Least-Squares Approximation

Methods for discrete data

1. Linear Least-Squares

The interpolation method of collocation by high degree polynomials normally runs into oscillation problems for (rather large) sets of measurement points with arguments $x_0, x_1, \ldots, x_{N-1}, x_N$ or $\vec{x}_0, \vec{x}_1, \ldots, \vec{x}_{N-1}, \vec{x}_N$ in the multi-variate case. Oscillation problems normally occur at the boundaries of the interval $(\min x_i, \max x_i)_{i=0,1,\ldots,N}$ and extrapolation (going beyond the boundaries) may not be applicable.

Moreover, in many cases interpolation may not be the appropriate method when the data arises from experiments that contain errors of a random nature. In such cases approximation methods may be more appropriate. These methods just approximate the data by a "combination" of a rather small set of basis functions $g_0, g_1, \ldots, g_{m-1}, g_m$ ($m << N$) and there is a need to judge how "good" the approximation is. In many cases such a judgement consists of minimizing a function measuring the Δy deviations (residuals) between the approximation and the data values. The important method of least-squares approximation uses the sum of squared residuals. Another important method is Chebyshev-approximation; it uses the maximal deviation (residual) to be minimized and therefore is called min-max-approximation.

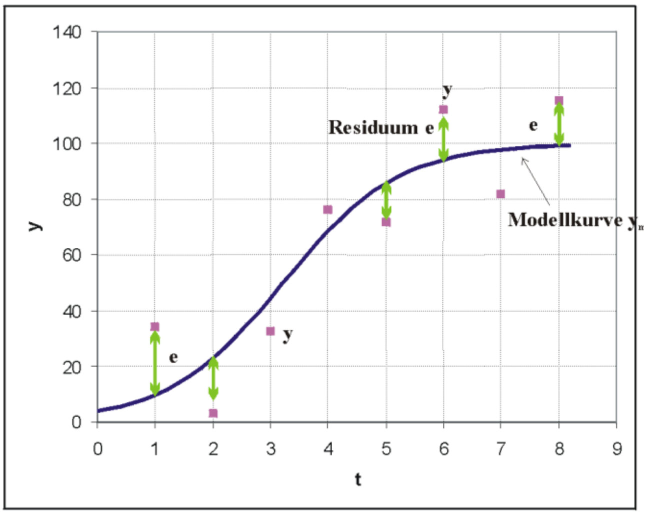

Figure 1.1: Data approximated by a modelling curve with residuals.

Widely used basis functions are polynomials, trigonometric functions, exponential functions or combinations of these.

Type of basis functions	Formulas
Trigonometric	$\{e^{ikx} \mid k \in Z\}$ or $\{\cos(kx), \sin(kx) \mid k \in Z\}$ $\quad (x \in (0, 2\pi))$
Polynomials	$\{x^j \mid j \in N_0\}$ (standard monomials) $\left\{ \left(\frac{x-\mu}{\sigma} \right)^j \middle| j \in N_0 \right\}$ (normalized standard monomials: μ and σ denote mean and standard deviation of the x-data. $\{\pi_j(x) \mid j \in N_0\}$ (Newton polynomials) $\{T_j(x) \mid j \in N_0\}$ (Chebyshev polynomials, $x \in (-1,1)$)

Polynomials	$\left\{ L_j(x) = \frac{e^x}{j!} \frac{d^j}{dx^j}\left(e^{-x} x^j\right) \middle\| j \in N_0 \right\}$ (Laguerre polynomials, $x \in (0,\infty)$) $\left\{ H_j(x) = (-1)^j e^{\frac{x^2}{2}} \frac{d^j}{dx^j}\left(e^{-\frac{x^2}{2}}\right) \middle\| j \in N_0 \right\}$ (Hermite polynomials, $x \in (-\infty,\infty)$)
Exponentials	$\left\{ e^{\alpha k x} \middle\| k \in Z \right\}$

Table 1.1: Some widely used uni-variate basis functions for least-squares approximation. Normalizing the standard monomials improves numeric stability significantly in many applications.

Example 1.1: Data smoothing by a Savitzky-Golay Filter

The data

```
{ {1,1.04},{2,1.37},{3,1.70},{4,2.00},{5,2.26},
  {6,2.42},{7,2.70},{8,2.78},{9,3.00},{10,3.14} }
```

consisting of 10 points is smoothed by six parabolas fitting only 5 (consecutive points). Figure 1.2 below shows the fitting least-squares parabola for the five points with x-arguments {3, 4, 5, 6, 7}.

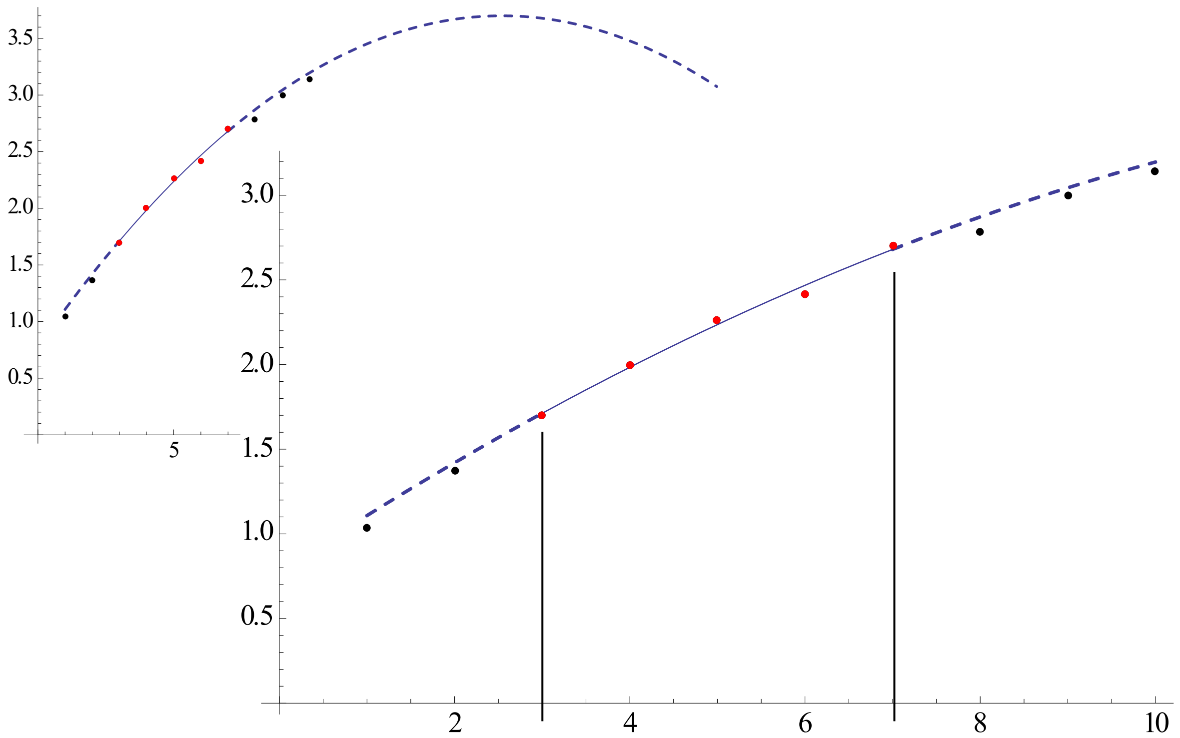

Figure 1.2: Parabola fitting only the 5 red points with x-arguments {3, 4, 5, 6, 7}. The parabola does not fit the whole data. A possible set of basis functions is {1, x, x^2}. The parabola formula is $0.776 + 0.342\,x - 0.01\,x^2$.

1.1 Normal equations and design matrix: A general framework

Given a measurement $(x_0, y_0), (x_1, y_1)_1, \ldots, (x_N, y_N) = \{(x_i, y_i)\}_{i=0,\ldots,N}$ or (in the multi-variate arguments case) $(\vec{x}_0, y_0), (\vec{x}_1, y_1)_1, \ldots, (\vec{x}_N, y_N) = \{(\vec{x}_i, y_i)\}_{i=0,\ldots,N}$ and a set of basis functions $g_0, g_1, \ldots, g_m = \{g_j\}_{j=0,\ldots,m}$ the <u>design matrix</u> G results from the trial to collocate all measurement points with a linear combination $\sum_{j=0}^{m} a_j g_j \quad (a_j \in R)$ of the basis functions. This requires that $y_i = \sum_{j=0}^{m} a_j g_j(x_i) \quad (i = 0, \ldots, N)$ and thus yields linear system of $N+1$ equations in $m+1$ unknowns $a_j \quad (j = 0, \ldots, m)$. Written in matrix form we get that:

$$\overbrace{\begin{pmatrix} g_0(x_0) & g_1(x_0) & \cdots & g_m(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ g_0(x_N) & g_1(x_N) & \cdots & g_m(x_N) \end{pmatrix}}^{\text{Design matrix } G} \cdot \begin{pmatrix} a_0 \\ \\ a_m \end{pmatrix} = \begin{pmatrix} y_0 \\ \\ y_N \end{pmatrix} \qquad \textbf{(1.1)}$$

Generally, the linear system (1.1) is overdetermined (m < N). Thus, in general, the residuals

$$r_i = y_i - \sum_{j=0}^{m} a_j g_j(x_i) \quad (i = 0, \ldots, N) \qquad \textbf{(1.2)}$$

are not all zero and the <u>squared sum S of residuals</u>

$$S = \sum_{i=0}^{N} \left(y_i - \sum_{j=0}^{m} a_j g_j(x_i) \right)^2 \quad \text{min!} \qquad \textbf{(1.2')}$$

has to be <u>minimized</u>.

Theorem 1.1: Normal equations

The squared sum of residuals S in (1.2') is minimal if and only if $\boxed{G^{tr} \cdot G \cdot a| = \quad G^{tr} \cdot y|}$ **(1.3)**

The symbols $a|$ and $y|$, respectively, are abbreviations for the column vectors formed by $a_j \quad (j = 0, \ldots, m)$ and $y_i \quad (i = 0, \ldots, N)$, respectively. G^{tr} stands for the transposed design matrix.

(<u>Proof 1</u>: Formally, (1.3) results from (1.1) by multiplication of G^{tr} on both sides. We use multivariate calculus to proof the theorem. A quadratic form like (1.2) is globally minimal if and only if the partial derivatives $\frac{\partial S}{\partial a_k} \quad (k = 0, \ldots, m)$ are 0. Computing the derivatives and interchanging the summation order yields:

$$\frac{\partial S}{\partial a_k}=\frac{\partial}{\partial a_j}\sum_{i=0}^{N}\left(y_i-\sum_{j=0}^{m}a_j g_j(x_i)\right)^2=\sum_{i=0}^{N}2\left(y_i-\sum_{j=0}^{m}a_j g_j(x_i)\right)(-g_k(x_i)=0\Leftrightarrow$$

$$\sum_{j=0}^{m}\left(\underbrace{\sum_{i=0}^{N}g_j(x_i)g_k(x_i)}_{<g_j,g_k>}\right)a_j=\underbrace{\sum_{i=0}^{N}y_i g_k(x_i)}_{<y|,g_k>}\qquad(k=0,\dots,m)$$

The last system of m+1 equations is exactly (1.3)

□)

(Proof 2: The following is a general geometric reasoning. In order to minimize the square sum of residuals (1.2) we are to find a vector $y^*| = \sum_{j=0}^{m} a_j\, g_j|$ in the vector space V spanned by the basis vectors $\{g_j|\}_{j=0,\dots,m}$ such that the Euclidean norm (distance) $S=\|y-y^*\|^2$ is minimal. A solution for this is the orthogonal projection of y onto V:

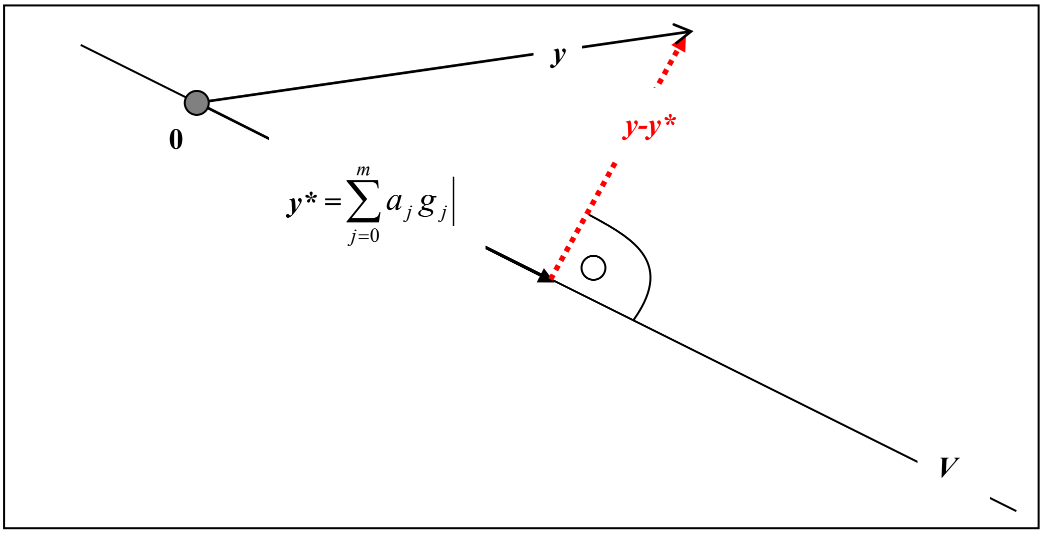

Figure 1.3: The orthogonal projection y^* is nearest to y .

Due to the orthogonality we get that

$$0=\ <y|-y^*|,g_k|>\ =\ <y|-\sum_{j=0}^{m}a_j\, g_j|,g_k|>\ \Leftrightarrow$$

$$0=\ <y|,g_k|>-\sum_{j=0}^{m}a_j<g_j|,g_k|>\ \Leftrightarrow\ \sum_{j=0}^{m}a_j<g_j|,g_k|>\ =\ <y|,g_k|>\quad(k=0,\dots,m)$$

The last equations are the same as those at the end in the first proof

□)

The equations (1.3) are called <u>normal equations</u>. The matrix product $G^{tr} \cdot G$ can be written as the matrix of all <u>inner products</u> $< g_j|, g_k| >:= \sum_{i=0}^{N} g_j(x_i) g_k(x_i) \quad (j,k = 0,...,m)$ of the column vectors of the design matrix. The right hand side of (1.3) then is the column vector formed by the inner products $< y|, g_k| >:= \sum_{i=0}^{N} y_i g_k(x_i) \quad (k = 0,...,m)$. Written in matrix form we get the following representation of (1.3):

$$\overbrace{\begin{pmatrix} < g_0|, g_0| > & < g_1|, g_0| > & \cdots & < g_m|, g_0| > \\ \vdots & \vdots & \ddots & \vdots \\ < g_0|, g_m| > & < g_1|, g_m| > & \cdots & < g_m|, g_m| > \end{pmatrix}}^{G^{tr}\,G} \cdot \begin{pmatrix} a_0 \\ \\ a_m \end{pmatrix} = \begin{pmatrix} < y|, g_0| > \\ \\ < y|, g_m| > \end{pmatrix} \tag{1.3'}$$

Obviously, the matrix $G^{tr} \cdot G$ is <u>symmetric</u> and has dimensions $(m+1)\times(m+1)$. If the basis functions $g_0, g_1, ..., g_m = \{g_j\}_{j=0,...,m}$ are a <u>linearly independent</u> set of functions, then $G^{tr} \cdot G$ is <u>positive definite</u> and has <u>full rank</u>. The linear system (1.3) then is regular and has a unique solution for the coefficients $a_j \quad (j = 0,...,m)$.

Example 1.2: Cont. Ex. 1.1

The data set of Example 1.1 is `{{3,1.70},{4,2.00},{5,2.26}, {6,2.42},{7,2.70}}`. The set of basis functions {1, x, x^2}. The system (1.1) has the following form:

$$\overbrace{\begin{pmatrix} 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \\ 1 & 7 & 49 \end{pmatrix}}^{\text{Design matrix } G} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1.70 \\ 2.00 \\ 2.26 \\ 2.42 \\ 2.70 \end{pmatrix}$$

The system of normal equations (1.3') is:

$$\overbrace{\begin{pmatrix} 5 & 25 & 135 \\ 25 & 135 & 775 \\ 135 & 775 & 4659 \end{pmatrix}}^{G^{tr}G} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 11.08 \\ 57.82 \\ 323.22 \end{pmatrix}$$

Generally, the system (1.3) should not be solved directly[1] for $a_j \quad (j = 0,...,m)$ because it does not have a „good" structure (sparse or band structure, e.g.) and because the relative error of the solution is sensitive to changes (errors) in the matrix entries or changes in the right hand side[2]. The best thing to happen with respect to (1.3) is a diagonal matrix $G^{tr} \cdot G$. In this case the column vectors of the design matrix G are orthogonal and the basis functions $g_0, g_1,..., g_m = \{g_j\}_{j=0,...,m}$ form an orthogonal set with respect to the arguments $x_0, x_1, \ldots, x_{N-1}, x_N$ [3].

[1] E.g., by Cholesky decomposition due to symmetry and positive definiteness.

[2] Such a system is called ill-conditioned or said to have a bad condition number.

[3] A theoretically important method to iteratively transform a set of basis vectors or functions into an orthogonal basis is the Gram-Schmidt orthogonalization. E.g., starting with the basis of Newton polynomials $\pi_i(x) \quad (i = 0,1,..., N)$ with respect to the arguments {0, 1, …, N} the Gram-Schmidt orthogonalization yields the polynomials $p_{k,N}(x) = \sum_{i=0}^{k} (-1)^i \binom{k}{i} \binom{k+i}{i} \frac{x^{(i)}}{N^{(i)}}$ with $N^{(i)} = N(N-1)(N-2)\cdots(N-i+1)$ and $x^{(i)} = x(x-1)(x-2)\cdots(x-i+1)$ defined as falling factorials. The Gram-Schmidt procedure is numerically not stable and has to be modified or replaced by other methods like Givens rotations or Householder transformations.

1.2 The QR-matrix-decomposition

A square real $n \times n$-matrix Q is called orthogonal or $Q \in O(n)$ if $Q^{tr} \cdot Q = Q \cdot Q^{tr} = I$ (the identitiy matrix = diagonal matrix with 1s in the diagonal). This is equivalent to saying that $Q^{tr} = Q^{-1}$: The inverse matrix is equal to the transpose and, of course, is also orthogonal.

An orthogonal matrix represents an isometry: It preserves length of vectors and angles between vectors:

$$Q \in O(n) \Leftrightarrow \langle x, y \rangle = \langle Qx, Qy \rangle \text{ for any vectors } x, y$$

Theorem 1.2: QR-Decomposition

Any matrix G of dimensions ($(N+1)\times(m+1)$) with N ￿ m with full rank m+1 can be decomposed as the product QR of an orthogonal $(N+1)\times(N+1)$-matrix Q and a rectangular upper-triangle $(N+1)\times(m+1)$ matrix R:

$$G = QR = Q\overbrace{\begin{pmatrix} \overbrace{\begin{matrix} r_{00} & r_{01} & \cdots & r_{0m} \\ 0 & r_{11} & \cdots & r_{1m} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{mm} \end{matrix}}^{\text{upper square triangle}} \\ \begin{matrix} - & - & - & - \\ 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{matrix} \end{pmatrix}}^{R}. \qquad (1.4)$$

(Proof: A constructive proof of this yielding a numerically stable procedure for determing Q and R can be found at http://en.wikipedia.org/wiki/QR_decomposition#Using_Householder_reflections or in the book by Schwarz/Köckler, Numerische Mathematik, 7th ed., Chapter 6.2.1|3 ￿)

With the aid of the decomposition of Theorem 1.2 the solution of (1.2) and (1.2'), i.e. minimization of the square sum of residuals, can be found in a numerically stable and efficient way. We apply Theorem 1.2 to the design matrix G in (1.1) and get $Ga| = QRa| = y| \Rightarrow Ra| = Q^{-1}y| = Q^{tr}y|$ and thus $Ra| = Q^{tr}y|$. Abbreviating $Q^{tr}y|$ by $\hat{y}|$ and $Q^{tr}r|$ by $\hat{r}|$ the system (1.2) translates into $\hat{r}| = \hat{y}| - Ra|$. In matrix notation this reads as follows:

$$
\begin{array}{lcllll}
\hat{r}_0 & = & \hat{y}_0 - r_{00}a_0 - & r_{01}a_1 & -\cdots- & r_{0m}a_m \\
\hat{r}_1 & = & \hat{y}_1 - 0 & r_{11}a_1 & -\cdots- & r_{1m}a_m \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\
\hat{r}_m & = & \hat{y}_m - 0 & \cdots & 0 & r_{mm}a_m \\
 & & & & & \\
\hat{r}_{m+1} & = & \hat{y}_{m+1} - 0 & \cdots & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\
\hat{r}_N & = & \hat{y}_N - 0 & \cdots & \cdots & 0
\end{array}
\qquad \textbf{(1.5)}
$$

Because of the orthogonality of Q (and thus also Q^{tr}) the square sum $\hat{S}$ of the residuals $\hat{r}|$ is equal to the square sum S of the original residuals $r|$. So minimizing S in (1.2) is equivalent to miminizing $\hat{S}$ in (1.5). But obviously $\hat{S}$ in (1.5) is minimal if and only if $\hat{r}_i = 0 \quad (i = 0,...,m)$. This yields a regular triangle $(m+1)\times(m+1)$-system of linear equations for the unknowns $a_j \quad (j = 0,...,m)$. It can be solved easily by backward-substitution beginning with solving for a_m. Then we have $S = \hat{S} = \hat{r}_{m+1}^{\;2} + \hat{r}_{m+2}^{\;2} + \cdots + \hat{r}_N^{\;2}$ as the minimal square sum of residuals. Moreover, the residuals $r|$ are determined by $Q\hat{r}|$.

1.3 Singular-value decomposition (SVD)

As the QR-decomposition the singular-value decomposition is a cleverly devised decomposition of a matrix useful for many operations involving matrices like solving linear systems, inverting matrices, finding eigenvalues and many more. Its computational costs are higher than those of the QR-decomposition but it is numerically even more stable. The singular-value decomposition is one of the most widely used matrix operations in applied linear algebra.

Theorem 1.3: SVD-Decomposition

Any real matrix G of dimensions $((N+1)\times(m+1))$ can be decomposed as the triple product $U D V^{tr}$ whereas U is an orthogonal $(N+1)\times(N+1)$-matrix, D is a $(N+1)\times(m+1)$ diagonal matrix and V again is orthogonal with dimensions $(m+1)\times(m+1)$.

$$G = UDV^{tr} = U \cdot \overbrace{\begin{pmatrix} \overbrace{\begin{matrix} d_{00} & 0 & \cdots & 0 \\ 0 & d_{11} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_{mm} \end{matrix}}^{diagonal} \\ \begin{matrix} - & - & - & - \\ 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{matrix} \end{pmatrix}}^{D} \cdot V^{tr} \tag{1.6}$$

(Proof: Cf. the book by Schwarz/Köckler, Numerische Mathematik, 7th ed., Chapter 6.3 □)

With the aid of the decomposition of Theorem 1.3 the solution of (1.2) and (1.2'), i.e. minimization of the square sum of residuals, can be found in a numerically stable and efficient way. We apply Theorem 1.3 to the design matrix G in (1.1) and get $Ga| = UDV^{tr} a| = y| \Rightarrow DV^{tr} a| = U^{-1} y| = U^{tr} y|$ and thus $DV^{tr} a| = U^{tr} y|$. Abbreviating $U^{tr} y|$ by $\hat{y}|$, $V^{tr} a|$ by $\hat{a}|$ and $U^{tr} r|$ by $\hat{r}|$ the system (1.2) translates into $\hat{r}| = \hat{y}| - D\hat{a}|$. In matrix notation this reads as follows:

$$\begin{array}{ccccccc} \hat{r}_0 & = & \hat{y}_0 - d_{00}\hat{a}_0 - & 0 & -\cdots- & 0 \\ \hat{r}_1 & = & \hat{y}_1 - 0 & d_{11}\hat{a}_1 & -\cdots- & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ \hat{r}_m & = & \hat{y}_m - 0 & \cdots & 0 & d_{mm}\hat{a}_m \\ & & & & & \\ \hat{r}_{m+1} & = & \hat{y}_{m+1} - 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \hat{r}_N & = & \hat{y}_N - 0 & \cdots & \cdots & 0 \end{array} \tag{1.7}$$

Because of the orthogonality of U (and thus also U^{tr}) the square sum $\widehat{S}$ of the residuals $\widehat{r}|$ is equal to the square sum S of the original residuals $r|$. So minimizing S in (1.2) is equivalent to miminizing $\widehat{S}$ in (1.7). But obviously $\widehat{S}$ in (1.7) is minimal if and only if $\widehat{r}_i = 0 \quad (i = 0,...,m)$. This yields a regular diagonal $(m+1)\times(m+1)$-system of linear equations for the unknowns $\widehat{a}_j \quad (j = 0,...,m)$. It can be solved directly due to its diagonal form. Then we have $S = \widehat{S} = \widehat{r}_{m+1}^{\,2} + \widehat{r}_{m+2}^{\,2} + \cdots + \widehat{r}_N^{\,2}$ as the minimal square sum of residuals.
In the final step we solve $V^{tr} a| = \widehat{a}|$ for the unknowns $a_j \quad (j = 0,...,m)$ by multiplication of V and get $a| = V \cdot \widehat{a}|$.

Example 1.3: Cont. Ex 1.2
The singular-value decomposition $\{U, D, V\}$ of G is[1]

$$\left\{\begin{pmatrix} -0.135724 & -0.624629 & -0.690747 & -0.0257609 & -0.337079 \\ -0.237802 & -0.537653 & 0.160648 & 0.280919 & 0.741388 \\ -0.368367 & -0.327557 & 0.492673 & -0.688191 & -0.201691 \\ -0.527419 & 0.0056597 & 0.305329 & 0.636669 & -0.472466 \\ -0.714958 & 0.461997 & -0.401383 & -0.203636 & 0.269848 \end{pmatrix},\right.$$

$$\left.\begin{pmatrix} 69.2244 & 0. & 0. \\ 0. & 2.63845 & 0. \\ 0. & 0. & 0.144857 \\ 0. & 0. & 0. \\ 0. & 0. & 0. \end{pmatrix}, \begin{pmatrix} -0.0286643 & -0.387418 & -0.921459 \\ -0.16424 & -0.907483 & 0.386651 \\ -0.986004 & 0.162424 & -0.0376172 \end{pmatrix}\right\}$$

For $\widehat{y}| = U^{tr} y|$, $\widehat{a}| = V^{tr} a|$ and the final coefficients $a| = V \cdot \widehat{a}|$ we get

$$\begin{pmatrix} -4.74558 \\ -1.61637 \\ -0.0843709 \\ -0.0463456 \\ 0.0391419 \end{pmatrix}, \begin{pmatrix} -0.0685536 \\ -0.61262 \\ -0.582441 \end{pmatrix} \text{ and } \begin{pmatrix} 0.776 \\ 0.342 \\ -0.01 \end{pmatrix}.$$

This is in accordance with Figure 1.2.

For $\widehat{r}| = U^{tr} r|$ and the residuals $r| = U \widehat{r}|$ we get:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ -0.0463456 \\ 0.0391419 \end{pmatrix} \text{ and } \begin{pmatrix} -0.012 \\ 0.016 \\ 0.024 \\ -0.048 \\ 0.02 \end{pmatrix}.$$

---

[1] The entries in the diagonal of D are the singular values. These are the roots of the m + 1 real non-negative eigenvalues of the symmetric (normal) matrix $G^{tr} \cdot G$ or $G \cdot G^{tr}$, respectively . The columns of U are, e.g., orthonormal eigenvectors of $G \cdot G^{tr}$ and those of V orthonormal eigenvectors of $G^{tr} \cdot G$.

1.4 Uniform arguments and orthogonal polynomials

In case of a set of basis functions $g_0, g_1, ..., g_m = \{g_j\}_{j=0,...,m}$ which is orthogonal with respect to the inner product $\langle g_j, g_k \rangle := \sum_{i=0}^{N} g_j(x_i) g_k(x_i)$ the l.h.s. matrix in (1.3') is diagonal and the solution of (1.3) is trivial:

$$a_j = \frac{< y|, g_j| >}{< g_j|, g_j| >} \quad (j = 0,1,...,m) \tag{1.8}$$

The minimal sum of squared residuals then is (cf. Figure 1.3):

$$S_{\min} = \|y - y^*\|^2 = \sum_{i=0}^{N} y_i^2 - \sum_{j=0}^{m} a_j^2 \underbrace{< g_j, g_j >}_{\|g_j\|^2} \tag{1.8'}$$

The next theorem gives a set of orthogonal polynomials for uniformly distributed arguments:

Theorem 1.4: Orthogonal polynomials for uniform arguments

If $\{x_0, x_1, \dots, x_{n-1}, x_N\} = \{x_0 + t \cdot h\}_{t=0,...,N}$ are uniformly distributed arguments of a measurement then putting $t = \frac{x - x_0}{h}$ the set of polynomials

$$p_{k,N}(t) = \sum_{i=0}^{k} (-1)^i \binom{k}{i} \binom{k+i}{i} \frac{t^{(i)}}{N^{(i)}} \quad (k = 0,1,..., N) \tag{1.9}$$

is orthogonal with respect to the inner product $\langle g, \tilde{g} \rangle := \sum_{i=0}^{N} g(x_i) \tilde{g}(x_i)$.

Here $N^{(i)} = N(N-1)(N-2)\cdots(N-i+1)$ and $t^{(i)} = t(t-1)(t-2)\cdots(t-i+1)$ are defined as falling factorials.

(Proof: Cf. Schaum's outline of numerical analysis for a detailed but non-easy proof, Probl. 21.26 ▯)

Example 1.4: Ex. 1.1 rev.

The five data points $\{(x_i, y_i)\}_{t=3,...,7}$ of Example 1.1 are `{{3,1.70},{4,2.00},{5,2.26},{6,2.42},{7,2.70}}`.

We have N = 4 and $t = x$ - 3. The first 3 orthogonal polynomials according to (1.8) are:

$\{p_{k,N}(t)\}_{k=0,...,2}$ = `{1, 1-t/2, 1-(3t)/2 + 1/2(-1+t)t}`

$\{p_{k,N}(x)\}_{k=0,...,2}$ = `{1, 1+(3-x)/2, 1-3/2(-3+x) + 1/2(-4+x)(-3+x)}`

Evaluation of the inner products $\langle g, \tilde{g} \rangle := \sum_{i=0}^{N} g(x_i)\tilde{g}(x_i) = \sum_{i=3}^{7} g(i)\tilde{g}(i)$ yields

$$\begin{pmatrix} 5 & 0 & 0 \\ 0 & \frac{5}{2} & 0 \\ 0 & 0 & \frac{7}{2} \end{pmatrix}$$

The system of normal equations (1.3) thus is diagonal on the l.h.s. The right hand side is

$$\begin{pmatrix} < y|, p_{0,N}| > \\ < y|, p_{1,N}| > \\ < y|, p_{2,N}| > \end{pmatrix} = \begin{pmatrix} \sum_{i=3}^{7} y_i p_{0,N}(i) \\ \sum_{i=3}^{7} y_i p_{1,N}(i) \\ \sum_{i=3}^{7} y_i p_{2,N}(i) \end{pmatrix} = \begin{pmatrix} 11.08 \\ -1.21 \\ -0.07 \end{pmatrix}$$

By (1.8) the solution of (1.3) thus is $a_k = \dfrac{< y|, p_{k,N}| >}{< p_{k,N}|, p_{k,N}| >} \quad (k = 0,1,2)$.

Numerically: $a_0 = \dfrac{11.08}{5} = 2.216, a_1 = \dfrac{-1.21}{5/2} = -0.484, a_2 = \dfrac{-0.07}{7/2} = -0.02$. These are the coefficients with respect to the basis $\{p_{k,N}(x)\}_{k=0,\dots,2}$. Expanding and simplifying the polynomials indeed yields the correct solution $0.776 + 0.342\,x - 0.01\,x^2$ written in standard form.

$$1$$

$$1 - \frac{t}{2}$$

$$1 - \frac{3\,t}{2} + \frac{1}{2}\,(-1 + t)\,t$$

$$1 - 3\,t + \frac{5}{2}\,(-1 + t)\,t - \frac{5}{6}\,(-2 + t)\,(-1 + t)\,t$$

$$1 - 5\,t + \frac{15}{2}\,(-1 + t)\,t - \frac{35}{6}\,(-2 + t)\,(-1 + t)\,t + \frac{35}{12}\,(-3 + t)\,(-2 + t)\,(-1 + t)\,t$$

Table 1.2: The first 5 orthogonal polynomials $\{p_{k,N}(t)\}_{k=0,\dots,N}$ for $N = 4$.

1.5 Chebyshev knots and Chebyshev orthogonal polynomials

As in polynomial collocation Chebyshev knots and Chebyshev polynomials as basis functions provide an efficient and elegant way to solve least-squares approximation with very good error properties.

The Chebyshev *T*-polynomials are defined by a simple trigonometric formula on the interval [-1, 1]:

$$T_n(x) = \cos(n \arccos x) \quad (n = 0,1,\ldots) \tag{1.10}$$

The table and figure below show the first few Chebyshev polynomials:

1
x
$2x^2 - 1$
$4x^3 - 3x$
$8x^4 - 8x^2 + 1$
$16x^5 - 20x^3 + 5x$
$32x^6 - 48x^4 + 18x^2 - 1$
$64x^7 - 112x^5 + 56x^3 - 7x$
$128x^8 - 256x^6 + 160x^4 - 32x^2 + 1$
$256x^9 - 576x^7 + 432x^5 - 120x^3 + 9x$
$512x^{10} - 1280x^8 + 1120x^6 - 400x^4 + 50x^2 - 1$

Table 1.3: The first 11 Chebyshev polynomials. As polynomials they are defined in the whole domain of complex numbers, of course.

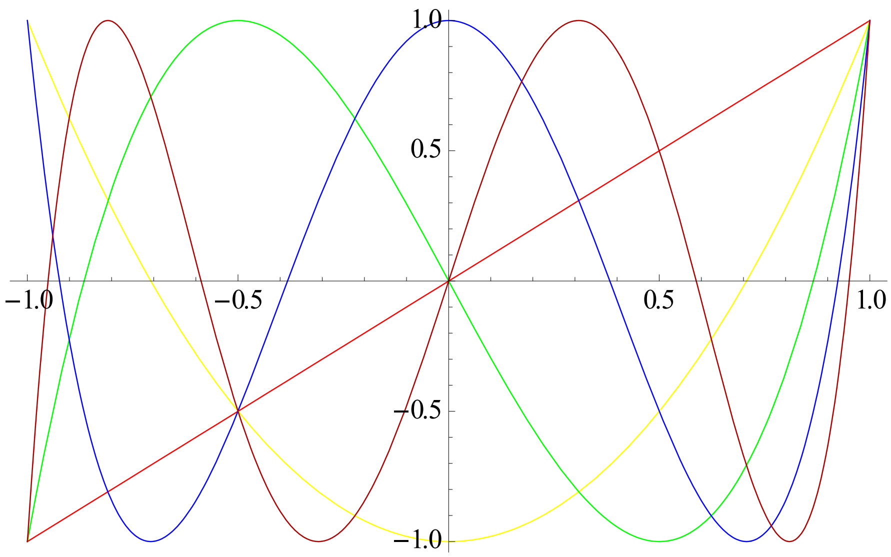

Figure 1.4: The first 6 Chebyshev polynomials except T_0 plotted in the domain [-1, 1]. But as polynomials they are defined in the whole domain of complex numbers.

Property 1.1: Elementary properties of the Chebyshev polynomials

a) $\max_{-1\le x\le 1} T_n(x) = 1 \quad \min_{-1\le x\le 1} T_n(x) = -1$

b) $\boxed{T_n(x) = 0 \Leftrightarrow x = \cos\left(\frac{2i+1}{2n}\pi\right) \quad (i = 0,1,\ldots,n-1)}$

c) $T_n(x) = \pm 1 \Leftrightarrow x = \cos\left(\frac{i}{n}\pi\right) \quad (i = 0,1,\ldots,n)$

(Proof: These are direct consequences of the trigonometric definition (1.10) □)

Property 1.2: Recursive relation

The Chebyshev polynomials fulfill the 2nd order recursion $\boxed{T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \quad (n \ge 2)}$ with initial conditions $T_1(x) = x,\ T_0(x) = 1$.

(Proof: By elementary trigonometry (addition theorem): $\cos((n+1)a) + \cos((n-1)a) = 2\cos a\ \cos(na)$. Substituting $a = \arccos x$ yields the recursion relation in the proposition □)

One of the most important property of Chebyhsev polynomials is that they are uniformly small on [-1, 1] and that they are optimal in this behaviour in the following sense.

Theorem 1.5: Min-Max property on [-1,1]

If $p(x)$ is a $(n+1)$-degree polynomial with leading coefficient 1 such that the maximal absolute value $\max_{-1\le x\le 1}|p(x)|$ is minimal then $\max_{-1\le x\le 1}|p(x)|$ is necessarily equal to $\frac{1}{2^n}$ as is the case for the normalized Chebyshev polynomial $\frac{1}{2^n}T_{n+1}(x)$.

(Proof: On one side we have that $\max_{-1\le x\le 1}\left|\frac{1}{2^n}T_{n+1}(x)\right| = \frac{1}{2^n}$ due to Property 1.1a and $\frac{1}{2^n}T_{n+1}(x)$ takes its extremal values $\pm\frac{1}{2^n}$ at $x_i = \cos\left(\frac{i}{(n+1)}\pi\right) \quad (i = 0,1,\ldots,n,n+1)$. On the other side suppose that some polynomial $p(x) = x^{n+1} + a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots a_1 x^1 + a_0$ were such that $\max_{-1\le x\le 1}|p(x)| < \frac{1}{2^n}$. Then define the polynomial $q(x) = p(x) - \frac{1}{2^n}T_{n+1}(x)$. This polynomial is of degree n or less and it does not vanish identically since this would require $\max_{-1\le x\le 1}|p(x)| = \frac{1}{2^n}$. Since $p(x)$ is dominated by $\frac{1}{2^n}T_{n+1}(x)$ at these points in absolute value, it follows that the values

$q(x_i)$ alternate in signs. Being continuous, $q(x)$ must have $n+1$ intermediate zeros. This contradicts the fact that q has degree at most n and is not identically zero. Thus $\max\limits_{-1\le x\le 1}|p(x)|\ge \frac{1}{2^n}$

□)

A polynomial as $\frac{1}{2^n}T_{n+1}(x)$ in Theorem 1.5 is called a <u>min-max polynomial</u>. By Property 1.1b) the zeros of $\frac{1}{2^n}T_{n+1}(x)$ are $x_i=\cos\left(\frac{2i+1}{2(n+1)}\pi\right)$ $(i=0,1,\dots,n)$. So we can write $\frac{1}{2^n}T_{n+1}(x)$ as $(x-x_0)(x-x_1)\cdots(x-x_n)=\pi_{n+1}(x)$ (<u>Newton polynomial</u> form) and we know now that $\max\limits_{-1\le x\le 1}|\pi_{n+1}(x)|$ is $\frac{1}{2^n}$ and this is minimal.

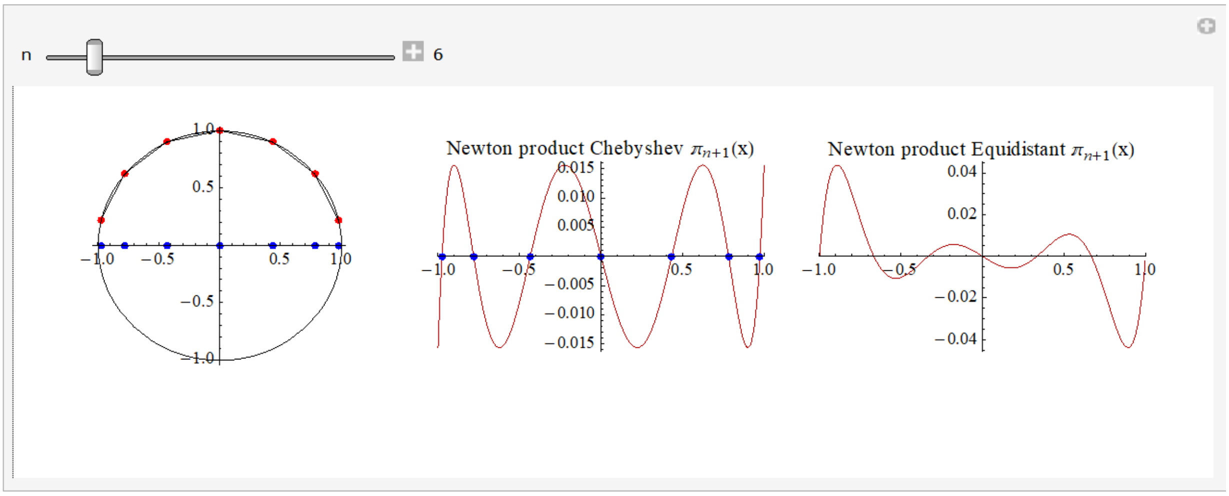

Figure 1.5: The situation of Theorem 1.5 for n = 6. The figure on the left shows the construction of the Chebyshev knots (blue points on the abscissa). The figure in the middle is a plot of $\frac{1}{2^n}T_{n+1}(x)$ and the figure on the right is a plot of $\pi_{n+1}(x)$ for equi-distant arguments.

Additionally, to Theorem 1.5 we have an <u>orthogonality relation for the Chebyshev polynomials</u>. The next theorem is an analogue to Theorem 1.4 for the Chebyshev arguments (knots) $x_i=\cos\left(\frac{2i+1}{2(n+1)}\pi\right)$ $(i=0,1,\dots,n)$

Theorem 1.6: Discrete orthogonality

The Chebyhsev polynomials are orthogonal in the following sense:

$$\langle T_j,T_k\rangle := \sum_{i=0}^{N} T_j(x_i)T_k(x_i)=\begin{cases}0 & j\ne k\\ (N+1)/2 & j=k\ne 0\\ N+1 & j=k=0\end{cases}\qquad (j,k=0,\dots,N)$$
$$x_i=\cos\left(\frac{2i+1}{2(N+1)}\pi\right)\qquad (i=0,1,\dots,N) \tag{1.11}$$

(Proof: Define the angles $a_i = \left(\frac{2i+1}{2(N+1)}\pi\right) \quad (i=0,1,...,N)$ and thus $x_i = \cos(a_i) \quad (i=0,1,...,N)$. By definition (1.10) we have $\langle T_j, T_k \rangle := \sum_{i=0}^{N} T_j(x_i) T_k(x_i) =$ $\sum_{i=0}^{N} \cos(ja_i)\cos(ka_i) = \frac{1}{2}\sum_{i=0}^{N}\cos((j+k)a_i) + \frac{1}{2}\sum_{i=0}^{N}\cos((j-k)a_i)$ by the addition theorem for the cosine function. Because the angles $a_i \quad (i=0,1,...,N)$ "lie" on a regular, symmetric polygon on the upper half circle (cf. Figure 1.5) each of these sums is zero except for the cases that $j+k=0$ or $j-k=0$, respectively. This proves the theorem for the case $j \neq k$. If $j=k=0$ the sums evaluate to $\frac{1}{2}(N+1)+\frac{1}{2}(N+1)=N+1$. If $j=k\neq 0$ then the sums evaluate to $\frac{1}{2}0+\frac{1}{2}(N+1)=\frac{1}{2}(N+1)$

□)

Example 1.5: A measurement of flow rates (volume per second) consists of N + 1= 231+1 data (there are 232 measurements). The data have to be approximated by polynomials of degree at most m = 4. Figures 1.6ab compare the cases of almost uniform arguments versus Chebyshev arguments.

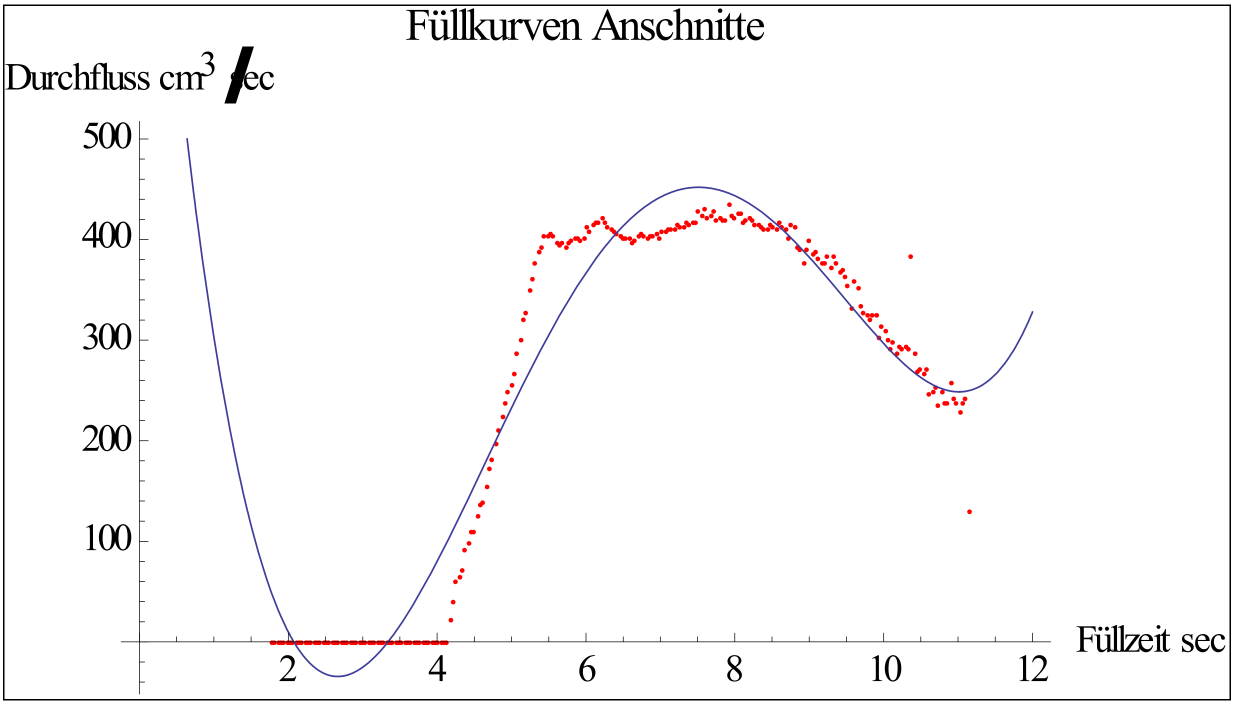

Figure 1.6a: A measurement at 232 almost uniform arguments and its least-squares approximation with respect to the basis $\{1, x, x^2, x^3, x^4\}$. The formula for the approximation curve is

$$999.371 - 951.948x + 285.299x^2 - 30.5158x^3 + 1.08026x^4$$

The square sum of the residuals is equal to `Smin = 338456.` The normal equations (1.3') have non-diagonal matrix $G^{tr}G$ and right hand side vector:

$$\begin{pmatrix} 232. & 1500.81 & 11419.6 & 95979.7 & 857822. \\ 1500.81 & 11419.6 & 95979.7 & 857822. & 7.98116\times10^{6} \\ 11419.6 & 95979.7 & 857822. & 7.98116\times10^{6} & 7.63635\times10^{7} \\ 95979.7 & 857822. & 7.98116\times10^{6} & 7.63635\times10^{7} & 7.45801\times10^{8} \\ 857822. & 7.98116\times10^{6} & 7.63635\times10^{7} & 7.45801\times10^{8} & 7.39911\times10^{9} \end{pmatrix}$$

$$\begin{pmatrix} 58921.2 \\ 456501. \\ 3.72083\times10^{6} \\ 3.16968\times10^{7} \\ 2.80255\times10^{8} \end{pmatrix}$$

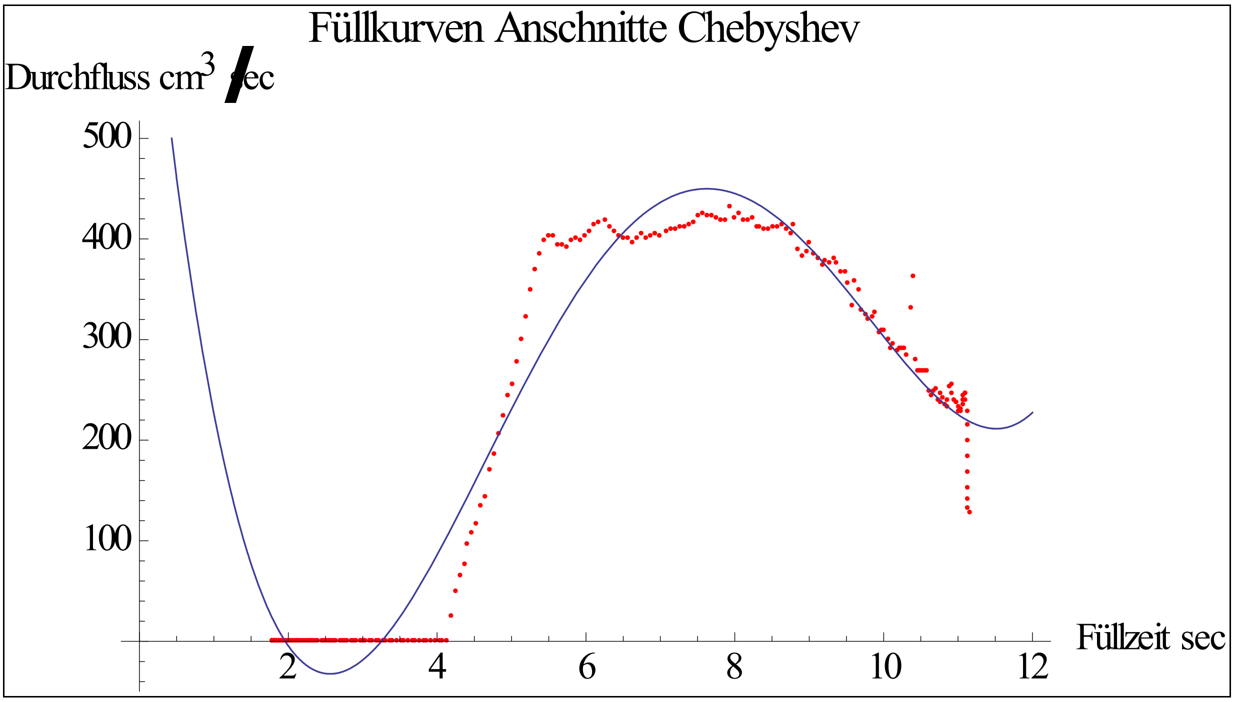

Figure 1.6b: A measurement at 232 Chebyshev arguments in the interval [a,b] = [1.7818, 11.14] and its least-squares approximation with respect to the basis

$$\{1, T_1(-1+2\frac{x-a}{b-a}), T_2(-1+2\frac{x-a}{b-a}), T_3(-1+2\frac{x-a}{b-a}), T_4(-1+2\frac{x-a}{b-a})\}.$$

The Chebyshev arguments are computed by:

$$x_i = a + \frac{b-a}{2}\left(\cos\left(\frac{2i+1}{2(N+1)}\pi\right)+1\right) \quad (i = 0,1,\ldots,N)$$

Here the affine transformation $x \mapsto a + \frac{b-a}{2}(x+1)$ mapping the interval [-1, 1] on [a, b] and its inverse $-1+2\frac{x-a}{b-a}$ mapping [a, b] to [-1, 1] have been used.

The formula for the approximation curve is

```
210.204 +
166.457 (-1.+0.213716 (-1.7818+x)) -
140.852 (-1.+2. (-1.+0.213716 (-1.7818+x))²) -
68.6974 (-3. (-1.+0.213716 (-1.7818+x))+4. (-1.+0.213716 (-1.7818+x))³) +
52.0094 (1.-8. (-1.+0.213716 (-1.7818+x))²+8. (-1.+0.213716 (-1.7818+x))⁴)

= 794.179 -780.861 x + 237.52 x² - 25.1147 x³ + 0.868006 x⁴
```

The square sum of the residuals is equal to `Smin = 291180`. The normal equations (1.3') have a diagonal matrix $G^{tr}G$ and right hand side vector:

$$\begin{pmatrix} 232. & -1.9984\times10^{-14} & -3.01981\times10^{-14} & -2.84217\times10^{-14} & -1.77636\times10^{-14} \\ -1.9984\times10^{-14} & 116. & -1.43219\times10^{-14} & -3.09752\times10^{-14} & -4.06342\times10^{-14} \\ -3.01981\times10^{-14} & -1.43219\times10^{-14} & 116. & -3.78586\times10^{-14} & -2.36478\times10^{-14} \\ -2.84217\times10^{-14} & -3.09752\times10^{-14} & -3.78586\times10^{-14} & 116. & -5.56222\times10^{-14} \\ -1.77636\times10^{-14} & -4.06342\times10^{-14} & -2.36478\times10^{-14} & -5.56222\times10^{-14} & 116. \end{pmatrix}$$

$$\begin{pmatrix} 48767.3 \\ 19309. \\ -16338.8 \\ -7968.9 \\ 6033.09 \end{pmatrix}$$

The next theorem reflects the idea what happens if $N \to \infty$ in Theorem 1.6.

Theorem 1.7: Continuous orthogonality

The Chebyshev polynomials fulfill the following continuous orthogonal relations:

$$\langle T_j, T_k \rangle_{cont} := \int_{-1}^{1} T_j(x) T_k(x) \frac{dx}{\sqrt{1-x^2}} = \begin{cases} 0 & j \neq k \\ \pi/2 & j = k \neq 0 \\ \pi & j = k = 0 \end{cases} \tag{1.12}$$

(Proof: By substitution $x = \cos u \Rightarrow dx = -\sin u \, du$ we get that

$$\int_{-1}^{1} T_j(x) T_k(x) \frac{dx}{\sqrt{1-x^2}} = -\int_{\pi}^{0} T_j(\cos u) T_k(\cos u) \frac{\sin u \, du}{\sqrt{1-\cos^2 u}} = \int_0^{\pi} \cos(ju)\cos(ku) du =$$

$$\frac{1}{2}\int_0^{\pi} \left(\cos((j+k)u) + \cos((j-k)u)\right) du = \frac{1}{2}\left(\frac{\sin((j+k)u)}{(j+k)} + \frac{\sin((j-k)u)}{(j-k)}\right)\Bigg|_0^{\pi} =$$

$$\begin{cases} 0 & j \neq k \\ \pi/2 & j = k \neq 0 \\ \pi & j = k = 0 \end{cases}$$ □)

From Theorem 1.7 we can derive a continuous version of Theorem 1.1. The idea here is to replace the discrete set of measurement $\{y_i\}_{i=0,...,N}$ with a function depending on a continuous argument $x \in (-1,1)$. This concept is called continuous least-squares approximation:

Theorem 1.8: Continuous Chebyshev least-squares approximation

If $y(x)$ is function on (-1,1) which is absolutely square-integrable with respect to the weight function $w(x) := \frac{1}{\sqrt{1-x^2}} \quad (x \in (-1,1))$ in the sense that $\int_{-1}^{1} |y(x)|^2 \frac{dx}{\sqrt{1-x^2}} < \infty$, then the continuous square-sum of residuals

$$S := \int_{-1}^{1} \left(y(x) - \sum_{j=0}^{m} a_j T_j(x) \right)^2 \frac{dx}{\sqrt{1-x^2}} \tag{1.13}$$

is minimal if and only if

$$a_j = \begin{cases} \frac{2}{\pi} \int_{-1}^{1} y(x) T_j(x) \frac{dx}{\sqrt{1-x^2}} & j > 0 \\ \frac{1}{\pi} \int_{-1}^{1} y(x) \frac{dx}{\sqrt{1-x^2}} & j = 0 \end{cases} \tag{1.14}$$

The polynomial $\sum_{j=0}^{m} a_j T_j(x)$ is called the continuous least-squares Chebyshev approximation of degree m.

The minimal square-sum of residuals then is given by $S_{\min} = \int_{-1}^{1} \frac{y(x)^2 dx}{\sqrt{1-x^2}} - \sum_{j=0}^{m} a_j^2 \int_{-1}^{1} \frac{T_j(x)^2 dx}{\sqrt{1-x^2}}$.

(Proof: The two proofs of Theorem 1.1 and the formulas (1.8) and (1.8') can be carried over to the present situation by the idea that $\langle T_j, T_k \rangle_{cont} := \int_{-1}^{1} T_j(x) T_k(x) \frac{dx}{\sqrt{1-x^2}}$ replaces $< g_j, g_k >$.

(1) As in the proof of Theorem 1.7 $\langle T_j, T_k \rangle_{cont} := \int_{-1}^{1} T_j(x) T_k(x) \frac{dx}{\sqrt{1-x^2}} = \int_{0}^{\pi} \cos(ju)\cos(ku)du$.

Setting the equi-distant Chebyshev angles $u_i = \frac{(2i+1)}{2(N+1)}\pi \quad (i = 0,1,...,N)$ we get $\Delta u_i = \frac{\pi}{N+1} \quad (i = 0,...,N-1)$ and by definition of the integral $\int_{0}^{\pi} \cos(ju)\cos(ku)du =$

$\lim_{N\to\infty}\frac{\pi}{N+1}\sum_{i=0}^{N}\cos(ju_i)\cos(ku_i) = \lim_{N\to\infty}\frac{\pi}{N+1}\sum_{i=0}^{N}\cos(ju_i)\cos(ku_i) = \lim_{N\to\infty}\frac{\pi}{N+1}\sum_{i=0}^{N}T_j(x_i)T_k(x_i) =$ $\pi\lim_{N\to\infty}\frac{1}{N+1}<T_j,T_k>$ for the Chebyshev arguments $x_i=\cos(u_i)=\cos\left(\frac{(2i+1)}{2(N+1)}\pi\right)\quad(i=0,1,...,N)$. In the same way we get that $\left\langle y,T_j\right\rangle_{cont}:=\int_{-1}^{1}y(x)T_j(x)\frac{dx}{\sqrt{1-x^2}} = \lim_{N\to\infty}\frac{\pi}{N+1}\sum_{i=0}^{N}y(x_i)T_j(x_i) = \pi\lim_{N\to\infty}\frac{1}{N+1}<y,T_j>$. The formulas (1.14) now follow from (1.11) in Theorem 1.6 and (1.8) if $N\to\infty$.

(2) The geometric proof (cf. Figure 1.3) with $<g_j,g_k>$ replaced by $\left\langle T_j,T_k\right\rangle_{cont}$ establishes the theorem by the continuous orthogonal relations (1.12)

□)

Corollary 1.1: Polynomial continuous least-squares Chebyshev approximation

If $y(x)$ is a polynomial of degree *n* then setting $N = n$ implies that discrete Chebyshev least-squares approximation with $x_i=\cos\left(\frac{(2i+1)}{2(N+1)}\pi\right)\quad(i=0,1,...,N)$ is the same as continuous Chebyshev least-squares approximation.

(Proof: Since $y(x)$ is a polynomial of degree n it can be written as a linear combination of Chebyshev polynomials $\sum_{k=0}^{n}a_kT_k(x)\quad(k=0,1,...,n=N)$.

Now by Theorem 1.6 for the Chebyshev knots $x_i=\cos\left(\frac{(2i+1)}{2(N+1)}\pi\right)\quad(i=0,1,...,N)$

$<y|,T_j|> = \left\langle\sum_{k=0}^{n}a_k\,T_k|,T_j|\right\rangle = \sum_{k=0}^{n}a_k<T_k|,T_j|> = a_j<T_j|,T_j|>$ which implies $a_j=\frac{<y|,T_j|>}{<T_j|,T_j|>}$.

A similar computation for $<y(x),T_j(x)>_{cont}$ applying Theorem 1.7 implies that

$a_j=\frac{<y(x),T_j(x)>_{cont}}{<T_j,T_j>_{cont}}$. This shows that discrete Chebyshev approximation is the same as continuous because the coefficients are the same

□)

Corollary 1.1 is useful because integrals as in (1.14) can be replaced by discrete sums. The next example shows how it is used together with truncation of a Chebyshev expansion.

Example 1.6: Chebyshev continuous least-squares parabola on the interval (0,1) for $y(t) = t^3$

First the interval (0,1) is transformed to (-1,1) by $x = 2t - 1$, then by Table 1.3

$$y(x) = \frac{(x+1)^3}{8} = \frac{1}{8}(x^3 + 3x^2 + 3x + 1) =$$

$$\frac{1}{8}\left(\frac{T_3(x) + 3T_1(x)}{4} + 3\frac{T_2(x) + T_0(x)}{2} + 3T_1(x) + T_0(x)\right) =$$

$$\frac{1}{8}\left(\frac{T_3(x) + 3T_1(x)}{4} + 3\frac{T_2(x) + T_0(x)}{2} + 3T_1(x) + T_0(x)\right) = \frac{1}{32}T_3(x) + \frac{3}{16}T_2(x) + \frac{15}{32}T_1(x) + \frac{5}{16}T_0(x).$$

By <u>truncation</u> we get that the Chebyshev least-squares parabola is equal to

$$\frac{3}{16}T_2(x) + \frac{15}{32}T_1(x) + \frac{5}{16}T_0(x) = \frac{3}{16}T_2(2t-1) + \frac{15}{32}T_1(2t-1) + \frac{5}{16}1.$$

The problem of Example 1.6 can also be solved by a discrete least-squares approximation of the Chebyshev data $(x_i, y_i = y(x_i))$ with $x_i = \cos\left(\frac{(2i+1)}{2(N+1)}\pi\right)$ $(i = 0,1,\ldots,N=3)$, i.e. $\left\{\cos(\frac{\pi}{8}), \cos(\frac{3\pi}{8}), \cos(\frac{5\pi}{8}), \cos(\frac{7\pi}{8})\right\}$, by a parabola $T_0(x) + a_1T_1(x) + a_2T_2(x)$. But the truncation method, of course, is much faster and more elegant.

The reason why the truncation method works for polynomial modell functions $y(x)$ $(-1 \le x \le 1)$ stems from the orthogonality of the Chebyshev-Polynomials (Theorem 1.7, formula (1.12)) and the main Theorem 1.8 (formula 1.14).

This will be demonstrated with the example $\frac{1}{32}T_3(x) + \frac{3}{16}T_2(x) + \frac{15}{32}T_1(x) + \frac{5}{16}T_0(x)$ = $b_3T_3(x) + b_2T_2(x) + b_1T_1(x) + b_0T_0(x)$ by proving that the coefficients b_j are optimal.

By Theorem 1.8 for $0 \le j \le 3$ the optimal coefficients are:

$$a_j = \begin{cases} \frac{2}{\pi}\int_{-1}^{1} y(x)T_j(x)\frac{dx}{\sqrt{1-x^2}} = \frac{2}{\pi}\int_{-1}^{1}\left(b_3T_3(x) + b_2T_2(x) + b_1T_1(x) + b_0T_0(x)\right)T_j(x)\frac{dx}{\sqrt{1-x^2}} & j > 0 \\ \frac{1}{\pi}\int_{-1}^{1}\left(b_3T_3(x) + b_2T_2(x) + b_1T_1(x) + b_0T_0(x)\right)\frac{dx}{\sqrt{1-x^2}} = 0 + \cdots + 0 + b_0\frac{1}{\pi}\underbrace{\int_{-1}^{1} T_0(x)T_0(x)\frac{dx}{\sqrt{1-x^2}}}_{\pi} & j = 0 \end{cases}$$

$$\ldots = \begin{cases} 0 + \cdots + 0 + b_j\frac{2}{\pi}\underbrace{\int_{-1}^{1} T_j(x)T_j(x)\frac{dx}{\sqrt{1-x^2}}}_{\pi/2} + 0 + \cdots + 0 = b_j & j > 0 \\ b_0 & j = 0 \end{cases}$$

∎

1.6 Multi-variate linear least-squares approximation

Multi-variate linear least-squares problems are solved exactly along the same lines as uni-variate problems. The methods developped in Sections 1.1 to 1.3 remain valid and keep the same form.

The x-arguments now are vectors $\vec{x}_0, \vec{x}_1, \ldots, \vec{x}_{n-1}, \vec{x}_N$ of dimension d and thus we need basis functions which take vectors as inputs (i.e. functions in several variables). For practical purposes most widely used are tensor products of uni-variate basis functions.

If $g_0, g_1, ..., g_m = \{g_j\}_{j=0,...,m}$ is a uni-variate basis then the *d*-fold tensor product

$$\boxed{\sum_{j_1=0}^{m_1}\sum_{j_2=0}^{m_2}\cdots\sum_{j_d=0}^{m_d} a_{j_1,j_2,...,j_d}\, g_{j_1}(x^{(1)}) g_{j_2}(x^{(2)})\cdots g_{j_d}(x^{(d)})} \qquad \textbf{(1.15)}$$

defines a basis of $m_1 \times m_2 \times \cdots \times m_d$ functions in d variables $x^{(1)}, x^{(2)}, \ldots, x^{(d-1)}, x^{(d)}$.

Type of basis functions	**2d-formulas**
Trigonometric	$\{e^{ikx^{(1)}} \cdot e^{i\ell x^{(2)}} \mid k, \ell \in Z\} \quad (x^{(1)}, x^{(2)} \in (0, 2\pi)$
Polynomials	$\{x^{(1)j} x^{(2)k} \mid j, k \in N_0\}$ (standard monomials) $\left\{\left(\frac{x^{(1)}-\mu_1}{\sigma_1}\right)^j \left(\frac{x^{(2)}-\mu_2}{\sigma_2}\right)^k \mid j, k \in N_0\right\}$ (normalized standard monomials: μ and σ denote mean and standard deviation of the x-data. $\{T_j(x^{(1)}) T_k(x^{(2)}) \mid j, k \in N_0\}$ (Chebyshev polynomials, $x^{(1)}, x^{(2)} \in (-1, 1)$) From the uni-variate orthogonal polynomials $p_{k,N}(t) = \sum_{i=0}^{k} (-1)^i \binom{k}{i}\binom{k+i}{i} \frac{t^{(i)}}{N^{(i)}} \quad (k = 0,1,...,N)$ we get an orthogonal 2d-basis $\{p_{k_1,N_1}(t_1) p_{k_2,N_2}(t_2) \mid (k_1, k_2) \in \{0,1,...,N_1\} \times \{0,1,...,N_2\}\}$ for a uniform 2-dimensional integer grid $(t_1, t_2) \in \{0,1,...,N_1\} \times \{0,1,...,N_2\}$. Note that superscripts in the variables here are not used in order to avoid confusion with the notation of the falling factorials.

Table 1.4: A few basis functions in two variables. The generalizations to more than two variables are straightforward. Cf. also Table 1.1.

Example 1.7: Multi-variate least-squares for 4695 three-dimensional data points

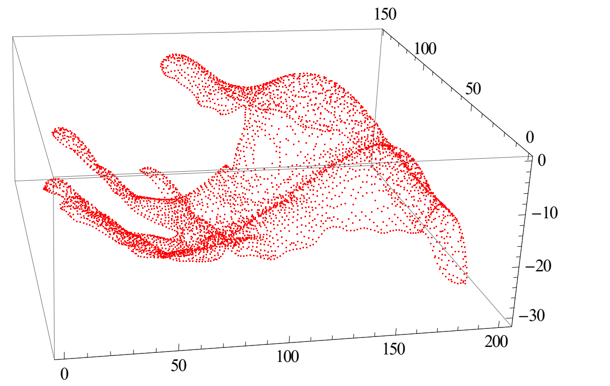

Figure 1.7a: 4695 data 3d-data points

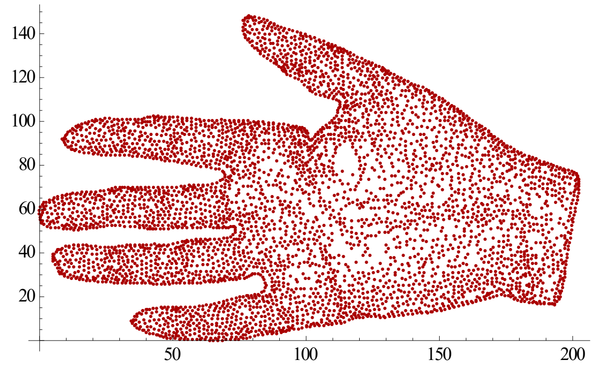

Figure 1.7b: Irregular grid in the argument plane (projections)

A. Degree 3 complete basis

Basis functions: $\{1, x, y, x^2, 2xy, y^2, x^3, 3x^2y, 3xy^2, y^3\}$

The first 4 data points are:

```
{ {82.0565,99.8271,-25.3321},{83.9503,99.6029,-24.8481},
  {85.8279,99.3811,-24.3083},{80.1541,99.9811,-25.753}, … }
```

The design matrix has dimensions 4695 x 10 and its first 3 rows are:

1.	82.0565	99.8271	6733.27	16382.9	9965.45	552509.	2.01649×10^6	2.45319×10^6	994822.
1.	83.9503	99.6029	7047.65	16723.4	9920.74	591653.	2.1059×10^6	2.49855×10^6	988134.
1.	85.8279	99.3811	7366.43	17059.3	9876.6	632245.	2.19625×10^6	2.54306×10^6	981548.

The first 16 rows of the diagonal matrix of the singular value decomposition are:

3.13397×10^8	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	1.17066×10^8	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	3.5396×10^7	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	7.76625×10^6	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	223954.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	86901.1	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	34224.4	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	629.702	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	282.232	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	4.50267
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.

Least-squares approximation: $35.2489 - 1.35126\, x + 0.0131384\, x^2 - 0.0000416223\, x^3 - 1.23016\, y + 0.0100748\, x\, y - 4.38128\times10^{-6}\, x^2\, y + 0.00911392\, y^2 - 0.0000704719\, x\, y^2 + 8.30584\times10^{-7}\, y^3$

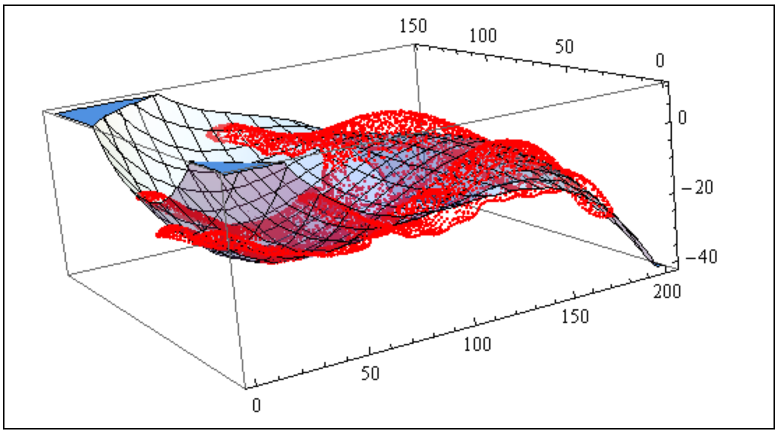

Figure 1.8a: 4th degree least-squares approximation and data points

B. Degree 4 complete basis

Basis functions: $\{1, x, y, x^2, 2xy, y^2, x^3, 3x^2y, 3xy^2, y^3, x^4, 4x^3y, 6x^2y^2, 4xy^3, y^4\}$

Least-squares approximation: $-2.75677+0.130319\ x-0.00670348\ x^2+0.0000625409\ x^3-1.9681\times10^{-7}\ x^4+0.000438675\ y-0.0145805\ x\ y+0.000200319\ x^2\ y-4.24924\times10^{-7}\ x^3\ y-0.00900462\ y^2+0.0000435083\ x\ y^2-6.40939\times10^{-7}\ x^2\ y^2+0.00015033\ y^3-1.15603\times10^{-8}\ x\ y^3-5.23578\times10^{-7}\ y^4$

The design matrix has dimensions 4695 x 15 and its first 4 rows are:

1.	82.0565	99.8271	6733.27	16382.9	9965.45	552509.	2.01649×10^6	2.45319×10^6	994822.	4.53369×10^7	2.20621×10^8	4.026×10^8	3.26526×10^8	9.93102×10^7
1.	83.9503	99.6029	7047.65	16723.4	9920.74	591653.	2.1059×10^6	2.49855×10^6	988134.	4.96694×10^7	2.35721×10^8	4.19507×10^8	3.31817×10^8	9.8421×10^7
1.	85.8279	99.3811	7366.43	17059.3	9876.6	632245.	2.19625×10^6	2.54306×10^6	981548.	5.42643×10^7	2.51333×10^8	4.36532×10^8	3.36977×10^8	9.75473×10^7

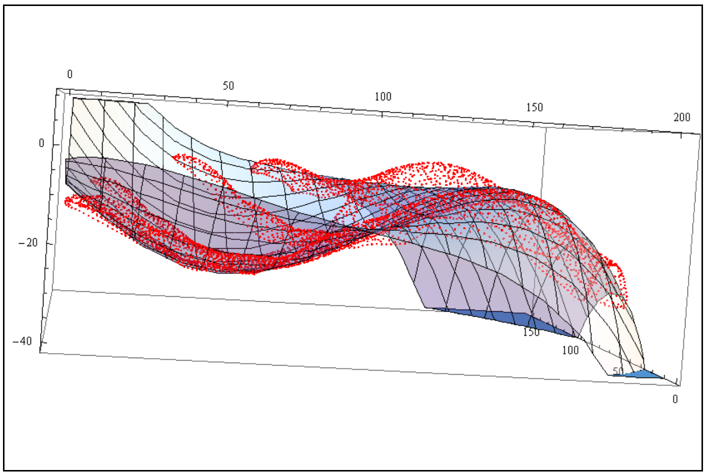

Figure 1.8b: 4th degree least-squares approximation and data points

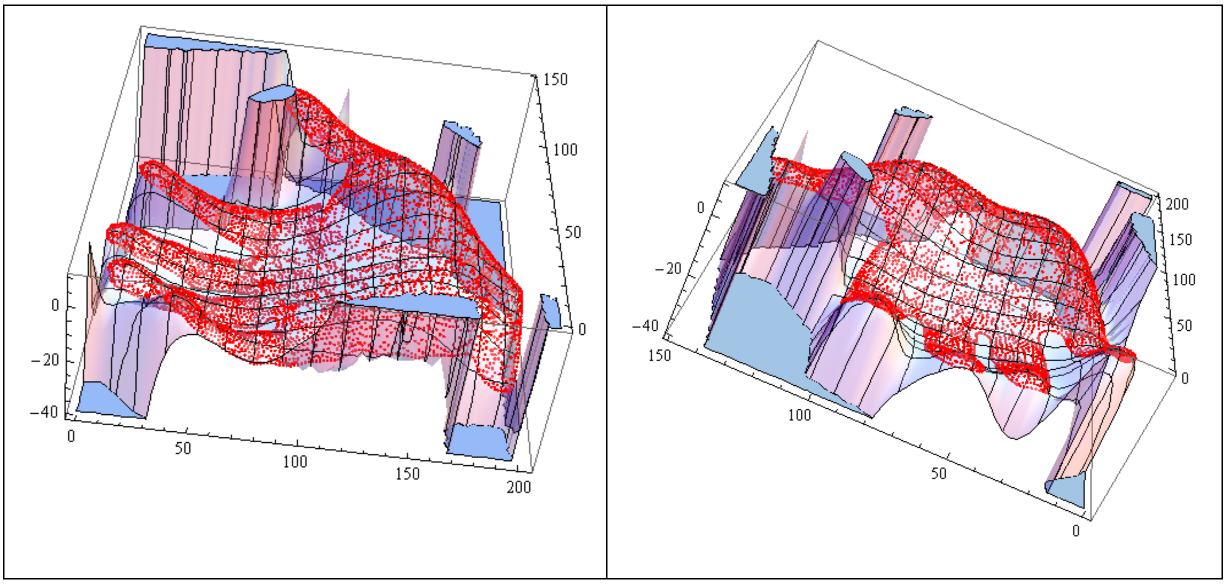

Figure 1.8cd: 64th degree least-squares approximation and data points in two different views. The design matrix has dimensions 4695 x 2145. The basis elements are generated by expanding the powers $(x+y)^k \quad k = 0,1,\ldots,64$. They consist of $k(k+1)/2$ = 65*33 different elements.

1.7 Legendre Polynome

In section 1.5 above Chebyshev polynomials as orthogonal basis functions provided an efficient and elegant way to solve the weighted continuous least-squares approximation problem with weighting function $w(x) = \frac{1}{\sqrt{1-x^2}} \quad (-1 < x < 1)$.

A similar mechanism can be established by the Legendre polynomials for the simple weighting function $w(x) = 1 \quad (-1 \le x \le 1)$.

The Legendre P-polynomials are defined here by Rodriguez' formula[1]:

$$P_n(x) = \frac{1}{2^n n!} \cdot \frac{d^n}{dx^n}\left(x^2 - 1\right)^n \qquad \textbf{(1.16)}$$

The table and figure below show the first few Legendre polynomials:

$P_0(x) = 1$	$P_1(x) = x$	$P_2(x) = \frac{1}{2}\left(3x^2 - 1\right)$	$P_3(x) = \frac{1}{2}\left(5x^3 - 3x\right)$
$P_4(x) = \frac{1}{8}\left(35x^4 - 30x^2 + 3\right)$		$P_5(x) = \frac{1}{8}\left(63x^5 - 70x^3 + 15x\right)$	
$P_6(x) = \frac{1}{16}\left(231x^6 - 315x^4 + 105x^2 - 5\right)$		$P_7(x) = \frac{1}{16}\left(429x^7 - 693x^5 + 315x^3 - 35x\right)$	

Table 1.5: The first Legendre polynomials. As polynomials they are defined in the whole domain of complex numbers, of course.

Property 1.3: Properties of the Legendre polynomials

a) $\max\limits_{-1\le x\le 1}|P_n(x)| = 1 \quad (n \in N_0) \qquad \min\limits_{-1\le x\le 1} P_n(x) = -1 \quad (n = 1,3,5...)$

b) $P_n(-x) = (-1)^n P_n(x)$ (symmetry or anti-symmetry)

c) $P_n(1) = 1,\ P_n(-1) = (-1)^n$

---

[1] Another definition uses a generating function approach:

$$\frac{1}{\sqrt{1-2xt+t^2}} = \sum_{n=0}^{\infty} P_n(x)t^n \quad (-1 \le x \le 1, -1 < t < 1)$$

Differentiating with respect to t yields $(x-t)\cdot\frac{1}{\sqrt{1-2xt+t^2}} = \left(1-2xt+t^2\right)\cdot\sum_{n=1}^{\infty} nP_n(x)t^{n-1}$. Substituting again $\sum_{n=0}^{\infty} P_n(x)t^n$ for $\frac{1}{\sqrt{1-2xt+t^2}}$ on the l.h.s. and comparing coefficients of the powers of t yields the Bonnet recursion formula: $(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x)$.

(Proof: Parts ab) are consequences of the definition (1.16) but a) is not easily derived. Part c) follows by induction from the Bonnet recursion in the footnote to (1.16).
A (rather complicated) proof deriving the Bonnet recursion from the Rodriguez formula can be found in Schaum's outline of Numerical Analysis, 2nd edition, Problem 15.17. □)

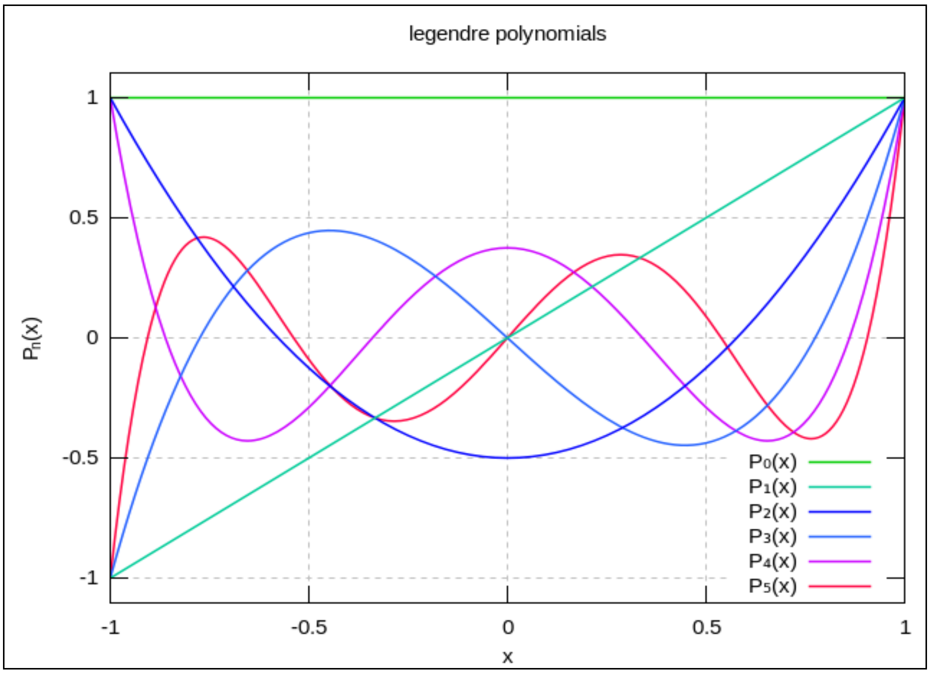

Figure 1.9: The first 6 Legendre polynomials plotted in the domain [-1, 1]. As polynomials they are defined in the whole domain of complex numbers.

For continuous least-squares approximation the next orthogonality property is fundamental.

Theorem 1.9: Orthogonality of the Legendre polynomials

The Legendre polynomials are orthogonal with respect to the integral inner product with weight $w(x)=1 \quad (-1 \le x \le 1)$:

$$\int_{-1}^{1} P_n(x)P_m(x)dx = \begin{cases} 0 & m \neq n \\ \dfrac{2}{2n+1} & m = n \end{cases} \qquad \textbf{(1.17)}$$

(Proof: A proof can be found in Schaum's outline of Numerical Analysis, 2nd ed., Problems 15.9 – 15.12. □)

From the orthogonality Theorem 1.9 we can derive a continuous version of Theorem 1.1.

Theorem 1.10: Continuous Legendre least-squares approximation

If $y(x)$ is function on [-1,1] which is absolutely square-integrable with respect to the weight function $w(x)=1 \quad (x \in [-1,1])$ in the sense that $\int_{-1}^{1} |y(x)|^2\, dx < \infty$, then the continuous square-sum of residuals

$$S := \int_{-1}^{1} \left(y(x) - \sum_{j=0}^{m} a_j P_j(x) \right)^2 dx \tag{1.18}$$

is minimal if and only if

$$a_j = \frac{2j+1}{2} \int_{-1}^{1} y(x) P_j(x)\, dx \quad (j = 0, \ldots, m) \tag{1.19}$$

The polynomial $\sum_{j=0}^{m} a_j P_j(x)$ is called the continuous least-squares Legendre approximation of degree m.

The minimal square-sum of residuals then is given by $S_{\min} = \int_{-1}^{1} y(x)^2 dx - \sum_{j=0}^{m} a_j^2 \int_{-1}^{1} P_j(x)^2 dx$.

(Proof: The two proofs of Theorem 1.1 and the formulas (1.8) and (1.8') can be carried over to the present situation by the idea that $\langle P_j, P_k \rangle_{cont} := \int_{-1}^{1} P_j(x) P_k(x)\, dx$ replaces $< g_j, g_k >$.

The geometric proof (cf. Figure 1.3) with $< g_j, g_k >$ replaced by $\langle P_j, P_k \rangle_{cont}$ establishes the theorem by the continuous orthogonal relations (1.17)
□)

The next example shows that for polynomials the coefficients a_j in (1.19) can be computed without solving integrals. It is an application of the truncation method.

Example 1.8: Legendre continuous least-squares parabola on the interval (0,1) for $y(t) = t^3$

First the interval (0,1) is transformed to (-1,1) by $x = 2t - 1$, then by Table 1.5

$$y(x) = \frac{(x+1)^3}{8} = \frac{1}{8}(x^3 + 3x^2 + 3x + 1) =$$

$$= \frac{1}{8}\left(\frac{2P_3(x) + 3P_1(x)}{5} + 3\frac{2P_2(x) + P_0(x)}{3} + 3P_1(x) + P_0(x)\right) =$$

$$= \frac{1}{20}P_3(x) + \frac{1}{4}P_2(x) + \frac{9}{20}P_1(x) + \frac{1}{4}P_0(x).$$

By truncation we get that the Legendre least-squares parabola is equal to

$$\frac{1}{4}P_2(x) + \frac{9}{20}P_1(x) + \frac{1}{4}P_0(x) = \frac{1}{4}P_2(2t-1) + \frac{9}{20}P_1(2t-1) + \frac{1}{4}.$$

The same problem could be solved by the integral formulas (1.19), of course:

$$a_0 = \frac{1}{2}\int_{-1}^{1} y(x)P_0(x)\,dx = \frac{1}{2}\int_{-1}^{1} y(x)dx = \frac{1}{4}$$
$$a_1 = \frac{3}{2}\int_{-1}^{1} y(x)P_1(x)\,dx = \frac{3}{2}\int_{-1}^{1} y(x)xdx = \frac{9}{20}$$
$$a_2 = \frac{5}{2}\int_{-1}^{1} y(x)P_2(x)\,dx = \frac{5}{2}\int_{-1}^{1} y(x)\left(\frac{3x^2-1}{2}\right)dx = \frac{1}{4}$$

Example 1.9: Legendre continuous least-squares parabola on the interval (0,1) for $y(t) = \sin(\pi t)$:

First the interval (0,1) is transformed to (-1,1) by $x = 2t - 1$, then $y(x) = \sin\left(\pi \cdot \frac{x+1}{2}\right)$. By Theorem 1.10 and formula (1.19):

$$a_0 = \frac{1}{2}\int_{-1}^{1} y(x)P_0(x)\,dx = \frac{1}{2}\int_{-1}^{1} \sin\left(\pi \cdot \frac{x+1}{2}\right)dx = \frac{2}{\pi}$$
$$a_1 = \frac{3}{2}\int_{-1}^{1} y(x)P_1(x)\,dx = \frac{3}{2}\int_{-1}^{1} \sin\left(\pi \cdot \frac{x+1}{2}\right)\cdot xdx = 0$$
$$a_2 = \frac{5}{2}\int_{-1}^{1} y(x)P_2(x)\,dx = \frac{5}{2}\int_{-1}^{1} \sin\left(\pi \cdot \frac{x+1}{2}\right)\cdot\left(\frac{3x^2-1}{2}\right)dx = \frac{10(\pi^2-12)}{\pi^3}$$

From these values we get for the parabola $\frac{10(\pi^2-12)}{\pi^3}P_2(x) + 0P_1(x) + \frac{2}{\pi}P_0(x)$ = $\frac{10(\pi^2-12)}{\pi^3}P_2(2t-1) + 0P_1(2t-1) + \frac{2}{\pi}P_0(2t-1)$.

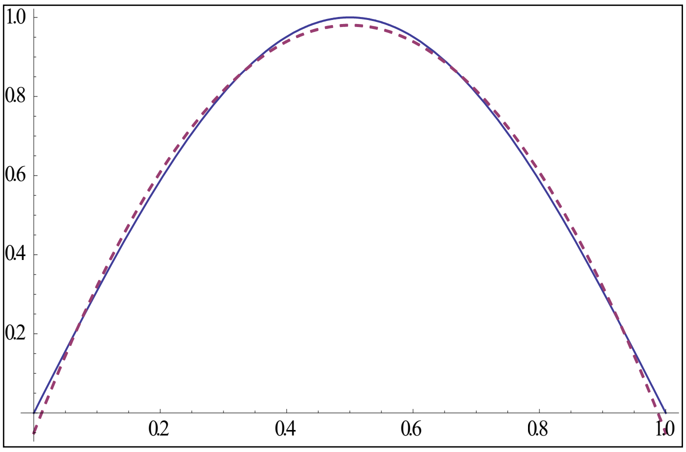

Figure 1.10: Plots showing $y(t) = \sin(\pi t)$ and the Legendre continuous least-squares parabola (dashed line) of Example 1.9 in the range $0 \le t \le 1$.

1.8 The matrix condition number

For a system of linear equations in matrix form $A \cdot x = b$ with square matrix A and right-hand side (r.h.s.) b the solution x obviously depends on b because $x = A^{-1} \cdot b$ provided that A is a regular (invertible) matrix.

The condition number of A, denoted by $\kappa(A)$, measures the maximum relative error of the solution x with respect to a relative error in the r.h.s. b:

Definition 1.10: Matrix condition number $\kappa(A)$

The condition number $\kappa(A)$ of a matrix A is defined by the maximum ratio of relative errors:

$$\kappa(A) = \max\left(\frac{\|\Delta x\| / \|x\|}{\|\Delta b\| / \|b\|} \right)$$

($b \neq 0, \Delta b \neq 0, A \cdot x = b$) if the matrix A is regular, and by

$\kappa(A) = \infty$ otherwise.

$\|\cdots\|$ denotes a vector norm, commonly this is the Euclidean norm $\|\cdots\|_2$ or the maximum norm $\|\cdots\|_\infty$.

In the case that $\kappa(A)$ is relatively small the matrix or the system of equations is called well-conditioned, otherwise one uses the term "ill-conditioned".

Property 1.4: If A is regular then $\kappa(A) = \|A\| \cdot \|A^{-1}\|$

wherein the (spectral) norm of a matrix is defined as $\|A\| = \max_{\|x\| \leq 1}(\|A \cdot x\|) = \max_{\|x\|=1}(\|A \cdot x\|)$. It depends on the vector norm $\|\cdots\|$.

(<u>Proof</u>: By linear computations: $\frac{\|\Delta x\| / \|x\|}{\|\Delta b\| / \|b\|} = \frac{\|A^{-1} \cdot \Delta b\| / \|x\|}{\|\Delta b\| / \|A \cdot x\|} = \left\| A^{-1} \cdot \Delta b / \|\Delta b\| \right\| \cdot \left\| A \cdot x / \|x\| \right\| =$

$\left\| A^{-1} \cdot \frac{\Delta b}{\|\Delta b\|} \right\| \cdot \left\| A \cdot \frac{x}{\|x\|} \right\|$. Since $\frac{\Delta b}{\|\Delta b\|}$ and $\frac{x}{\|x\|}$ are unit vectors the assertion follows by the definition of the spectral norm of a matrix

□)

The next propositions establish relations between the matrix condition number and the singular-value decomposition (SVD) or eigenvalues, respectively:

Property 1.5: For a regular matrix A and the Euclidean vector norm[1] we have:

a) $\kappa(A)=\frac{|\sigma_{\max}|}{|\sigma_{\min}|}$ in which $|\sigma_{\max}|$ and $|\sigma_{\min}|$ denote the absolute maximal and the absolute minimal singular value of A, respectively.

b) If, additionally, the matrix A is normal: $\kappa(A)=\frac{|\lambda_{\max}|}{|\lambda_{\min}|}$ in which $|\lambda_{\max}|$ and $|\lambda_{\min}|$ denote the absolute maximal and the absolute minimal eigenvalue, respectively.

c) $\kappa(A^*\cdot A)=\frac{|\sigma_{\max}|^2}{|\sigma_{\min}|^2}$ where again $|\sigma_{\max}|$ and $|\sigma_{\min}|$ denote the absolute maximal and the absolute minimal singular value of A, respectively.

Note to b): Normal means that A commutes with its conjugate transpose A^*: $A\cdot A^* = A^*\cdot A$.
A special case of this are symmetric matrices.

(Proof: Note that $|\sigma_{\max}|=\|A\|$ and $|\sigma_{\min}|^{-1}=\|A^{-1}\|$ by the definition of the singular values. This yields a).
For b) and normal matrices use the fact that the singular values correspond to the absolute eigenvalues. Note that, generally, the singular values are the square roots of the non-negative, real eigenvalues of $A^*\cdot A$ (or equivalently $A\cdot A^*$).
Finally, c) is concluded from a) and b) since the matrix $A^*\cdot A$ as a symmetric matrix is normal
□)

Example 1.10: Continuation of Example 1.7A

For the design matrix G and its singular-value decomposition (cf. Example 1.7A above) we conclude by Proposition 1.5c that $\kappa(G^*\cdot G)=\frac{|\sigma_{\max}|^2}{|\sigma_{\min}|^2}=\left(3.13397\times 10^8 / 4.50267\right)^2$. Thus the 10 x 10 system of normal equations is ill-conditioned.

Example 1.11: Statistical normalization (cont. Example 1.10)

The variables in the (complete) basis of total degree 3 in Example 1.7A are statistically normalized[2] by the substitutions:

$$x \mapsto \frac{x-\mu_x}{\sigma_x}, \quad y \mapsto \frac{y-\mu_y}{\sigma_y} \qquad \textbf{(1.20)}$$

[1] For the maximum norm $\|\cdots\|_\infty$ and a regular, lower triangular matrix A: $\kappa(A)\geq\frac{\max|a_{ii}|}{\min|a_{ii}|}$.

[2] The new variables thus have mean 0 and standard deviation 1.

The singular-value decomposition for the new design matrix yields (10 rows of the diagonal matrix with singular values in the diagonal):

$$\begin{pmatrix} 235.075 & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\ 0. & 220.228 & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\ 0. & 0. & 171.286 & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\ 0. & 0. & 0. & 128.405 & 0. & 0. & 0. & 0. & 0. & 0. \\ 0. & 0. & 0. & 0. & 107.539 & 0. & 0. & 0. & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. & 99.3667 & 0. & 0. & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. & 0. & 73.5078 & 0. & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. & 0. & 0. & 29.4099 & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 24.0546 & 0. \\ 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 21.5343 \end{pmatrix}$$

It is obvious that the 10 x 10 system of normal equations now is well-conditioned; the condition number being $(10.9161)^2$.