Inhalt:

Allgemeine Problemstellung

I. Gegeben ist eine Messreihe von Daten (x_j, y_j), $j = 0, \ldots, n$. Man bestimme eine "einfache" Funktion $u : [a, b] \to \mathbb{R}$ mit
Interpolation: $u(x_j) = y_j$ für $j = 0, \ldots, n$ (siehe Abschnitte 3.1, 7.1, 7.2)
Approximation: $u(x_j) \approx y_j$ für $j = 0, \ldots, n$ (siehe Abschnitte 6.1, 7.3)

II. Gegeben ist eine Funktion $f : [a, b] \to \mathbb{R}$. Man bestimme eine "einfache" Funktion $u : [a, b] \to \mathbb{R}$,
Interpolation: die zu gegebenen Argumenten x_j, $j = 0, \ldots, n$, die Funktionswerte $u(x_j) = f(x_j)$ besitzt (siehe Abschnitte 3.1, 7.1, 7.2)
Approximation: für die $\|f - u\|$ möglichst klein ist, wobei $\|.\|$ eine Norm auf $C[a, b]$ ist. (siehe Abschnitte 3.3, 7.3)

Als "einfach" bezeichnet man z.B. Polynome, rationale Funktionen oder trigonometrische Polynome.

Matlab/Octave: Interpolation von 4 Messwerten durch ein kubisches Polynom

```
x=[0,1,2,3];
y=[1,-1,1,-1]; % oder y=cos(pi*x)
c=polyfit(x,y,3); % kubisches Interpolationspolynom
xx=linspace(0,3,101);
plot(x,y,'o');
hold on
plot(xx,polyval(c,xx));
```

Approximation der Bevölkerungszahl der USA 1790-1990:

```
load census; % laedt `cdate' und `pop'
c=polyfit(cdate,pop,2); % quadratisches Ausgleichspolynom
xx=linspace(1790,1990,101);
plot(cdate,pop,'o');
hold on
plot(xx,polyval(c,xx));
```

3.1 Polynominterpolation

3.1.1 Vektorraum der Polynome vom Höchstgrad n

Für $n \in \mathbb{N}_0$ ist

$$\mathcal{P}_n = \{p : \mathbb{R} \to \mathbb{R} \mid p(x) = a_0 + a_1 x + \cdots + a_n x^n,\ a_j \in \mathbb{R}\}$$

der Vektorraum der Polynome mit $\mathrm{Grad}(p) \leq n$. Es gilt $\dim(\mathcal{P}_n) = n + 1$, und die *Monome* $e_j : \mathbb{R} \to \mathbb{R}$ mit $e_j(x) = x^j$, $j = 0, \ldots, n$, sind eine Basis von $\mathcal{P}_n$.

Bemerkung: Mit Polynomen kann man auf einem kompakten Intervall jede stetige Funktion beliebig genau annähern:

Satz von Weierstrass

Es sei I ein kompaktes Intervall und $f \in C(I)$. Dann gibt es zu jedem $\varepsilon > 0$ ein Polynom p (mit $\mathrm{Grad}(p)$ abhängig von ε), so dass

$$\|f - p\|_\infty := \max_{x \in I} |f(x) - p(x)| < \varepsilon.$$

3.1.2 Erinnerung: Horner-Schema

Zur Berechnung von

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n = a_0 + x \cdot (a_1 + x\,(a_2 + \cdots + x\,(a_{n-1} + x a_n) \cdots))$$

verwendet man das Horner-Schema aus Kapitel 1:

- Eingabe: $a_k^{(0)} = a_k$, $k = 0, \ldots, n$, und Stelle ξ
- Setze $a_n^{(1)} = a_n^{(0)}$.
  Berechne für $k = n-1, \ldots, 0$
  $$a_k^{(1)} = a_k^{(0)} + \xi a_{k+1}^{(1)}.$$
- Ergebnis: $p(\xi) = a_0^{(1)}$

Rechenaufwand: n Multiplikationen/Additionen

3.1.3 Erinnerung: vollständiges Horner-Schema

Zur Berechnung der Taylor-Entwicklung von

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n = b_0 + b_1(x-\xi) + \cdots b_n(x-\xi)^n$$

an der Stelle ξ verwendet man das vollständige Horner-Schema von Übungsblatt 1:

- Eingabe: $a_k^{(0)} = a_k$ für $k = 0, \ldots, n$, und Stelle ξ
- Für $j = 0, \ldots, n-1$,
  setze $a_n^{(j+1)} = a_n^{(j)}$,
  berechne für $k = n-1, \ldots, j$

$$a_k^{(j+1)} = a_k^{(j)} + \xi a_{k+1}^{(j+1)}.$$

- Ergebnis: $b_j = a_j^{(j)}$

Rechenaufwand: $n(n+1)/2$ Multiplikationen/Additionen

3.1.4 Definition: Interpolationspolynom

Gegeben seien Punkte $(x_j, y_j) \in \mathbb{R}^2$, $j = 0, \ldots, n$, mit paarweise verschiedenen $x_j \in \mathbb{R}$.

- Die Zahlen x_j heißen die *Stützstellen* (oder *Knoten*) der *Lagrange-Interpolation*.
- Die Zahlen y_j heißen die *Daten* (oder *Knotenwerte*).
- Ein Polynom $p \in \mathcal{P}_n$ mit $p(x_j) = y_j$ für alle $j = 0, \ldots, n$ heißt *Interpolationspolynom*.

3.1.5 Hauptsatz:

Die Lagrange-Interpolationsaufgabe ist eindeutig lösbar. D.h. zu paarweise verschiedenen Stützstellen $x_j \in \mathbb{R}$, $j = 0, \ldots, n$, und beliebigen Daten $y_j \in \mathbb{R}$ existiert genau ein Polynom $p \in \mathcal{P}_n$ mit $p(x_j) = y_j$ für alle $j = 0, \ldots, n$.

4 Varianten zur Darstellung/Berechnung des Interpolationspolynoms $p \in \mathcal{P}_n$ liefern jeweils unterschiedliche Beweise von Satz 3.1.5.

3.1.6 Das Interpolationspolynom in der Monombasis

Zu paarweise verschiedenen Stützstellen $x_j \in \mathbb{R}$, $j = 0, \ldots, n$, und beliebigen Daten $y_j \in \mathbb{R}$ ist das Interpolationspolynom gegeben durch $p(x) = \sum_{k=0}^{n} a_k x^k$, wobei der Koeffizientenvektor $(a_0, a_1, \ldots, a_n)^T \in \mathbb{R}^{n+1}$ die eindeutige Lösung des linearen Gleichungssystems

$$M\vec{a} = \vec{y}, \qquad M = \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ & \vdots & & & \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}$$

ist. Die Matrix M ist regulär und hat die *Vandermonde-Determinante*

$$\det M = \prod_{0 \le i < j \le n} (x_j - x_i).$$

Beweis: Der Determinanten-Multiplikationssatz ergibt

$$
\begin{aligned}
\det M &= \det \begin{pmatrix} 1 & & & \\ -1 & 1 & & \\ \vdots & & \ddots & \\ -1 & & & 1 \end{pmatrix} \underbrace{\begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ & \vdots & & & \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{pmatrix} \begin{pmatrix} 1 & -x_0 & & \\ & \ddots & \ddots & \\ & & 1 & -x_0 \\ & & & 1 \end{pmatrix}}_{= \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & x_1 - x_0 & x_1(x_1 - x_0) & \cdots & x_1^{n-1}(x_1 - x_0) \\ & \vdots & & & \\ 1 & x_n - x_0 & x_n(x_n - x_0) & \cdots & x_n^{n-1}(x_n - x_0) \end{pmatrix}} \\
&= \det \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & x_1 - x_0 & x_1(x_1 - x_0) & \cdots & x_1^{n-1}(x_1 - x_0) \\ & \vdots & & & \\ 0 & x_n - x_0 & x_n(x_n - x_0) & \cdots & x_n^{n-1}(x_n - x_0) \end{pmatrix} = \prod_{j=1}^{n} (x_j - x_0) \cdot \det \widetilde{M},
\end{aligned}
$$

wobei $\widetilde{M}$ die Vandermonde-Matrix zu den Knoten $x_1, \ldots, x_n$ ist. Das Herausziehen des Produktes erfolgt mit der Linearität von $\det A$ bezüglich jeder Zeile von A. Die Aussage des Satzes folgt per Induktion.

3.1.7 Das Interpolationspolynom in der Lagrange-Basis

Zu $n+1$ paarweise verschiedenen Stützstellen $x_j \in \mathbb{R}$, $j = 0, \ldots, n$, definieren wir die *Lagrange-Grundpolynome*

$$L_{n,k}(x) = \prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{x - x_j}{x_k - x_j} \in \mathcal{P}_n, \quad k = 0, \ldots, n.$$

- Die Lagrange-Grundpolynome $\{L_{n,k} : k = 0, \ldots, n\}$ bilden eine Basis von $\mathcal{P}_n$.
- Es gilt
$$L_{n,k}(x_j) = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$
- Das Interpolationspolynom zu den Stützstellen x_j, $j = 0, \ldots, n$, und Daten $y_j \in \mathbb{R}$ ist
$$p(x) = \sum_{k=0}^{n} y_k L_{n,k}(x).$$

3.1.8 Bemerkung:

(i) Mit dem Ansatz $p(x) = \sum_{k=0}^{n} c_k L_{n,k}(x)$ ergeben die Interpolationsbedingungen das lineare Gleichungssystem

$$A\vec{c} = \vec{y}, \qquad A = \left(L_{n,k}(x_j)\right)_{j,k=0,\dots,n} = \mathrm{I},$$

dessen Lösung $\vec{c} = \vec{y}$ sofort abzulesen ist.

(ii) Kurz-Schreibweise für die Lagrange-Grundpolynome: mit Hilfe des *Knotenpolynoms*

$$w(x) = \prod_{j=0}^{n}(x - x_j) \in \mathcal{P}_{n+1}$$

ist

$$L_{n,k}(x) = \frac{w(x)}{(x - x_k)w'(x_k)}, \quad k = 0, \dots, n.$$

(iii) Der Nachteil der Lagrange-Darstellung des Interpolationspolynoms $p \in \mathcal{P}_n$ ist, dass bei Hinzunahme eines weiteren Stützpunktes (x_{n+1}, y_{n+1}) oder bei der Änderung eines Stützpunktes (x_j, y_j) die Basisfunktionen $L_{n,k}$ sich völlig ändern. Deshalb ist diese Darstellung des Interpolationspolynoms für die meisten praktischen Zwecke zu aufwändig.

3.1.9 Das Interpolationspolynom in der Newton-Basis

Zu $n+1$ paarweise verschiedenen Stützstellen $x_j \in \mathbb{R}$, $j = 0, \ldots, n$, definieren wir die *Newton-Grundpolynome*

$$N_0(x) = 1; \qquad N_k(x) = \prod_{j=0}^{k-1} (x - x_j) \in \mathcal{P}_k, \quad k = 1, \ldots, n.$$

- Die Newton-Grundpolynome $\{N_k : k = 0, \ldots, n\}$ bilden eine Basis von $\mathcal{P}_n$.
- Es gilt
$$N_k(x_j) = 0 \quad \text{für} \quad k > j.$$
- Das Interpolationspolynom zu den Stützstellen x_j, $j = 0, \ldots, n$, und Daten $y_j \in \mathbb{R}$ ist
$$p(x) = \sum_{k=0}^{n} y[x_0, \ldots, x_k] N_k(x),$$
wobei $y[x_0, \ldots, x_k]$ die k-te dividierte Differenz zu den Punkten (x_j, y_j), $j = 0, \ldots, n$, bezeichnet.

3.1.10 Bemerkung:

(i) Mit dem Ansatz $p(x) = \sum_{k=0}^{n} c_k N_k(x)$ ergeben die Interpolationsbedingungen das lineare Gleichungssystem

$$A\vec{c} = \vec{y}, \qquad A = \big(N_k(x_j)\big)_{j,k=0,\ldots,n};$$

die Matrix A ist eine untere Dreiecksmatrix, das Auflösen kann also durch Vorwärtseinsetzen erfolgen. Die dividierten Differenzen bilden einen numerisch stabileren Algorithmus zur Lösung.

(ii) In der Newton-Darstellung ist die Teilsumme

$$p_{0,m}(x) = \sum_{k=0}^{m} y[x_0,\ldots,x_k]N_k(x) \in \mathcal{P}_m, \qquad 0 \le m \le n,$$

das Interpolationspolynom zu den Daten $(x_0, y_0), \ldots, (x_m, y_m)$.

Deshalb kann auch ein weiterer Punkt (x_{n+1}, y_{n+1}) leicht hinzugenommen werden: das neue Interpolationspolynom $p_{0,n+1}$ ergibt sich als

$$p_{0,n+1}(x) = p_{0,n}(x) + y[x_0,\ldots,x_{n+1}]N_{n+1}(x).$$

(iii) Die dividierte Differenz $y[x_0,\ldots,x_n]$ ist der *Höchstkoeffizient* des Interpolationspolynoms in der Monom-Darstellung:

$$p(x) = a_0 + \ldots + a_{n-1}x^{n-1} + a_nx^n \quad \text{mit} \quad a_n = y[x_0,\ldots,x_n].$$

3.1.11 Definition: dividierte Differenzen

Zu $n+1$ paarweise verschiedenen Stützstellen $x_j \in \mathbb{R}$, $j = 0, \dots, n$, und Daten y_j sind die *dividierten Differenzen* rekursiv definiert durch

- Ordnung $k = 0$:
$$y[x_j] = y_j, \quad j = 0, \dots, n,$$
- Ordnung $1 \le k \le n$:
$$y[x_j, \dots, x_{j+k}] = \frac{y[x_{j+1}, \dots, x_{j+k}] - y[x_j, \dots, x_{j+k-1}]}{x_{j+k} - x_j}, \quad j = 0, \dots, n-k.$$

Berechnungs-Schema:

$$\begin{array}{rrr|c|cccc}
x_n - x_0 \cdots x_2 - x_0 & & x_1 - x_0 & x_0 & y_0 & y[x_0, x_1] & y[x_0, x_1, x_2] \cdots y[x_0, \dots, x_n] \\
\cdots x_3 - x_1 & & x_2 - x_1 & x_1 & y_1 & y[x_1, x_2] & y[x_1, x_2, x_3] \cdots \\
 & & \vdots & \vdots & \vdots & \vdots & \\
 & & x_n - x_{n-1} & x_{n-1} & y_{n-1} & y[x_{n-1}, x_n] & \\
 & & & x_n & y_n & &
\end{array}$$

3.1.12 Lemma (zur Newton-Darstellung des Interpolationspolynoms):

Es seien $n+1$ paarweise verschiedene Stützstellen $x_j \in \mathbb{R}$, $j = 0, \ldots, n$, und Daten $y_j \in \mathbb{R}$ gegeben. Mit

$$p_{j,j+k} \in \mathcal{P}_k, \qquad 0 \le k \le n, \quad 0 \le j \le n-k$$

bezeichnen wir das Interpolationspolynom zu den Punkten

$$(x_j, y_j), \ldots, (x_{j+k}, y_{j+k})$$

.

Dann gilt

$$\begin{aligned} p_{j,j+k}(x) \quad &= \quad y[x_j] + y[x_j, x_{j+1}](x - x_j) + \cdots + \\ &\qquad y[x_j, \ldots, x_{j+k}](x - x_j) \cdots (x - x_{j+k-1}). \end{aligned}$$

Beweis: Induktion nach k:

- Für $k = 0$ und $0 \leq j \leq n$ ist $p_{j,j}(x) = y_j = y[x_j]$ das konstante Interpolationspolynom zum Punkt (x_j, y_j).
- Sei $k \geq 1$ und $0 \leq j \leq n - k$. Nach Induktionsannahme interpoliert

$$p_{j,j+k-1}(x) = y[x_j] + \cdots + y[x_j, \ldots, x_{j+k-1}](x - x_j) \cdots (x - x_{j+k-2}) \in \mathcal{P}_{k-1}$$

die Punkte $(x_j, y_j), \ldots, (x_{j+k-1}, y_{j+k-1})$, und ebenso interpoliert

$$p_{j+1,j+k}(x) = y[x_{j+1}] + \cdots + y[x_{j+1}, \ldots, x_{j+k}](x - x_{j+1}) \cdots (x - x_{j+k-1}) \in \mathcal{P}_{k-1}$$

die Punkte $(x_{j+1}, y_{j+1}), \ldots, (x_{j+k}, y_{j+k})$. Deshalb interpoliert

$$q(x) = \frac{(x - x_j)p_{j+1,j+k}(x) + (x_{j+k} - x)p_{j,j+k-1}(x)}{x_{j+k} - x_j}$$

die Punkte $(x_j, y_j), \ldots, (x_{j+k}, y_{j+k})$, ist also das gesuchte Interpolationspolynom $p_{j,j+k}$. Der Höchstkoeffizient von q (also der Vorfaktor von x^k) berechnet sich aus den Höchstkoeffizienten von $p_{j,j+k-1}$ und $p_{j+1,j+k}$,

$$a := HK(q) = \frac{y[x_{j+1}, \ldots, x_{j+k}] - y[x_j, \ldots, x_{j+k-1}]}{x_{j+k} - x_j} = y[x_j, \ldots, x_{j+k}].$$

Andererseits gilt (durch Hinzunahme des Punktes (x_{j+k}, y_{j+k}) zu $p_{j,j+k-1}$, siehe Bemerkung 3.1.10(ii))

$$q(x) = p_{j,j+k}(x) = p_{j,j+k-1}(x) + a(x - x_j) \cdots (x - x_{j+k-1}).$$

Damit hat $p_{j,j+k}$ die behauptete Form.

3.1.13 Bemerkung:

(i) Die dividierte Differenz $y[x_0, \ldots, x_n]$ ist der Höchstkoeffizient des Interpolationspolynoms $p \in \mathcal{P}_n$ zu den Punkten $(x_0, y_0), \ldots, (x_n, y_n)$.

(ii) Die dividierte Differenz $y[x_0, \ldots, x_n]$ ist invariant gegenüber einer Index-Permutation in der Aufzählung der Punkte $(x_0, y_0), \ldots, (x_n, y_n)$. Insbesondere brauchen die Stützstellen x_j nicht sortiert vorzuliegen.

(iii) Bei Hinzunahme eines Punktes (x_{n+1}, y_{n+1}) wird das Schema der dividierten Differenzen unten um eine Diagonale ergänzt.

Zur Auswertung des Interpolationspolynoms an einer Stelle $\xi \in \mathbb{R}$ eignet sich die Rekursion im Beweis von Lemma 3.1.12.

3.1.14 Das Interpolationspolynom in der Neville-Aitken-Form

Zu $n+1$ paarweise verschiedenen Stützstellen $x_j \in \mathbb{R}$, $j = 0, \dots, n$, und Daten y_j berechnet man den Wert $p(\xi) = p_{0,n}(\xi)$ des Interpolationspolynoms rekursiv gemäß

- $k = 0$: $\quad p_{j,j}(\xi) = y_j$ für $j = 0, \dots, n$,
- $1 \le k \le n$: $\quad p_{j,j+k}(\xi) = p_{j,j+k-1}(\xi) + (\xi - x_j)\dfrac{p_{j+1,j+k}(\xi) - p_{j,j+k-1}(\xi)}{x_{j+k} - x_j}$

  für $j = 0, \dots, n-k$.

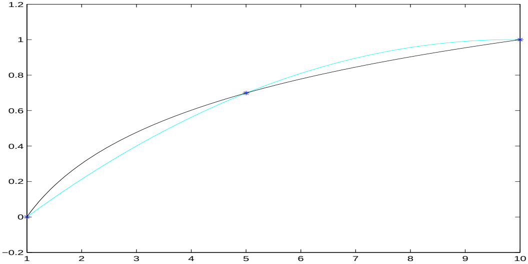
Schema:

$$
\begin{array}{c|c|cccccc}
x_0 & y_0 & p_{0,1}(\xi) & p_{0,2}(\xi) & p_{0,3}(\xi) & \dots & p_{0,n-1}(\xi) & p_{0,n}(\xi) \\
x_1 & y_1 & p_{1,2}(\xi) & p_{1,3}(\xi) & p_{1,4}(\xi) & \dots & p_{1,n}(\xi) & \\
x_2 & y_2 & p_{2,3}(\xi) & p_{2,4}(\xi) & p_{2,5}(\xi) & \dots & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & & & \\
x_{n-1} & y_{n-1} & p_{n-1,n}(\xi) & & & & & \\
x_n & y_n & & & & & &
\end{array}
$$

Erweiterte Problemstellung:

- Gegeben sei eine Funktion
$$f \in C^{n+1}[a, b].$$
Zu paarweise verschiedenen Stützstellen $x_j \in [a, b]$ werden die Daten $y_j = f(x_j)$ dem Graphen von f entnommen.
- $\mathrm{conv}(x_0, \ldots, x_n)$ bezeichnet das kleinste Intervall, das alle x_j, $j = 0, \ldots, n$, enthält, also die konvexe Hülle der Menge $\{x_0, \ldots, x_n\}$).

Vergleich von $f(x) = \log_{10}(x)$ auf $[a, b] = [1, 10]$ (schwarz) und dem quadratischen Interpolationspolynom zu den Stützstellen $x_j = 1, 5, 10$ (cyan)



3.1.15 Satz: Interpolationsfehler

Es seien $f : [a,b] \to \mathbb{R}$ und paarweise verschiedene Stützstellen $x_0, \ldots, x_n \in [a,b]$ gegeben. $p \in \mathcal{P}_n$ sei das Interpolationspolynom zu den Punkten $(x_j, f(x_j))$, $j = 0, \ldots, n$. Weiter sei $x \in [a,b]$.
Dann ist der Interpolationsfehler $f(x) - p(x)$ gegeben in *Newton-Form*

$$f(x) - p(x) = f[x_0, \ldots, x_n, x] \prod_{j=0}^{n} (x - x_j),$$

mit der dividierten Differenz zu den Punkten $(x_0, f(x_0)), \ldots, (x_n, f(x_n)), (x, f(x))$, bzw. in *Lagrange-Form*

$$f(x) - p(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \prod_{j=0}^{n} (x - x_j),$$

mit einem $\xi_x \in \operatorname{conv}(x_0, \ldots, x_n, x)$, falls $f \in C^{n+1}[a,b]$ gilt.

Die dividierte Differenz $f[x_0, \dots, x_n]$ zu den Punkten $(x_0, f(x_0)), \dots, (x_n, f(x_n))$ besitzt zwei interessante Darstellungen.

3.1.16 Folgerung:

Es seien $n \in \mathbb{N}_0$, $f \in C^n[a, b]$ und $x_j \in [a, b]$, $j = 0, \dots, n$, (paarweise verschiedene) Stützstellen.

a) Es existiert $\xi \in \operatorname{conv}(x_0, \dots, x_n)$ mit

$$f[x_0, \dots, x_n] = \frac{f^{(n)}(\xi)}{n!}.$$

b) Für $n \geq 1$ gilt

$$f[x_0, \dots, x_n] =$$
$$\int_0^1 \int_0^{t_1} \cdots \int_0^{t_{n-1}} f^{(n)}(x_0 + t_1(x_1 - x_0) + \dots + t_n(x_n - x_{n-1}))\, dt_n \cdots dt_2\, dt_1.$$

Beachte:

$$\int_0^1 \int_0^{t_1} \cdots \int_0^{t_{n-1}} dt_n \cdots dt_2\, dt_1 = \mathrm{vol}(\text{Standard-Simplex im } \mathbb{R}^n) = 1/n!$$

Im Beweis von b) führt man die innere Integration aus:

$$(x_n - x_{n-1}) \int_0^1 \int_0^{t_1} \cdots \int_0^{t_{n-1}} f^{(n)}(x_0 + t_1(x_1 - x_0) + \ldots + t_n(x_n - x_{n-1}))\, dt_n \cdots dt_2\, dt_1 =$$

$$\int_0^1 \int_0^{t_1} \cdots \int_0^{t_{n-2}} \Big(f^{(n-1)}(x_0 + t_1(x_1 - x_0) + \ldots + t_{n-1}(x_n - x_{n-2})) -$$
$$f^{(n-1)}(x_0 + t_1(x_1 - x_0) + \ldots + t_{n-1}(x_{n-1} - x_{n-2})) \Big)\, dt_{n-1} \cdots dt_2\, dt_1 =$$

$$f[x_0, x_1, \ldots, x_{n-2}, x_n] - f[x_0, x_1, \ldots, x_{n-2}, x_{n-1}] = \qquad \text{nach Ind.-Annahme}$$

$$f[x_0, x_1, \ldots, x_{n-2}, x_n] - f[x_{n-1}, x_0, x_1, \ldots, x_{n-2}] \qquad \text{Vertauschung der Stützstellen}$$

3.1.17 Bemerkung: Die Definition der dividierten Differenz von $f \in C^n[a, b]$ für zusammenfallende Knoten geschieht mittels sogenannter “Konfluenz”: für zusammenfallende Knoten $x_0 = x_1$ ist

$$f[x_0, x_0] = \lim_{h \to 0} f[x_0, x_0 + h] = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0).$$

Die Integraldarstellung bleibt in diesem Fall gültig,

$$f[x_0, x_0] = f'(x_0) = \int_0^1 f'(x_0 + t(x_0 - x_0))\, dt = \int_0^1 f'(x_0)\, dt.$$

Bei mehrfacher Wiederholung der Stützstelle $x_j = \cdots = x_{j+k}$ ist

$$f[\underbrace{x_j, \ldots, x_j}_{(k+1)-\text{mal}}] = \frac{f^{(k)}(x_j)}{j!} = \int_0^1 \int_0^{t_1} \cdots \int_0^{t_{k-1}} f^{(k)}(x_j)\, dt_k \cdots dt_2\, dt_1.$$

Sind die Stützstellen $x_0 \le x_1 \le \cdots \le x_n$ angeordnet, so werden im Schema 3.1.11 die nicht-existierenden Quotienten (Teilen durch Null) durch die entsprechenden Ableitungsterme ersetzt. Dadurch bleibt die rekursive Berechnung von $f[x_0, \ldots, x_n]$ gültig, auch wenn Stützstellen zusammenfallen.

Für mehrfache Stützstellen stellt sich eine modifizierte Interpolationsaufgabe.

3.1.18 Hermite-Interpolation

Es seien $x_0, \dots, x_m \in \mathbb{R}$ paarweise verschieden,
$\mu_0, \dots, \mu_m \in \mathbb{N}_0$ und

Daten $y_j^{(k)}$, $j = 0, \dots, m$, $k = 0, \dots, \mu_j$ gegeben. Weiter sei $n = \sum_{j=0}^{m} (1 + \mu_j) - 1$.

Ein Polynom $p \in \mathcal{P}_n$ mit

$$p^{(k)}(x_j) = y_j^{(k)} \quad \text{für alle} \quad j = 0, \dots, m, \quad k = 0, \dots, \mu_j,$$

heißt *Hermite-Interpolationspolynom*.

3.1.19 Satz zur Hermite-Interpolation

Die Hermite-Interpolationsaufgabe ist eindeutig lösbar.

Mit dem *erweiterten Knotenvektor*

$$(\xi_0, \xi_1, \ldots, \xi_n) = (\underbrace{x_0, \ldots, x_0}_{\mu_0+1-\text{fach}}, \ldots, \underbrace{x_m, \ldots, x_m}_{\mu_m+1-\text{fach}})$$

und der Definition dividierter Differenzen mit mehrfachen Knoten ist das Interpolationspolynom gegeben in der Newton-Form

$$p(x) = \sum_{k=0}^{n} y[\xi_0, \ldots, \xi_k](x - \xi_0) \cdots (x - \xi_{k-1}).$$

3.1.20 Bemerkung: Die Darstellungen des Interpolationsfehlers für Daten

$$y_j^{(k)} = f^{(k)}(x_j), \qquad j = 0, \dots, m, \quad k = 0, \dots, \mu_j,$$

bleiben exakt wie in Satz 3.1.15 erhalten.

3.1.21 Diskussion: Interpolationsfehler bei f mit beschränkten Ableitungen

Mit $\max_{x\in[a,b]} |f^{(n+1)}(x)| =: M_{n+1}$ gilt

$$|f(x) - p_{0,n}(x)| \le \frac{M_{n+1}}{(n+1)!} \prod_{j=0}^{n} |x - x_j|.$$

Für äquidistante Knoten $x_j = a + jh$, $j = 0, \ldots, n$, $h = (b-a)/n$, ist weiterhin $\prod_{j=0}^{n} |x - x_j| \le n! h^{n+1}$, also insgesamt

$$|f(x) - p_{0,n}(x)| \le \frac{M_{n+1}}{n+1} \left(\frac{b-a}{n}\right)^{n+1}.$$

- Falls
$$\frac{M_{n+1}}{n+1} \left(\frac{b-a}{n}\right)^{n+1} = \mathbf{o}(1) \quad \text{für} \quad n \to \infty$$
gilt, so konvergiert die Folge $(p_{0,n})_{n\ge 0}$ der Interpolationspolynome gleichmäßig gegen f.

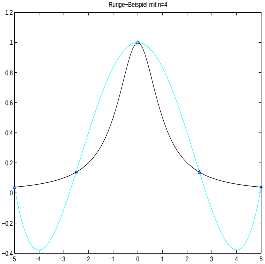
- Ist die Folge $(M_n)_{n\ge 0}$ sogar beschränkt (z.B. für $f(x) = e^x$ auf $[a,b]$), so ist die Konvergenz sehr schnell:
$$|f(x) - p_{0,n}(x)| = \mathcal{O}\left(\frac{(b-a)^{n+1}}{n^{n+2}}\right) \quad \text{für} \quad n \to \infty.$$

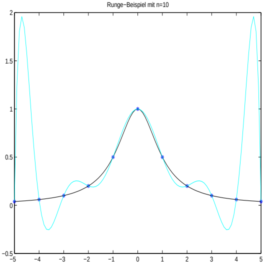
3.1.22 Diskussion: Interpolationsfehler bei f mit wachsenden Ableitungen

Ein klassisches Beispiel von Runge ist die Funktion

$$f : [-5, 5] \to \mathbb{R}, \qquad f(x) = \frac{1}{1+x^2}.$$

Interpolation mit äquidistanten Knoten $x_j = -5 + jh$, $j = 0, \ldots, n$, $h = 10/n$, führt schon für $n = 10$ zu unbrauchbarem Interpolationspolynom. Tatsächlich **divergiert** die Folge der Interpolationspolynome $(p_{0,n})_{n \geq 0}$.





3.2 Extrapolation zum Limes

3.2.1 Beispiel (vgl. 0.0.4 in der Einleitung):

Berechne: $a_0 = \lim_{h \to 0} f(h)$ für $f(x) = \frac{\tan x - x}{x^3}$.

- Numerische Rechnung (doppelt genau) ergibt für $h_j = 10^{-(j+1)}$ mit $j = 0, 1, 2$ die Ergebnisse

$$0.33467208545054, \quad 0.33334666720702, \quad 0.33333346673159$$

- Mit der Taylor-Reihe

$$\tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \frac{17}{315}x^7 + \cdots$$

ergibt sich

$$f(x) = \frac{1}{3} + \frac{2}{15}x^2 + \frac{17}{315}x^4 + \cdots ;$$

insbesondere ist f gerade, besitzt also eine Entwicklung mit geraden Potenzen von x.

- Daher ist es sinnvoll, a_0 anzunähern durch den Wert $p_{0,1}(0)$ des *linearen* Interpolationspolynoms zu den Punkten $(h_0^2, f(h_0))$, $(h_1^2, f(h_1))$, also nach dem Neville-Schema

$$p_{0,1}(0) = f(h_1) + \frac{1}{(h_0/h_1)^2 - 1}\,(f(h_1) - f(h_0)) = 0.33333327914396.$$

- Analog ergibt das *lineare* Interpolationspolynom zu den Punkten $(h_1^2, f(h_1))$, $(h_2^2, f(h_2))$

$$p_{1,2}(0) = f(h_2) + \frac{1}{(h_1/h_2)^2 - 1}\,(f(h_2) - f(h_1)) = 0.33333333339345.$$

- Weiterführung zum *quadratischen* Interpolationspolynom zu den Punkten $(h_0^2, f(h_0))$, $(h_1^2, f(h_1))$, $(h_2^2, f(h_2))$ liefert

$$p_{0,2}(0) = p_{1,2}(0) + \frac{1}{(h_0/h_2)^2 - 1}(p_{1,2}(0) - p_{0,1}(0)) = 0.33333333339888,$$

  also keine weitere Verbesserung zum exakten Wert $1/3$.

3.2.2 Beispiele: Differenzenquotienten

- Für eine $(r+1)$-mal stetig differenzierbare Funktion $f : [a, b] \to \mathbb{R}$ und $x_0 \in [a, b]$ gilt

$$a(h) := \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0) + \sum_{j=1}^{r} \frac{f^{(j+1)}(x_0)}{(j+1)!} h^j + \mathbf{o}(h^r).$$

Für verschiedene Werte $h_j > 0$, $j = 0, \ldots, r$, stellt man das Neville-Schema zur Berechnung der Interpolationspolynome

$$p_{j,j+k} \quad \text{zu den Punkten} \quad (h_j, a(h_j)), \ldots, (h_{j+k}, a(h_{j+k}))$$

und für die Auswertung bei $\xi = 0$ auf. Aus den Näherungswerten

$$a_{j,0} = p_{j,j}(0) = a(h_j) \quad \text{für} \quad j = 0, \ldots, r$$

werden (bessere) Näherungswerte berechnet:

Für $1 \leq k \leq r$ und $0 \leq j \leq r - k$ setze

$$a_{j+k,k} = p_{j,j+k}(0) = p_{j+1,j+k}(0) + \frac{1}{(h_j/h_{j+k}) - 1}(p_{j+1,j+k}(0) - p_{j,j+k-1}(0)).$$

- Wählt man stattdessen den symmetrischen Differenzenquotienten

$$b(h) := \frac{f(x_0+h) - f(x_0-h)}{2h} = f'(x_0) + \sum_{j=1}^{[r/2]} \frac{f^{(2j+1)}(x_0)}{(2j+1)!}\, h^{2j} + \mathbf{o}(h^r),$$

so werden Interpolationspolynome

$$q_{j,j+k} \quad \text{zu den Punkten} \quad (h_j^2, b(h_j)), \ldots, (h_{j+k}^2, b(h_{j+k}))$$

bei $\xi = 0$ ausgewertet. Aus den Näherungswerten

$$b_{j,0} = q_{j,j}(0) = b(h_j) \quad \text{für} \quad j = 0, \ldots, [r/2]$$

werden (bessere) Näherungswerte berechnet:

Für $1 \leq k \leq [r/2]$ und $0 \leq j \leq [r/2] - k$ setze

$$b_{j+k,k} = q_{j,j+k}(0) = q_{j+1,j+k}(0) + \frac{1}{(h_j/h_{j+k})^2 - 1}(q_{j+1,j+k}(0) - q_{j,j+k-1}(0)).$$

Für $f(x) = e^x$ an der Stelle $x_0 = 0$, $h_j = 2^{-j}$ für $j = 1, \dots, 5$:

- Extrapolation für Differenzenquotienten ($q = 1$ in Satz 3.2.3)

```
tab=extrapolation_tab('bsp_diffqu_exp',2.^-[1:5],1)

tab =

   1.29744254140026                  0                  0                  0
   1.13610166675097   0.97476079210167                  0                  0
   1.06518762453461   0.99427358231826   1.00077784572378                  0
   1.03191134268575   0.99863506083689   1.00008888700977   0.99999046433634
   1.01578903997129   0.99966673725682   1.00001062939680   0.99999944973780
```

- Extrapolation für symmetrische Differenzenquotienten ($q = 2$ in Satz 3.2.3)

```
tab=extrapolation_tab('bsp_symdiffqu_exp',2.^-[1:5],2)

tab =

   1.04219061098749                  0                  0                  0
   1.01044926723267   0.99986881931440                  0                  0
   1.00260620192892   0.99999184682767   1.00000004866189                  0
   1.00065116883507   0.99999949113712   1.00000000075775   0.99999999999737
   1.00016276836414   0.99999996820716   1.00000000001183   0.99999999999999
```

3.2.3 Satz: Richardson-Extrapolation

Die Funktion $a : \mathbb{R}_+ \to \mathbb{R}$ besitze die Entwicklung

$$a(h) = a_0 + \sum_{j=1}^{n+1} a_j h^{j \cdot q} + \mathbf{o}(h^{(n+1)q}), \qquad h \to 0.$$

Hierbei sind $q > 0$ und $a_j \in \mathbb{R}$, $j = 0, \dots, n+1$.
Weiter sei $(h_j)_{j \in \mathbb{N}_0}$ eine monoton fallende Folge positiver Zahlen mit

$$0 < \frac{h_{j+1}}{h_j} \le \rho < 1, \quad j \in \mathbb{N}_0.$$

Dann erfüllt das Interpolationspolynom

$$p_{j,j+n} \in \mathcal{P}_n \quad \text{zu den Punkten} \quad (h_j^q, a(h_j)), \dots, (h_{j+n}^q, a(h_{j+n}))$$

die Beziehung

$$a(0) - p_{j,j+n}(0) = \mathcal{O}(h_j^{(n+1)q}), \quad j \to \infty.$$

3.2.4 Lemma

Die Lagrange-Grundpolynome $L_{n,j}$ zu paarweise verschiedenen Stützstellen $x_0, \ldots, x_n \in \mathbb{R}$ erfüllen

$$\sum_{j=0}^{n} x_j^k L_{n,j}(x) = \begin{cases} x^k, & \text{für } 0 \leq k \leq n, \\ x^{n+1} - w(x), & \text{für } k = n+1. \end{cases}$$

Beweis: Für $x \in \mathbb{R}$ und $0 \leq k \leq n$ ist

$$\sum_{j=0}^{n} x_j^k L_{n,j}(x) = x^k,$$

weil das Monom $e_k : \mathbb{R} \to \mathbb{R}$, $e_k(x) = x^k$ "sich selbst interpoliert". Für $k = n+1$ ergibt die Fehlerdarstellung in der Newton-Form

$$x^{n+1} - \sum_{j=0}^{n} x_j^{n+1} L_{n,j}(x) = e_{n+1}[x_0, \ldots, x_n, x] \cdot w(x),$$

und die dividierte Differenz $(n+1)$-ter Ordnung von e_{n+1} ist 1.

3.2.5 Extrapolations-Tafel: Das Neville-Schema zur Berechnung der $a_{j+k,k} := p_{j,j+k}(0) \approx a(0)$ wird als **untere** Dreiecksmatrix aufgeschrieben:

$$
\begin{array}{c|c|cccccc}
h_0 & a_{00} = a(h_0) & & & & & & \\
h_1 & a_{10} = a(h_1) & a_{11} & & & & & \\
h_2 & a_{20} = a(h_2) & a_{21} & a_{22} & & & & \\
\vdots & \vdots & \vdots & \vdots & & & & \\
h_j & a_{j0} = a(h_j) & a_{j1} & a_{j2} & \dots & a_{j,j-1} & a_{jj} & \\
\vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \ddots
\end{array}
$$

mit Hilfe der Rekursion (mit dem entsprechenden q in Satz 3.2.3)

$$
j = 0, 1, 2 \dots : \qquad a_{j0} = a(h_j)
$$

$$
k = 1, \dots, j : \quad a_{jk} = a_{j,k-1} + \frac{1}{\left(h_{j-k}/h_j\right)^q - 1} \left(a_{j,k-1} - a_{j-1,k-1}\right).
$$

3.2.6 Bemerkung: Schrittweiten-Folgen und monotone Konvergenz

(i) Gebräuchliche Schrittweiten-Folgen $(h_j)_{j\geq 0}$ sind

$$\left(\frac{1}{2^j}\right)_{j\in\mathbb{N}_0},$$

$$\left(\frac{1}{n_j}\right)_{j\in\mathbb{N}_0}, \quad n_j = 1, 2, 3, 4, 6, 8, 12, \ldots, 2^{(j+1)/2}\ j \text{ ungerade}, 3 * 2^{(j-2)/2}\ j \geq 2 \text{ gerade}.$$

Unzulässig ist die Folge $\left(\frac{1}{j}\right)_{j\in\mathbb{N}}$, da $\lim\limits_{j\to\infty} \dfrac{j}{j+1} = \rho = 1$.

(ii) Nach Satz 3.2.3 gilt für die Einträge der k-ten Spalte

$$a_{jk} - a(0) = \mathcal{O}(h_{j-k}^{(k+1)q}), \quad j \to \infty,$$

falls die Schrittweiten-Folge $(h_j)_{j\geq 0}$ die Voraussetzungen des Satzes erfüllt. Noch genauer ist sogar für (unbekanntes!) $a_{k+1} \neq 0$

$$a_{jk} - a(0) = (-1)^k a_{k+1} \prod_{i=j-k}^{k} h_i^q + \mathbf{o}(h_{j-k}^{(k+1)q}), \quad j \to \infty,$$

woraus man auf "schließlich monotone" Konvergenz der Folge $(a_{jk})_{j\geq k}$ gegen $a(0)$ schließen kann.

(iii) Führt man zusätzlich die Folge

$$b_{j,k} = 2a_{j+1,k} - a_{j,k}, \qquad j \geq k,$$

mit, so ergibt sich wegen $|a_{j+1,k} - a(0)| \ll |a_{j,k} - a(0)|$ die Beziehung

$$b_{j,k} - a(0) \approx a(0) - a_{j,k},$$

also (heuristisch) eine Einschließung ($\rightarrow$ Abbruchkriterium!)

$$a_{j,k} \leq a(0) \leq b_{j,k} \qquad \text{oder} \qquad a_{j,k} \geq a(0) \geq b_{j,k}.$$

Besitzt die Funktion a sogar eine Reihenentwicklung

$$a(h) = a(0) + \sum_{j=1}^{\infty} a_j h^{qj}$$

(z.B. falls a analytisch ist), so kann auch der Grenzwert $\lim_{k\to\infty} a_{kk}$ entlang der Diagonalen der Extrapolations-Tafel betrachtet werden.

3.2.7 Satz: Konvergenz entlang der Diagonalen

Falls in der Reihenentwicklung unendlich viele $a_j \neq 0$ sind und falls

$$\inf_{j\in\mathbb{N}_0} \frac{h_{j+1}}{h_j} > 0 \quad \text{und} \quad \sup_{j\in\mathbb{N}_0} \frac{h_{j+1}}{h_j} < 1,$$

so konvergiert die Folge $(a_{kk})_{k\geq 0}$ der Diagonalelemente der Extrapolations-Tafel schneller gegen $a(0)$ als die Folge $(a_{j,k_0})_{j\geq k_0}$ entlang einer beliebigen Spalte k_0; d.h.

$$\lim_{k\to\infty} \frac{|a_{kk} - a(0)|}{|a_{k,k_0} - a(0)|} = 0.$$

3.3 Gauß-Approximation

Wir betrachten weiterhin die Approximation von Funktionen

$$f \in C[a,b] = \{f : [a,b] \to \mathbb{K} : f \text{ ist stetig}\}.$$

- $C[a,b]$ ist ein $\mathbb{K}$-Vektorraum, seine Dimension ist unendlich.
- $\mathcal{P}_n$ (genauer die Einschränkungen der Polynome vom Grad kleiner oder gleich n auf $[a,b]$) ist ein $(n+1)$-dimensionaler Teilraum von $C[a,b]$

Die Gaußapproximation ist die *Orthogonalprojektion* von $C[a,b]$ auf $\mathcal{P}_n$ bezüglich eines gegebenen Skalarprodukts.

3.3.1 Definition: Skalarprodukt

Es sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $s : V \times V \to \mathbb{K}$ heißt Skalarprodukt auf V, wenn

(S1) $s(\alpha x + \beta y, z) = \alpha s(x,z) + \beta s(y,z)$ für alle $x, y, z \in V$, $\alpha, \beta \in \mathbb{K}$;

(S2) $s(x,y) = \overline{s(y,x)}$ für alle $x, y \in V$;

(S3) $(x,x) > 0$ für alle $x \in V \setminus \{0\}$.

(V, s) heißt *Skalarproduktraum* oder *Prä-Hilbertraum*, speziell für $\mathbb{K} = \mathbb{R}$ auch *euklidischer Raum* und für $\mathbb{K} = \mathbb{C}$ *unitärer Raum*.

Schreibweise: $\langle x, y\rangle = s(x,y)$

Wichtige Ergänzung:

(i) Das Skalarprodukt induziert eine *Norm*

$$\|x\| = \sqrt{\langle x, x\rangle}, \quad x \in V.$$

(ii) Es gilt die *Cauchy-Schwarz-Ungleichung*

$$|\langle x, y\rangle| \leq \|x\| \, \|y\|, \quad x, y \in V.$$

Beweis: klar für $v = 0$ oder $w = 0$; für $v, w \neq 0$ setze o.B.d.A. $\|v\| = \|w\| = 1$ und betrachte

$$0 \leq \langle v - \langle v, w\rangle w, v - \langle v, w\rangle w\rangle = 1 - |\langle v, w\rangle|^2.$$

3.3.2 Beispiel: Skalarprodukte auf $C[a, b$

Es seien $a, b \in \mathbb{R}$, $a < b$.

a) Das "Standard-Skalarprodukt" auf $C[a, b]$ ist

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}\, dx.$$

(SP1) und (SP2) sind sofort klar, (SP3) folgt aus der Stetigkeit:

$$\langle f, f \rangle = \int_a^b |f(x)|^2\, dx \geq 0 \quad \text{für alle } f \in C[a, b];$$

für $f \neq 0$ existiert ein Intervall $U = [x_0 - \delta, x_0 + \delta] \cap [a, b]$ mit $|f(x)| > 0$ für alle $x \in U$, also ist $\int_a^b |f(x)|^2\, dx > 0$ für $f \neq 0$.

Die induzierte Norm ist die L_2-Norm

$$\|f\| = \left(\int_a^b |f(x)|^2\, dx \right)^{1/2}.$$

Durch

$$\|f - g\| = \left(\int_a^b |f(x) - g(x)|^2\, dx \right)^{1/2}$$

wird die Abweichung von f und g *im quadratischen Mittel* erfasst.

b) Das *gewichtete* Skalarprodukt auf $C[a, b]$: Die Funktion $w : (a, b) \to \mathbb{R}$ erfülle

$$w(x) > 0 \text{ für alle } x \in (a, b), \qquad \int_a^b w(x)\, dx < \infty.$$

w heißt *Gewichtsfunktion*. Dann ist

$$\langle f, g\rangle_w = \int_a^b f(x)\overline{g(x)}\; w(x)\, dx$$

ein Skalarprodukt auf $C[a, b]$ mit induzierter Norm

$$\|f\|_w = \left(\int_a^b |f(x)|^2\, w(x)\, dx\right)^{1/2}.$$

Beispiel: $w(x) = \frac{1}{\sqrt{1-x^2}}$ auf $[-1, 1]$ ergibt

$$\langle f, g\rangle_w = \int_{-1}^1 f(x)\overline{g(x)}\frac{dx}{\sqrt{1-x^2}}.$$

3.3.3 Bemerkung und Bezeichnungen: Es sei V ein Skalarproduktraum.

(i) Es gilt die *Parallelogramm-Identität*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2), \qquad x, y \in V.$$

Aus der Parallelogramm-Identität folgt umgekehrt die *Polarisierung* für $\mathbb{K} = \mathbb{R}$

$$\langle x, y \rangle = \frac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2),$$

und für $\mathbb{K} = \mathbb{C}$

$$\langle x, y \rangle = \frac{1}{2}(\|x+y\|^2 + i\|x+iy\|^2 - (1+i)(\|x\|^2 + \|y\|^2)).$$

(ii) Der Cosinus des *Winkels* zwischen $x, y \in V$ mit $x, y \neq 0$ ist

$$\cos \angle(x, y) = \frac{\langle x, y \rangle}{\|x\| \, \|y\|}.$$

(iii) $x, y \in V$ sind *orthogonal*, wenn $\langle x, y \rangle = 0$ gilt.

3.3.4 Geometrisches Verständnis der Gauß-Approximation:

Die Gaußapproximation von $f \in C[a,b]$ durch Polynome vom Grad kleiner oder gleich n berechnet dasjenige $p \in \mathcal{P}_n$ mit

$$\|f - p\| = \min_{q \in \mathcal{P}_n} \|f - q\|.$$

Dieser kürzeste Abstand wird genau dann erzielt, wenn die *Orthogonalitätsbedingung*

$$\langle f - p, q \rangle = 0 \quad \text{für alle} \quad q \in \mathcal{P}_n$$

erfüllt ist; d.h. die Differenz $f - p$ ist orthogonal zu jedem $q \in \mathcal{P}_n$. Siehe hierzu Satz 3.3.5.

Anschaulich: Die Gaußapproximation von $f \in C[a,b]$ ist die Orthogonalprojektion von f auf den Teilraum der Polynome.

3.3.5 Satz: Die Orthogonalitätsbedingung

Es sei V ein Skalarproduktraum und $S \preceq V$ ein endlichdimensionaler Teilraum. Dann sind äquivalent:

(i) $p \in S$ ist eine beste Approximation von $f \in V$; d.h.

$$\|f - p\| = \min_{q \in S} \|f - q\|.$$

(ii) Es gilt die *Orthogonalitätsbedingung*

$$\langle f - p, q \rangle = 0 \quad \text{für alle} \quad q \in S.$$

Das Element $p \in S$ ist durch (i) oder (ii) eindeutig bestimmt. Es heißt *Orthogonalprojektion von f auf S*.

Bemerkung: In der Approximationstheorie und der Funktionalanalysis wird gezeigt, dass die Äquivalenz sogar für jeden **abgeschlossenen** Teilraum (auch mit $\dim S = \infty$) eines Hilbertraumes gilt.

Die Gram-Matrix dient der allgemeinen Beschreibung der Orthogonalprojektion.

3.3.6 Definition und Satz: Gram-Matrix

Es seien V ein Skalarproduktraum, $\psi_1, \dots, \psi_n \in V$. Die Matrix

$$M = \begin{pmatrix} \langle \psi_1, \psi_1 \rangle & \cdots & \langle \psi_n, \psi_1 \rangle \\ \vdots & & \vdots \\ \langle \psi_1, \psi_n \rangle & \cdots & \langle \psi_n, \psi_n \rangle \end{pmatrix}$$

heißt *Gram-Matrix* der Elemente $\psi_1, \dots, \psi_n$. Es gilt:

a) M ist hermitesch und positiv-semidefinit.

b) M ist genau dann positiv-definit, wenn die Familie $(\psi_1, \dots, \psi_n)$ linear unabhängig ist.

c) Sind $\psi_1, \dots, \psi_n$ linear unabhängig und $S = \mathrm{Span}(\psi_1, \dots, \psi_n)$, so ist die Orthogonalprojektion von $f \in V$ auf S gegeben durch

$$p = \sum_{j=1}^{n} c_j \psi_j,$$

wobei der Vektor $\vec{c} = (c_j)_{j=1,\dots,n}$ die eindeutige Lösung des linearen Gleichungssystems

$$M\vec{c} = \begin{pmatrix} \langle f, \psi_1 \rangle \\ \vdots \\ \langle f, \psi_n \rangle \end{pmatrix} \qquad \text{ist.}$$

3.3.7 Bemerkung: Es seien $\psi_1, \dots, \psi_n \in V$ und M die zugehörige Gram-Matrix.

(i) M ist genau dann eine Diagonalmatrix, wenn die Elemente $\psi_1, \dots, \psi_n$ paarweise orthogonal sind, und $M = \mathrm{I}$ genau dann, wenn die Elemente $\psi_1, \dots, \psi_n$ ein Orthonormalsystem in V bilden.

(ii) Gedächtnisstütze: Die Transponierte der Gram-Matrix kann man kurz schreiben als

$$M^T = \overline{M} = \begin{pmatrix} \langle \psi_1, \\ \vdots \\ \langle \psi_n, \end{pmatrix} (\psi_1\rangle, \dots, \psi_n\rangle)\,.$$

Dabei wird die $n \times n$-Matrix der Einträge $\langle \psi_j, \psi_k \rangle$ gebildet. Diese Vektornotation hilft z.B. beim Basiswechsel.

3.3.8 Korollar:

a) Die Orthogonalprojektion $\Pi_S : V \to S$ ist eine *lineare Abbildung*.

b) Falls $(\phi_1, \ldots, \phi_n)$ eine Orthonormalbasis von S ist, so ist die Orthogonalprojektion gegeben durch

$$\Pi_S(f) = \sum_{j=1}^{n} \langle f, \phi_j \rangle \phi_j, \qquad f \in V.$$

3.3.9 Beispiele:

a) Auf dem $\mathbb{R}$-Vektorraum $C[0,1]$ ist das Standardskalarprodukt

$$\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$$

definiert. Zur Monom-Basis $e_k : [0,1] \to \mathbb{R}$, $e_k(x) = x^k$, $k = 0, \dots, n$, von $\mathcal{P}_n$ gehört die Gram-Matrix

$$H_{n+1} = \left(\frac{1}{j+k-1} \right)_{j,k=1,\dots,n+1},$$

dies ist die Hilbert-Matrix von Übungsblatt 3. Sie ist sehr schlecht konditioniert. Für die Gauß-Approximation sollte man also eine andere Methode als in Satz 3.3.6 verwenden!!

($\to$ Koordinaten-Transformation der Legendre-Polynome auf das Intervall $[0,1]$.)

b) Auf dem $\mathbb{R}$-Vektorraum $C[-1,1]$ ist das gewichtete Skalarprodukt

$$\langle f, g\rangle_w = \int_{-1}^{1} f(x)g(x)\,\frac{dx}{\sqrt{1-x^2}}$$

definiert. Die *Tschebyscheff-Polynome 1. Art* $T_n \in \mathcal{P}_n$ lauten für $x \in [-1,1]$

$$T_n(x) = \cos(n \arccos(x)), \qquad n =\in \mathbb{N}_0.$$

Dies sind tatsächlich Polynome vom Grad n: klar ist

$$T_0(x) = 1, \quad T_1(x) = x \qquad \text{für } \; x \in [-1,1],$$

und aus der trigonometrischen Identität $\cos((n+1)t) + \cos((n-1)t) = 2\cos(nt)\cos t$ folgt die *Rekursion der Tschebyscheff-Polynome*

$$\boxed{T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \qquad n \geq 1,}$$

also für $n = 2, 3, 4, 5$

$$T_2(x) = 2x^2-1, \quad T_3(x) = 4x^3-3x, \quad T_4(x) = 8x^4-8x^2+1, \quad T_5(x) = 16x^5-20x^3+5x.$$

(Natürlich sind die Polynome auf ganz $\mathbb{R}$ (sogar $\mathbb{C}$) mittels der Rekursion definiert.)

Für $j, k \geq 0$ ergibt sich das Skalarprodukt

$$\langle T_j, T_k\rangle_w = \int_{-1}^{1} T_j(x)T_k(x)\,\frac{dx}{\sqrt{1-x^2}} = \int_0^{\pi} \cos(jt)\cos(kt)\,dt = \begin{cases} \pi, & \text{für } j = k = 0, \\ \pi/2, & \text{für } j = k > 0, \\ 0, & \text{für } j \neq k. \end{cases}$$

Die zugehörige Gram-Matrix ist eine Diagonalmatrix. Die Gauß-(Tschebyscheff)-Approximation vom Grad n der Funktion $f \in C[-1,1]$ ist gegeben durch

$$p = \sum_{k=0}^{n} c_k T_k$$

mit den Koeffizienten

$$c_0 = \frac{1}{\pi} \int_{-1}^{1} f(x) \frac{dx}{\sqrt{1-x^2}}, \qquad c_k = \frac{2}{\pi} \int_{-1}^{1} f(x) T_k(x) \frac{dx}{\sqrt{1-x^2}} \quad \text{für} \quad k > 0.$$

Vorteilhaft für die Gauß-Approximation ist die Verwendung von Orthogonalsystemen. Aus der Linearen Algebra ist bekannt (evtl. für Vektoren im $\mathbb{K}^n$):

3.3.10 Verfahren: Gram-Schmidt-Orthogonalisierung

Es sei V ein Skalarproduktraum und $\psi_1, \ldots, \psi_n \in V \setminus \{0\}$. Weiter sei

$$S = \mathrm{Span}(\psi_1, \ldots, \psi_n), \qquad 1 \leq r = \dim S \leq n.$$

Der folgende Algorithmus liefert eine Orthonormalbasis $\phi_1, \ldots, \phi_r$ von S (mit Aussortieren linear abhängiger ψ_j):

1. Setze $k = 1$ und $\phi_1 = \frac{1}{\|\psi_1\|}\,\psi_1$.
2. Für $j = 2, \ldots, n$

$$\tau = \psi_j - \sum_{\ell=1}^{k} \langle \psi_j, \phi_\ell \rangle \phi_\ell.$$

Falls $\tau \neq 0$ setze $k = k+1$ und $\phi_k = \frac{1}{\|\tau\|}\,\tau$.

3.3.11 Beispiel: Legendre-Polynome auf [−1, 1

$C[-1,1]$ als $\mathbb{R}$-Vektorraum mit dem Standardskalarprodukt

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx$$

besitzt die Monom-Basis $e_k : [-1,1] \to \mathbb{R}$, $e_k(x) = x^k$, für $k = 0, 1, \dots, n$.

Gram-Schmidt-Orthonormalisierung ergibt (verwende L_j für die τ im Algorithmus 3.3.10):

$$L_0(x) = 1, \qquad \phi_0(x) = \sqrt{1/2},$$
$$L_1(x) = x, \qquad \phi_1(x) = \sqrt{3/2}\, x,$$

und mit dem folgenden Satz 3.3.12

$$\boxed{L_{k+1}(x) = x \cdot L_k(x) - \frac{k^2}{4k^2-1} \cdot L_{k-1}(x), \qquad k = 1, 2, \dots}$$

$$\phi_k(x) = \frac{(2k)!}{(k!)^2}\sqrt{\frac{2k+1}{2^{2k+1}}}\; L_k(x).$$

Für $n = 2, 3, 4, 5$ ist

$$L_2(x) = x^2 - \frac{1}{3}, \quad L_3(x) = x^3 - \frac{3}{5}x, \quad L_4(x) = x^4 - \frac{6}{7}x^2 + \frac{3}{35}, \quad L_5(x) = x^5 - \frac{10}{9}x^3 + \frac{5}{21}x.$$

Damit ist die Gauß-(Legendre)-Approximation von f durch Polynome vom Grad n gegeben durch

$$p = \sum_{k=0}^{n} c_k \phi_k$$

mit den Koeffizienten

$$c_k = \int_{-1}^{1} f(x)\phi_k(x)\,dx \quad \text{für} \quad k = 0, \ldots, n.$$

Die Normalisierungskonstante der ϕ_k berechnet man z.B. mit der sog. *Rodriguez-Formel* für die Legendre-Polynome:

$$L_n(x) = \frac{n!}{(2n)!} \, \frac{d^n}{dx^n} \left[(x^2-1)^n \right].$$

Beachte: Höchstkoeffizient ist

$$\frac{L_n^{(n)}(0)}{n!} = \frac{1}{(2n)!} \, \underbrace{\frac{d^{2n}}{dx^{2n}} \left[(x^2-1)^n \right]}_{=(2n)!} = 1.$$

Mit partieller Integration (alle Randterme sind Null) ergibt sich

$$\begin{aligned}
\langle L_n, L_n \rangle &= \frac{(n!)^2}{((2n)!)^2} \int_{-1}^{1} \frac{d^n}{dx^n} \left[(x^2-1)^n \right] \, \frac{d^n}{dx^n} \left[(x^2-1)^n \right] \, dx \\
&= (-1)^n \frac{(n!)^2}{((2n)!)^2} \int_{-1}^{1} (x^2-1)^n \, \underbrace{\frac{d^{2n}}{dx^{2n}} \left[(x^2-1)^n \right]}_{=(2n)!} \, dx \\
&= (-1)^n \frac{(n!)^2}{(2n)!} \int_{-1}^{1} (x-1)^n \, (x+1)^n \, dx \\
&= \frac{(n!)^2}{(2n)!} \, \frac{(n!)^2}{(2n)!} \int_{-1}^{1} (x+1)^{2n} \, dx \\
&= \frac{(n!)^4}{((2n)!)^2} \, \frac{2^{2n+1}}{2n+1},
\end{aligned}$$

und daraus die obige Normierung der ϕ_k.

Bemerkung:
Eine andere Normalisierung, nämlich $\widetilde{L}_n(1) = 1$, erzielt man mit der Rekursion

$$\boxed{\begin{aligned} &\widetilde{L}_0(x) = 1, \qquad \widetilde{L}_1(x) = x, \\ &\widetilde{L}_{k+1}(x) = \tfrac{2k+1}{k+1}\widetilde{L}_k(x) - \tfrac{k}{k+1}\widetilde{L}_{k-1}(x). \end{aligned}}$$

Hierbei ist

$$\langle L_k, L_k \rangle = \frac{2}{2k+1},$$

also ist ϕ_k in 3.3.11

$$\phi_k(x) = \sqrt{\frac{2k+1}{2}}\, \widetilde{L}_k(x).$$

Für $n = 2, 3, 4, 5$ ist

$$\widetilde{L}_2(x) = \frac{3}{2}x^2 - \frac{1}{2}, \quad L_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x, \quad L_4(x) = \frac{35}{8}x^4 - \frac{15}{4}x^2 + \frac{3}{8}, \quad L_5(x) = \frac{63}{8}x^5 - \frac{35}{4}x^3 + \frac{15}{8}x.$$

3.3.12 Satz: 3-Term Rekursion der Orthogonalpolynome

Das Skalarprodukt $\langle \cdot, \cdot \rangle$ auf $C[-1,1]$ besitze die Symmetrie-Eigenschaft

$$\langle p, xq \rangle = \langle xp, q \rangle \quad \text{für alle Polynome} \quad p, q.$$

Dann führt die Gram-Schmidt-Orthonormalisierung der Monom-Basis $\{1, x, \ldots, x^n\}$ auf die folgenden Polynome $\widetilde{p}_k$ (mit Höchstkoeffizient 1) und ϕ_k (mit $\|\phi_k\| = \sqrt{\langle \phi_k, \phi_k \rangle} = 1$):

$$\begin{aligned}
\widetilde{p}_0(x) &= 1, \quad \widetilde{p}_1(x) = x - \beta_0, \\
\widetilde{p}_{k+1}(x) &= (x - \beta_k)\widetilde{p}_k(x) - \gamma_k \widetilde{p}_{k-1}(x), && k = 1, 2, \ldots, \\
\phi_k &= \tfrac{1}{\|\widetilde{p}_k\|}\widetilde{p}_k, && k = 0, 1, 2, \ldots
\end{aligned}$$

mit

$$\beta_k = \frac{\langle x\widetilde{p}_k \widetilde{p}_k \rangle}{\|\widetilde{p}_k\|^2} \;\text{ für } k \geq 0, \qquad \gamma_k = \frac{\|\widetilde{p}_k\|^2}{\|\widetilde{p}_{k-1}\|^2} \;\text{ für } k \geq 1$$

Achtung: $\|p\|^2 = \langle p, p \rangle$ mit dem **gegebenen** Skalarprodukt!

Beweis: $\widetilde{p}_0 = 1$ und $\widetilde{p}_1(x) = x - \langle x, \phi_0 \rangle \phi_0 = x - \beta_0$ sind anhand der Definitionen abzulesen.

Für $k \geq 1$ setze

$$q_{k+1}(x) = (x - \beta_k)\widetilde{p}_k(x) - \gamma_k \widetilde{p}_{k-1}(x).$$

Dann ergibt die Orthogonalität $\widetilde{p}_k \perp \mathcal{P}_{k-1}$

$$\begin{aligned}
\langle q_{k+1}, \widetilde{p}_k \rangle &= \langle x\widetilde{p}_k, \widetilde{p}_k \rangle - \beta_k \|\widetilde{p}_k\|^2 - \gamma_k \underbrace{\langle \widetilde{p}_k, \widetilde{p}_{k-1} \rangle}_{=0} = 0, \\
\langle q_{k+1}, \widetilde{p}_{k-1} \rangle &= \langle x\widetilde{p}_k, \widetilde{p}_{k-1} \rangle - \beta_k \underbrace{\langle \widetilde{p}_k, \widetilde{p}_{k-1} \rangle}_{=0} - \underbrace{\gamma_k \|\widetilde{p}_{k-1}\|^2}_{=\langle \widetilde{p}_k, \widetilde{p}_k \rangle} \\
&= \langle \widetilde{p}_k, \underbrace{x\widetilde{p}_{k-1} - \widetilde{p}_k}_{\in \mathcal{P}_{k-1}} \rangle = 0.
\end{aligned}$$

Weiterhin ergibt sich für $j < k - 1$ sofort $\langle q_{k+1}, \widetilde{p}_j \rangle = 0$.

Wir haben gezeigt, dass q_{k+1} ein Polynom vom Grad $k + 1$ mit dem Höchstkoeffizienten 1 ist, das orthogonal zu $\mathcal{P}_k$ ist. Weil das orthogonale Komplement von $\mathcal{P}_k$ in $\mathcal{P}_{k+1}$ eindimensional ist, folgt also $q_{k+1} = \widetilde{p}_{k+1}$.

3.3.13 Bemerkung:

Bei der Gauß-Approximation bzgl. des Standard-Skalarprodukts wird die Abweichung im quadratischen Mittel minimiert. Dabei wird die Maximalabweichung

$$\|f - p\| = \max_{x \in [a,b]} |f(x) - p(x)|$$

häufig insbesondere in der Nähe der Intervallenden groß. Deshalb verwendet man bei der Berechnung der besten Gauß-Approximation p gerne das gewichtete Skalarprodukt

$$(f,g) = \int_a^b f(x)\overline{g(x)} \frac{dx}{\sqrt{(x-a)(b-x)}},$$

das den Fehler $f - p$ am Rand höher gewichtet als in der Mitte des Intervalls $[a, b]$.

Die Orthogonalpolynome zu diesem Skalarprodukt sind die auf das Intervall $[a, b]$ transformierten Tschebyscheff-Polynome 1. Art, siehe Beispiel 3.3.9(b):

$$T_{n,[a,b]}(x) = T_n\left(\frac{2x-a-b}{b-a}\right)$$

mit der Normalisierungskonstanten

$$\|T_{n,[a,b]}\|^2 = \int_a^b T_n\left(\frac{2x-a-b}{b-a}\right)^2 \frac{dx}{\sqrt{(x-a)(b-x)}} = \begin{cases} \frac{\pi(b-a)}{2}, & \text{für } n = 0, \\ \frac{\pi(b-a)}{4}, & \text{für } n > 0. \end{cases}$$

Die Gauß-Legendre Approximation (links) zum Standard-Skalarprodukt, und die Gauß-Tschebyscheff Approximation (rechts) zur Gewichtsfunktion $w(x) = 1/\sqrt{1-x^2}$: