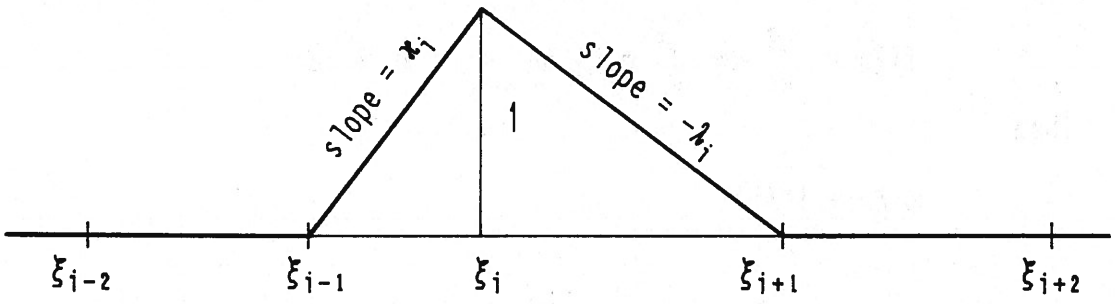


Fig. 6.1. The function $q_i(\xi)$

Inserting these functions into the above equation gives a linear system

$$\begin{bmatrix} -\lambda_2 & & & \\ (\alpha_3+\lambda_3), & -\lambda_3 & & \\ -\alpha_4, & (\alpha_4+\lambda_4), & -\lambda_4 & \\ & \cdot & \cdot & \cdot \\ & & -\alpha_{n-1}, & (\alpha_{n-1}+\lambda_{n-1}), & -\lambda_{n-1} \end{bmatrix} \cdot \begin{bmatrix} f(\xi_3) \\ f(\xi_4) \\ \cdot \\ \cdot \\ f(\xi_n) \end{bmatrix} = 0$$

This system is regular and homogeneous. Hence the solution must be zero. Q.e.d.

6.3. Existence and uniqueness of splines.

Recall the decomposition of the space W into U and $U^{\perp}$. $U = U(\Lambda,\theta)$ contains the images of $N = N(\Lambda)$ under the operator θ. Recall that the space of splines $S = S(\Lambda,\theta)$ was defined as the set of pre-images of vectors in $U^{\perp}$.

Theorem: The space V is the direct sum of $N(\Lambda)$ and $S(\Lambda,\theta)$.

Remark: This means that any vector $x \in V$ is uniquely represented as $x = x_N + x_S$, $x_N \in N$, $x_S \in S$. Confer section A.5.1.

Proof: We start by noting that any pre-image x of a vector $y \in U$ may be uniquely decomposed as $x = x_a + x_b$, $x_a \in N(\Lambda)$, $x_b \in N(\theta)$. [If $\theta(x) = y$ with $y \in U$, then y is also the image of some vector x_a in $N(\Lambda)$: $y = \theta(x_a)$. Put $x_b = x - x_a$, then $\theta(x_b) = \theta(x) - \theta(x_a) = y - y = 0$. Thus $x_b \in N(\theta)$]. Next we show that any $x \in V$ may be decomposed as

$$x = x_N + x_S, \quad x_N \in N, \quad x_S \in S$$

Let $y = \theta(x)$. Split $y = y_1 + y_2$, $y_1 \in U$, $y_2 \in U^{\perp}$. Let x_1, x_2 be pre-images of y_1, y_2, respectively. Put $x_3 = x - x_1 - x_2$. Then $\theta(x_3) = y - y_1 - y_2 = 0$. Thus $x = x_1 + x_2 + x_3$ with x_1 being a pre-image of a vector in U, x_2 being in S, and x_3 in $N(\theta)$. As shown earlier, we may split x_1 into x_4 and x_5, $x_4 \in N(\Lambda)$, $x_5 \in N(\theta)$. Vectors in $N(\theta)$ have zero images. They may therefore be viewed as vectors of S. Calling $x_N = x_4$, $x_S = x_2 + x_3 + x_5$, we get the desired decomposition.

Finally, we show that only the zero vector is common to N and S. If $x \in N \cap S$, then $y = \theta(x)$ is in U as well as in $U^{\perp}$; hence $y = 0$. Thus $x \in N(\theta)$. However it

was postulated in section 6.1 that the only vector common to $N = N(\Lambda)$ and $N(\theta)$ is the zero vector.

We have shown that any vector $x \in V$ may be uniquely decomposed as $x_N + x_S$. This proves existence and uniqueness of the spline x_S. The original vector (function) x and x_S interpolate the same data y. For $\Lambda(x) = \Lambda(x_N + x_S) = \Lambda(x_S)$.

6.4. Minimum properties of splines.

Theorem: Let $x \in V$ and let x_S be its spline. Then x_S solves the following two extremum problems:

$$(I) \quad \min_{z \in S} \|\theta(x) - \theta(z)\|_W = \|\theta(x) - \theta(x_S)\|$$

$$(II) \quad \min_{z \in V,\, \Lambda(z) = \Lambda(x)} \|\theta(z)\|_W = \|\theta(x_S)\|$$

Remark: Problem (II) was the problem stated in section 6.1. It shows that splines have "the minimum norm property". Among all functions interpolating the same data as x does, $\theta(x_S)$ has the smallest norm in W. Problem (I) is complementary to problem (II). Confer chapter A.11 for a general discussion of complementary least squares problems. Problem (I) shows that x_S has the "best approximation property". Out of the space of splines, the spline x_S is closest to a given function x.

Proof: Given $x \in V$, decompose $x = x_N + x_S$. In case of problem (I) we let

z vary over the spline space S. Then

$$x - z = x_N + (x_S - z), \quad x_N \in N, \quad x_S - z \in S$$

Thus

$$\|\theta(x-z)\|^2 = \|\theta(x_N)\|^2 + \|\theta(x_S-z)\|^2$$

This is minimal for $z = x_S$.

In case of problem (II) we let z vary over the set $\{z \mid \Lambda(z) = \Lambda(x)\}$. Thus $z-x \in N$ or $z = x + u$, $u \in N$. Hence

$$z = x_N + u + x_S, \quad x_N + u \in N$$
$$\|\theta(z)\|^2 = \|\theta(x_N+u)\|^2 + \|\theta(x_S)\|^2$$

The minimum is obtained for $u = -x_N$. This proves (II).

6.5. Other examples.

The examples 1 and 1a treated earlier refer to the case of the free spline with nodes $\xi_1, \ldots, \xi_n$. It has been shown that the free spline interpolating $f(\xi)$ at $\xi_1, \ldots, \xi_n$ has certain minimal properties (I) and (II). Similar properties hold for the constrained spline and for the periodic spline. It is then preferable to put $\xi_1 = \alpha$ and $\xi_n = \beta$. Among all functions $s(\xi)$ interpolating $f(\xi)$ at $\xi_1, \ldots, \xi_n$, and having the same derivatives as $f(\xi)$ at the boundary nodes ξ_1

and ξ_n, the constrained spline minimizes $\|s''(\xi)\|$. A similar statement holds for the periodic spline.

Also a generalization to the bicubic splines is possible. It is interesting to note that in the two-dimensional case the operator θ is given by

$$\theta(f) = f_{\xi\xi\eta\eta}$$

Thus the squared norm in W is given by

$$\|f\|_W^2 = \int_{\xi_1}^{\xi_n} \int_{\eta_1}^{\eta_n} f_{\xi\xi\eta\eta}(\xi,\eta)^2 \, d\xi \, d\eta$$

6.6. Prediction as a special case of spline interpolation.

Let

$$V = R^{n+m}$$

A vector $x \in V$ is represented as

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad x_1 \in R^n, \quad x_2 \in R^m$$

Let Λ be implied by

$$\ell = (I, 0)\, x = x_1$$

Let the space W coincide with V and let the norm in $V = W$ be implied by

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$$

We denote, as usual

$$Q = P^{-1} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$$

Let the operator θ be the identity:

$$\theta = I$$

Consider problem (II): Given $\ell \in R^n$, find $x \in V$ such that

$x_1 = \ell$ ... interpolation requirement

$\|x\|$ = minimum ... minimum norm requirement

This reduces to the problem

$$x^T P x = \text{minimum}$$

subject to

$$(I, 0)\, x = \ell$$

The problem is formally equivalent to a conditioned adjustment problem with zero observations, residuals x and discrepancies $-\ell$; i.e.

$$(0 + x)^T P (0 + x) = \text{minimum}$$

subject to

$$(I \;\; 0)(0 + x) = \ell$$

Thus the solution is

$$x = P^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} k, \quad \text{i.e.} \quad x = \begin{bmatrix} Q_{11} \\ Q_{21} \end{bmatrix} k$$

where the correlates k follow from the normal equations

$$(I \;\; 0)\, P^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} k - \ell = 0$$

i.e.

$$Q_{11}\, k = \ell, \quad k = Q_{11}^{-1}\, \ell$$

Thus

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} Q_{11} \\ Q_{21} \end{bmatrix} Q_{11}^{-1} \ell = \begin{bmatrix} \ell \\ Q_{21}Q_{11}^{-1}\ell \end{bmatrix}$$

The formula

$$x_2 = Q_{21}Q_{11}^{-1}\ell = Q_{21}Q_{11}^{-1}x_1$$

is the familiar prediction formula. (Usually written $s_2 = \Sigma_{21}\Sigma_{11}^{-1}s_1$).

Replacing the solution x by s, which stands for "spline", we know that the set of all splines is obtained by letting ℓ vary all over R^n.

Any spline $s \in S$ is represented as

$$s = \begin{bmatrix} I \\ Q_{21}Q_{11}^{-1} \end{bmatrix} \ell \qquad \text{for some } \ell \in R^n.$$

We know that interpolating splines are also the solution of problem (I):

Given $x \in V$, find s such that

$$\|x - s\| = \text{minimum}$$

subject to

$$s \in S$$

Let us verify this directly. The problem is restated as follows

$$(x - s)^T P (x - s) = \text{minimum}$$

subject to

$$s = \begin{bmatrix} I \\ Q_{21}Q_{11}^{-1} \end{bmatrix} \ell$$

This is formally identical to an adjustment problem by parameters with observations x, adjusted observations s and parameters ℓ. The solution is obtained via the normal equations

$$(I \quad Q_{11}^{-1}Q_{12}) \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \cdot \begin{bmatrix} I \\ Q_{21}Q_{11}^{-1} \end{bmatrix} \ell = (I \quad Q_{11}^{-1}Q_{12}) \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Rewriting this as

$$Q_{11}^{-1}(Q_{11} \quad Q_{12})Q^{-1} \begin{bmatrix} Q_{11} \\ Q_{12} \end{bmatrix} Q_{11}^{-1}\ell = Q_{11}^{-1}(Q_{11} \quad Q_{12})Q^{-1}x$$

or

$$Q_{11}^{-1}Q_{11}Q_{11}^{-1}\ell = Q_{11}^{-1}x_1$$

i.e.

$$\ell = x_1$$

which was to be shown!

6.7. Noise-free collocation with trend parameters.

Consider the problem

Given $\ell \in R^n$, find $x \in R^m$, $y \in R^p$ such that

$$\ell = A x + B y \quad \text{and} \quad y^T P y = \text{minimum}$$

The vector ℓ is called observation vector. (Its coordinates are linear functionals). The vector $x \in R^m$ comprises the trend parameters. The vector $B y$ is sometimes called "the signal".

We put

$$V = R^{m+p}$$

such that for $z \in V$ we have

$$z = \begin{bmatrix} x \\ y \end{bmatrix}$$

The operator Λ is given by

$$\ell = \Lambda(z) = (A \quad B) \begin{bmatrix} x \\ y \end{bmatrix}$$

The space R^p of signals is taken as the space W. Its inner product is given by the matrix P.

The operator θ is taken as

$$w = \theta(z) = (0, \ I)\ z \ = \ y$$

Our above stated extremum problem is precisely problem (II) of section 6.4. and section 6.1.

Given $\ell \in R^n$, find $z \in V$ such that

$$\|\theta(z)\|_W = \text{minimum}$$

subject to

$$\Lambda(z) = \ell$$

The special problem of this section is also formally equivalent to the socalled "general adjustment problem". Let us solve the problem in the familiar way by means of Lagrange multipliers. Lagrange's function is

$$\Phi(x,y) = y^T P\, y \ - \ 2\, \lambda^T (A\, x \ + \ B\, y \ - \ \ell)$$

Thus

$$\frac{1}{2}\frac{\partial \Phi}{\partial y} = P\, y \ - \ B^T \lambda \ = \ 0$$

$$\frac{1}{2}\frac{\partial \Phi}{\partial x} = A^T \lambda \qquad = \ 0$$

The first of these equations gives

$$y \ = \ P^{-1} B^T \lambda$$

Inserting into $A\,x + B\,y = \ell$ we get the system

$$\begin{aligned} B\,P^{-1}B^T\lambda + A\,x &= \ell \\ A^T\lambda &= 0 \end{aligned}$$

Solving for x by partial reduction we find the system and its solution

$$A^T(B\,P^{-1}B^T)^{-1}A\,x = A^T(B\,P^{-1}B^T)^{-1}\ell$$

$$\boxed{x = (A^T(B\,P^{-1}B^T)^{-1}A)^{-1}A^T(B\,P^{-1}B^T)^{-1}\ell}$$

From the first equation one finds

$$\lambda = (B\,P^{-1}B^T)^{-1}(\ell - Ax)$$

From $y = P^{-1}B^T\lambda$ one finds

$$\boxed{y = P^{-1}B^T(B\,P^{-1}B^T)^{-1}(\ell - Ax)}$$

These are the formulas for noise free collocation with trend parameters x. Note that the usual notation is

$$s = B\,y$$

$$\Sigma_{ss} = B\,P^{-1}B^T \qquad \Sigma_{ys} = P^{-1}B^T$$

We prefer to introduce the notation

$$s = \begin{bmatrix} u \\ v \end{bmatrix}$$

for the solution of our problem. Given $\ell \in R^n$, we find u, v from the equations

$$\begin{aligned} B\,P^{-1}B^T\lambda + A\,u &= \ell \\ A^T\lambda &= 0 \\ v &= P^{-1}B^T\lambda \end{aligned}$$

Letting ℓ run through R^n, we obtain the set S of splines s. Let us check, whether also our abstract approach arrives at this set.

The null-space of Λ is given by the solutions of

$$A\,x + B\,y = 0$$

The space $U^\perp$ consists of all vectors w which are orthogonal to all y's which fulfill $A\,x + B\,y = 0$ together with a suitable x. Thus

$$A\,x + B\,y = 0$$

must imply

$$w^T P\,y = 0$$

It follows that the row vector

$$(0, \quad w^T P)$$

must be a linear combination of the rows of $(A, \quad B)$:

$$(0, \quad w^T P) \quad = \quad \lambda^T (A, \quad B)$$

or

$$A^T \lambda \quad = \quad 0$$

$$w \quad = \quad P^{-1} B^T \lambda$$

The pre-images of such w's are the splines. Thus the set S of splines is given by

$$s = \begin{bmatrix} u \\ P^{-1} B^T \lambda \end{bmatrix} \qquad \text{with u arbitrary and} \quad A^T \lambda = 0$$

Let us solve the problem (I) of section 6.4:

Given

$$z \quad = \quad \begin{bmatrix} x \\ y \end{bmatrix} \quad \varepsilon \quad V$$

find $s \varepsilon S$ such that

$$\| z - s \|_w \quad = \quad \text{minimum}$$

Using the above representation for s, as well as the definition of the inner product in W, we get

$$(y - P^{-1}B^T\lambda)^T P (y - P^{-1}B^T\lambda) = \text{minimum}$$

subject to

$$A^T\lambda = 0$$

This problem is formally equivalent to an adjustment problem by parameters which fulfill additional conditions. (Observations ... y, parameters ... λ, adjusted observations $P^{-1}B^T\lambda$, conditions $A^T\lambda = 0$).

Introducing the Lagrangean

$$\Phi(\lambda,\mu) = (y - P^{-1}B^T\lambda)^T P (y - P^{-1}B^T\lambda) + 2\,\mu^T A^T\lambda$$

we find

$$B\,P^{-1}B^T\lambda + A\,\mu = B\,y$$
$$A^T\lambda = 0$$

If we identify $\ell = B\,y$ and $\mu = x$, we arrive at the earlier equations. This concludes the successful verification of the equivalence of problem (I) and problem (II).

7. Approximation with splines.

7.1. Introduction.

Frequently we want that our data are not exactly interpolated but only approximated. Our spline functions shall not precisely match the data; there will be residuals or discrepancies. Data may be distributed irregularly. In the 1-dimensional case we can design spline functions matching irregular data exactly. However, even in the 1-dimensional case this is not always desirable. The data may be noisy, and we want that our approximating functions filter out some of the noise. In the two- or higher dimensional case it is virtually impossible to interpolate irregular data by bi- or n-cubic splines.

7.2. Approximation in one dimension.

Let (ξ_k,η_k), k=1,...,N denote the data. This means that at the locations ξ_k, function values $\eta_k = f(\xi_k)$ are prescribed. Let x_i, i=0,...,n denote the nodes of the spline s(x), i.e. the location of the discontinuities of its third derivative. The x_i generally do not coincide with the ξ_k. Occasional coincidence is, however, not excluded. Also the number of nodes n is typically less than N, the number of data. Let y_i denote the ordinates of the spline s(x) at the locations x_i. The y_i are now unknown. The approximating spline is parameterized in terms of y_i and y'_i, i=0,...,n. Thus we have 2n+2 unknowns. The equations of section 4.3, enforcing continuity of the second derivatives at x_i, i=1,...,n-1, represent a set of n-1 constraints among x_i, y_i. Additional boundary constraints at x_0 and x_n may or may not be prescribed.

The discrepancy $v_i = -[\eta_i - s(\xi_i)]$ may be linearly expressed in terms of y_i, y'_i. This can be done by means of the formula in section 4.2 giving the coefficients $a^{(i,i+1)}$ as linear functions of y_i, y'_i. Then one has:

$$v_k = -[\eta_k - \sum_{l=0}^{3} a_l^{(i,i+1)} \xi_k^l]$$

The size of the discrepancies can be measured by their squared norm

$$\|v\|^2 = \sum_{k=1}^{N} v_k^2$$

This expression can be minimized subject to the formulated constraints. We obtain a mixed adjustment model. It is the model of variation of parameters with additional constraints. It is a feasible way to obtain an approximating spline. A better way is outlined in the next section.

7.3. Basis splines with local support.

This remarkable type of spline is already described in Schoenberg (1946). We assume equidistant data with

$$x_i = i, \quad i=0,\ldots,n$$

We start with the function B(x) defined as follows:

$$B(x) = \begin{cases} 0 & \ldots \quad x \leq -2 \\ \frac{1}{6}(x+2)^3 & \ldots \quad -2 \leq x \leq -1 \\ \frac{1}{6}(x+2)^3 - \frac{4}{6}(x+1)^3 & \ldots \quad -1 \leq x \leq 0 \\ \frac{1}{6}(-x+2)^3 - \frac{4}{6}(-x+1)^3 & \ldots \quad 0 \leq x \leq 1 \\ \frac{1}{6}(-x+2)^3 & \ldots \quad 1 \leq x \leq 2 \\ 0 & \ldots \quad 2 \leq x \end{cases}$$

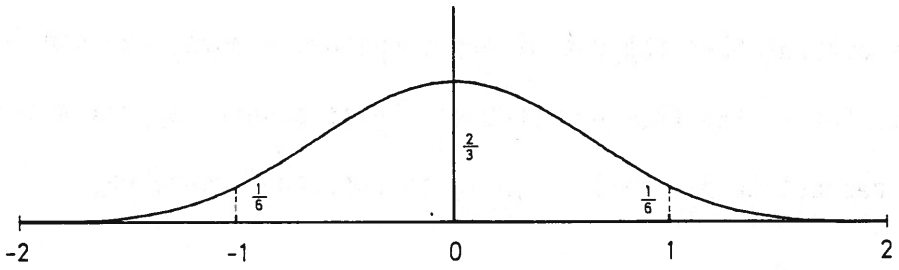


Fig. 7.1 Basis spline B(x)

The function B(x) is shown in figure 7.1. It is a spline. Its second derivative is continuous. The decisive property of B(x) is its local support. It vanishes outside a finite interval, namely the interval $-2 \leq x \leq 2$.

We now add two auxiliary nodes of $x_{-1} = -1$ and $x_{n+1} = n+1$. We associate with any node x_i, $i=-1,0,\ldots,n,n+1$, a basis spline $B_i(x)$ by shifting the function B(x):

$$B_i(x) = B(x-x_i) = B(x-i), \qquad i=-1,\ldots,n+1$$

We consider a linear combination of the basis splines

$$s(x) = \sum_{i=-1}^{n+1} \beta_i B_i(x)$$

Obviously s(x) is a spline. It is parameterized in terms of its coordinates β_i with respect to the basis built up of the B_i's. (Splines with nodes x_i, $i=0,\ldots,n$, form an n+3-dimensional vector space. The B_i, $i=-1,\ldots,n+1$, are a basis. This basis is different from that one specified in section 4.8).

It becomes obvious that the use of basis splines removes the continuity constraints for s''(x) from our problem. Least squares approximation may now be done with respect to β_i, $i=-1,\ldots,n+1$, in straightforward way.

We could have used the basis splines defined in section 4.8 for our approximation procedure. Theoretically this is sound; practically, it is not. The basis splines $a_i(x)$, $i=0,\ldots,n$, $\alpha_0(x)$, $\alpha_n(x)$ introduced in section 4.8 do not have local support.

Consequently, the normal equations of the adjustment problem are not sparse. The splines $B_i(x)$ have local support. Hence the normal equations are sparse and can be solved much faster. (The computational effort is $O(n)$ versus $O(n^3)$ in the case of a full system). The basis splines of section 4.8 are good for exact interpolation. For approximation, the $B_i(x)$ are much better.

7.3. Two dimensions.

We assume a grid of nm square meshes of unit side length. The basic two-dimensional spline is

$$B(x,y) = B(x)\ B(y)$$

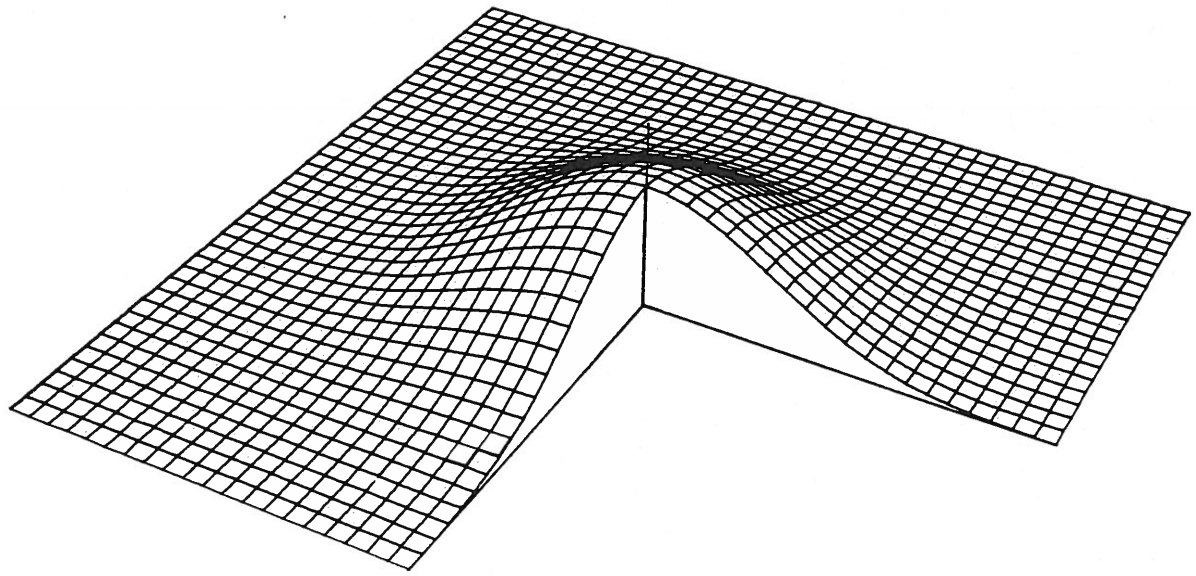


Fig. 7.2

Fig 7.2 shows this function. A complete basis is again obtained by shifting

$$B_{ij}(x,y) = B(x-i,y-j) = B(x-i)\,B(y-j)$$

$$i=-1,\ldots,m+1; \quad j=-1,\ldots,n+1$$

Using these basis functions, a spline is represented as

$$s(x,y) = \sum_{i=-1}^{m+1} \sum_{j=-1}^{n+1} \beta_{ij} B_{ij}(x,y)$$

The approximation problem with irregular data can be solved routinely. The normal equations will be sparse. The computational effort to solve this system is $O\{(nm)^{1.5}\}$ as opposed to $O\{(nm)^{3}\}$ if basis functions not having a local support are used.

MITTEILUNGEN DER GEODÄTISCHEN INSTITUTE
DER TECHNISCHEN UNIVERSITÄT GRAZ

(früher: Institutsmitteilungen des I. Geodätischen Institutes
der Technischen Hochschule in Graz)

Bisher erschienene Folgen:

Nr., Autor/Herausgeber, Titel, Seiten, Preis (ö.S.)

1 HUBENY K. (1959): Formeln und Tafeln zur Berechnung der 2. Hauptaufgabe auf dem Bessel'schen Ellipsoid für Strecken bis 300 km im Bereich zwischen den geographischen Breiten 45° und 57°30'. 33 Seiten, öS 45.

2 HUBENY K. (1959): Formeln und Tafeln zur Berechnung der geodätischen Hauptaufgaben über Normalschnitte für beliebige Ellipsoide und beliebige Entfernungen. 29 Seiten, öS 45.

3 MORITZ H. (1959): Untersuchungen über eine direkte Lösung der 2. Hauptaufgabe auf dem Rotationsellipsoid für beliebige Entfernungen. 16 Seiten, öS 35.

4 HUBENY K. (1960): Die Lösung der geodätischen Hauptaufgaben nach Bessel-Jordan. Erweiterte und neue Formeln sowie Tafeln für die Ellipsoide von Bessel und Hayford im Bereich zwischen den geographischen Breiten 45° und 58°. 65 Seiten, öS 50.

5 HUBENY K. (1960): Formeln und Tafeln zur Berechnung der geodätischen Hauptaufgaben über beliebige Entfernungen (Internationales Ellipsoid). 44 Seiten, öS 50.

6 HUBENY K., K. RINNER (1966): Vorlesungen am II. Fortbildungskurs für Praktiker an der Technischen Hochschule in Graz vom 5. bis 7. Oktober 1964. 156 Seiten, öS 50.

7 RINNER K. (1967): Geodätische Programme im Rechenzentrum Graz (Stand 9. Oktober 1967). 191 Seiten, öS 50.

8 RINNER K. (1968): Vorlesungen am III. Fortbildungskurs für Praktiker an der Technischen Hochschule in Graz vom 9. bis 12. Oktober 1967. 282 Seiten, öS 50.

9 RINNER K., G. BRANDSTÄTTER (1971): Forschungsberichte über Erdgezeiten und Satellitengeodäsie. 121 Seiten, öS 50.

10 FELDBACHER F., K. HUBENY, K. RINNER (1971): Beiträge zur ellipsoidischen Geometrie und zu Mikrowellen- und Lasermessungen für große Entfernungen. 78 Seiten, öS 50.

11 RINNER K. (1972): Proceedings of the International Symposium "Satellite and Terrestrial Triangulation"; 2 volumes: 1) Sessions of the West European Sub-Commission of the International Commission for Artificial Satellites, I.A.G.; 2) Sessions of the Special Study Group 1.26 of the I.A.G. 612 Seiten, öS 300.

12 BITTMANN O., G. KRAJICEK, P. MEISSL (1973): Microcomputer Compucorp 320 G und 322 G, die Benützung und Anwendungsbeispiele für die Vermessungstechnik. 53 Seiten, öS 30.

13 RINNER K. (1973): Berichte über Forschungsarbeiten. 57 Seiten. öS 50.

14 FRIEDL J., G. KRAJICEK, P. MEISSL (1974): Taschenrechner Hewlett-Packard HP-45, die Benützung und Anwendungsbeispiele für die Vermessungstechnik. 60 Seiten, öS 50.

15 BARTELME N., P. MEISSL (1974): Strength Analysis of Distance Networks. 57 Seiten, öS 50.

16 CHESI G., K. RINNER (1974): Tabellen zur meteorologischen Reduktion von Entfernungsmessungen mit dem Geodimeter 8. 100 Seiten, öS 100.

17 BITTMANN O., P. MEISSL (1974): Empfohlene Algorithmen zur Programmierung geodätischer Rechenaufgaben. I. Einfache Koordinatenrechnungen in der Ebene, 35 Seiten, öS 50.

18 MEISSL P., K. RINNER (1975): Vorträge am IV. Fortbildungskurs für Praktiker des Vermessungswesens an der Technischen Universität in Graz vom 25. bis 27. November 1974. 290 Seiten, öS 130.

19 LACHAPELLE G. (1975): Determination of the Geoid using Heteregeneous Data. 121 Seiten, öS 150.

20 MEISSL P., H. MORITZ, K. RINNER (1975): Contributions of the Graz Group to the XVI. General Assembly of IUGG/IAG in Grenoble. 308 Seiten, öS 150.

21 BENZ F., K. RINNER (1976): Verfahren zur Verminderung des Einflusses der Bodenreflexion bei der Entfernungsmessung mit Mikrowellen. 97 Seiten, öS 100.

22 RINNER K. (1976): Bericht über Laser- und Mikrowellenmessungen im Testnetz Steiermark. 109 Seiten, öS 100.

23 RINNER K. (1976): Bericht zur Meeresgeodäsie und Satellitengeodäsie. 111 Seiten, öS 100.

24 MEISSL P. (1976): Empfohlene Algorithmen zur Programmierung geodätischer Rechenaufgaben. II. Punktverwaltung mittels Massenspeicher. 69 Seiten, öS 50.

25 MEISSL P., K. STUBENVOLL (1977): Ein Computer-Programmsystem zur Verdichtung trigonometrischer Netze. 129 Seiten, öS 100.

26 KRYNSKI J., H. NOE, K.P. SCHWARZ, H. SONKEL (1977): Numerical Studies and Programs for Interpolation and Collocation. 67 Seiten, öS 50.

27 HUBENY K., A. REITHOFER (1977): Isotherme Koordinatensysteme und konforme Abbildungen des Rotationsellipsoides mit Tafeln und Programmen zur konformen Abbildung für die Ellipsoide von Bessel, Hayford, Krasowsky und für das Referenzellipsoid 1967. 222 Seiten, öS 200.

28 SONKEL H. (1977): Die Darstellung geodätischer Integralformeln durch bikubische Spline-Funktionen. 161 Seiten, öS 150.

29 LEBERL F. (1977): Proceedings of the International Symposium on Image Processing - Interactions with Photogrammetry and Remote Sensing, 3-5 October, 1977, Graz. 250 Seiten, öS 160.

30 ALLMER F. (1977): Dr. Ing. h.c. Eduard Ritter von Orel, dem Erfinder des Stereo-Autographen zum 100. Geburtstag. 41 Seiten, öS 50.

31 KRYNSKY J. (1978): Possibilities of Low-low Satellite Tracking for Local Geoid Improvement. 67 Seiten, öS 70.

32 GERONTOPOULOS P. (1978): Molodensky's Problem in the Plane. 160 Seiten, öS 200.

33 LEBERL F. (1980): Beiträge zur Radargrammetrie und digitalen Bildverarbeitung. 230 Seiten, öS 200.

34 HUBENY K. (1980): Die Klothoide (Formeln, Tafeln, Beispiele). 122 Seiten, öS 120.

35 RINNER K. et al. (1980): Festschrift zur Emeritierung von Prof. Dr. K. Hubeny. 200 Seiten, öS 200.

36 NOE H. (1980): Numerical Investigations on the Problem of Molodensky. 80 Seiten, öS 90.

37 BARTELME N., B. HOFMANN-WELLENHOF, P. MEISSL (1980): Empfohlene Algorithmen zur Programmierung geodätischer Rechenaufgaben. III. Zugriff auf Meßdatendatei. 200 Seiten, öS 180.

38 LICHTENEGGER H., K. RINNER (1982): Verzeichnis der Habilitationen, Dissertationen, Diplomarbeiten sowie von Seminar- und Proseminararbeiten 1960-1981. 81 Seiten, öS 100.

39 CHEN C.-Y. (1982): Geodetic Datum and Doppler Positioning (Dissertation). 255 Seiten, öS S 200.

40 MORITZ H. et al. (1982): Geodaesia Universalis. Festschrift Karl Rinner zum 70. Geburtstag. 382 Seiten, öS 250.

41 MORITZ H., H. SONKEL (1982): Geodesy and Global Geodynamics. 689 Seiten, öS 400.

42 RINNER K., H. LICHTENEGGER (1982): Proceedings of the International Symposium "Education in Geodesy", Sept. 27-29, 1982, Graz, Austria.

43 MEISSL P. (1982): Least Squares Adjustment. A Modern Approach. 440 Seiten, öS 450.

44 ALLMER F. (1983): Biographie Peter Meissl (in Vorbereitung -in preparation).