1 Euklidische Approximation

Sei V ein reeller euklidischer Vektorraum.
Das Skalarprodukt in V wird mit $\langle \cdot, \cdot \rangle_V$ und die Norm mit $\| \cdot \|_V$ bezeichnet.
$V_N \subset V$ sei ein Teilraum der Dimension $N < \infty$ mit Basis $\{\phi_n\}_{n=1,\dots,N}$.

(1.1) Problemstellung: Sei $v \in V$. Bestimme $v_N \in V$ mit

$$\|v - v_N\|_V = \min_{w_N \in V_N} \|v - w_N\|_V .$$

(1.2) Die Matrix

$$A = \left(\langle \phi_n, \phi_k \rangle_V \right)_{n,k=1,\dots,N} \in \mathbb{R}^{N \times N}$$

ist symmetrisch und positiv definit.

(1.3) Problem (1.1) ist eindeutig lösbar. Es gilt

$$v_N = \sum_{n=1}^{N} x_n \phi_n ,$$

wobei $x \in \mathbb{R}^N$ die eindeutige Lösung des linearen Gleichungssystems

$$Ax = b$$

mit $b = \left(\langle v, \phi_k \rangle_V \right)_{k=1,\dots,N} \in \mathbb{R}^N$ ist.

2 Direkte Lösungsverfahren für lineare Gleichungen

(2.1) Sei $L \in \mathbb{R}^{N\times N}$ eine normierte untere Dreiecksmatrix und $b \in \mathbb{R}^N$, d.h. $\operatorname{diag} L = I_N$ und $L[1:n, n+1] = 0_n$ für $n = 1, \ldots, N-1$. Dann ist L regulär und das lineare Gleichungssystem $Ly = b$ ist mit $O(N^2)$ Operationen lösbar.
Entsprechend ist für eine reguläre obere Dreiecksmatrix $R \in \mathbb{R}^{N\times N}$ (d.h. $R[n,n] \neq 0$ für alle n und $R[n+1, 1:n] = 0_n^T$ für $n < N$) das LGS $Rx = y$ in $O(N^2)$ Operationen lösbar.

(2.2) Die normierten unteren Dreiecksmatrizen bilden eine Gruppe.
Die regulären oberen Dreiecksmatrizen bilden eine Gruppe.

(2.4) Wenn eine Matrix $A \in \mathbb{R}^{N\times N}$ eine LR-Zerlegung $A = LR$ mit einer normierten untere Dreiecksmatrix L und einer regulären obere Dreiecksmatrix R besitzt, dann ist A regulär und das LGS $Ax = b$ ist mit $O(N^2)$ Operationen lösbar.

(2.5) Eine Matrix $A \in \mathbb{R}^{N\times N}$ besitzt genau dann eine LR-Zerlegung von A, wenn alle Hauptuntermatrizen $A[1:n, 1:n]$ regulär sind.
Die LR-Zerlegung ist eindeutig und lässt sich mit $O(N^3)$ Operationen berechnen.

(2.6) Eine Matrix $A \in \mathbb{R}^{N\times N}$ heißt *strikt diagonal-dominant*, falls $|A[n,n]| > \sum_{k=1,\ k\neq n}^{N} |A[n,k]| \qquad \forall n.$

(2.7) Wenn A strikt diagonal dominant ist, dann existiert eine LR-Zerlegung.

(2.8) Sei $A \in \mathbb{R}^{N\times N}$ symmetrisch und positiv definit. Dann existiert genau eine Cholesky-Zerlegung $A = LL^T$ mit einer regulären unteren Dreiecksmatrix L.

LR- und Cholesky-Zerlegung

```
function x = lr_solve(A,b)
  N = size(A,1);
  for n=1:N-1
    A(n+1:N,n) = A(n+1:N,n) / A(n,n);
    A(n+1:N,n+1:N) = A(n+1:N,n+1:N) - A(n+1:N,n) * A(n,n+1:N);
  end
  x = b;
  for n=2:N
    x(n) = x(n) - A(n,1:n-1) * x(1:n-1);
  end
  for n=N:-1:1
    x(n) = (x(n) - A(n,n+1:N) * x(n+1:N)) / A(n,n);
  end
return
```

```
function x = cholesky_solve(A,b)
  N = size(A,1);
  for n=1:N
    A(n:N,n) = A(n:N,n) - A(n:N,1:n-1) * A(n,1:n-1)';
    A(n:N,n) = A(n:N,n) / sqrt(A(n,n));
  end
  x = b
  for n=1:N
    x(n) = (x(n) - A(n,1:n-1) * x(1:n-1)) / A(n,n);
  end
  for n=N:-1:1
    x(n) = (x(n) - A(n+1:N,n)' * x(n+1:N)) / A(n,n);
  end
return
```

2 Direkte Lösungsverfahren für lineare Gleichungen

(2.9) Sei $\pi \in S_N$ eine Permutation. Dann heißt $P_\pi = \left(e_{\pi^{-1}(1)} | \ldots | e_{\pi^{-1}(N)}\right) \in \mathbb{R}^{N\times N}$ Permutationsmatrix zu π. Es gilt $(P_\pi A)[n,k] = A[\pi(n),k]$ und $(AP_\pi)[n,k] = A[n,\pi^{-1}(k)]$.

(2.10) Die Permutationsmatrizen in $\mathbb{R}^{N\times N}$ bilden eine Gruppe. Es gilt $P_\sigma P_\pi = P_{\pi\circ\sigma}$ und $P_{\pi^{-1}} = P_\pi^T$.

(2.11) Sei $A \in \mathbb{R}^{N\times N}$ regulär. Dann existiert eine Permutationsmatrix P, so dass PA eine LR-Zerlegung $PA = LR$ besitzt und für die Einträge $|L[m,n]| \le 1$ gilt.
Sie lässt sich mit $O(N^3)$ Operationen berechnen.
Die Lösung von $Ax = b$ berechnet sich aus $Ly = Pb$ und $Rx = y$.

Sei $|\cdot|$ eine Vektornorm, und sei $\|\cdot\|$ eine zugeordete Matrixnorm, d. h.,

$$|Ax| \le \|A\|\,|x|, \qquad x \in \mathbb{R}^N,\ A \in \mathbb{R}^{M\times N}.$$

(2.12) Sei $A \in \mathbb{R}^{N\times N}$ regulär, und sei $\triangle A \in \mathbb{R}^{N\times N}$ so klein, dass $\|\triangle A\| < \|A^{-1}\|^{-1}$ gilt. Dann ist die Matrix $\tilde{A} = A + \triangle A$ regulär.
Sei $b \in \mathbb{R}^N$, $b \neq 0_N$, $\triangle b \in \mathbb{R}^N$ klein und $\tilde{b} = b + \triangle b$.
Dann gilt für die Lösungen $x \in \mathbb{R}^N$ von $Ax = b$ und $\tilde{x} \in \mathbb{R}^N$ von $\tilde{A}\tilde{x} = \tilde{b}$

$$\frac{|\triangle x|}{|x|} \le \frac{\kappa(A)}{1 - \kappa(A)\frac{\|\triangle A\|}{\|A\|}} \left(\frac{|\triangle b|}{|b|} + \frac{\|\triangle A\|}{\|A\|}\right).$$

Dabei ist $\triangle x = \tilde{x} - x$, $\frac{|\triangle x|}{|x|}$ der *relative Fehler*, und $\kappa(A) = \|A\|\,\|A^{-1}\|$ die *Kondition* von A.

LR-Zerlegung mit Pivot-Suche

```
function x = lr_pivot_solve(A,b)
  N = size(A,1);
  p = (1:N)';
  for n = 1:N-1
    [r,m] = max(abs(A(n:N,n)));
    m = m+n-1;
    if abs(A(m,n))<eps
      error('*** ERROR *** Matrix fast singulär');
    end
    if (m ~= n)
      A([n m],:) = A([m n],:);   p([n m]) = p([m n]);
    end
    A(n+1:N,n) = A(n+1:N,n) / A(n,n);
    A(n+1:N,n+1:N) = A(n+1:N,n+1:N) - A(n+1:N,n) * A(n,n+1:N);
  end
  x = b(p);
  for n=2:N
    x(n) = x(n) - A(n,1:n-1) * x(1:n-1);
  end
  for n=N:-1:1
    x(n) = (x(n) - A(n,n+1:N) * x(n+1:N)) / A(n,n);
  end
return
```

(2.13) a) Eine Gleitkommazahlen zur Basis $B \in \{2,3,...\}$ der Mantissenlänge M und Exponentenlänge E ist die Menge

$$\mathrm{FL} = \left\{ \pm B^e \sum_{m=1}^{M} a_m B^{-m} : e = e^- + \sum_{k=0}^{E-1} c_k B^k,\ a_m, c_k \in \{0,1,...,B-1\} \right\}$$

b) Eine Gleitkommaarithmetik wird durch eine Abbildung $\mathrm{fl} : \mathbb{R} \longrightarrow \mathrm{FL}$ mit $\mathrm{fl}(x) = x$ für $x \in \mathrm{FL}$ definiert. Sei bestimmt die Rundung: $x \oplus y = \mathrm{fl}(x+y)$, $x \odot y = \mathrm{fl}(x \cdot y)$, etc.
Die zugehörige Maschinengenauigkeit ist $\mathrm{eps} = \sup \left\{ \frac{|x - \mathrm{fl}(x)|}{|x|} : x \notin \mathrm{FL} \right\}$.

(2.14) Sei $f : \mathbb{R}^N \longrightarrow \mathbb{R}^K$ eine differenzierbare Funktion und $x \in \mathbb{R}^N$. Dann heißt

a) $\kappa_{\mathrm{abs}}^{kn} = \left| \frac{\partial}{\partial x_n} f_k(x) \right|$ *absolute Konditionszahl*.

b) $\kappa_{\mathrm{rel}}^{kn} = \left| \frac{\partial}{\partial x_n} f_k(x) \right| \frac{|x_n|}{|f_k(x)|}$ *relative Konditionszahl*.

2 Kondition und Stabilität

(2.15)

a) Ein Problem heißt *sachgemäß gestellt*, wenn es eindeutig lösbar ist und die Lösung stetig von den Daten abhängt.

b) Die *Kondition* eines Problems ist eine Maß dafür, wie stark die Abhängigkeit der Lösung von Störungen in den Daten ist.

c) Die *Stabilität* eines numerischen Algorithmus ist eine Maß dafür, wie stark die Daten-Abhängigkeit der numerischen Lösung im Vergeich zu der exakten Lösung ist.

(2.16) Wir verwenden für $x \in \mathbb{R}^N$ und $A \in \mathbb{R}^{M\times N}$

$$|x|_1 = \sum_{n=1}^{N} |x[n]|\,, \qquad |x|_2 = \sqrt{x^T x}\,, \qquad |x|_\infty = \max_{n=1,\dots,N} |x[n]|$$

und die zugeordnete Operatornorm $\|A\|_p = \sup_{x\neq 0_N} \frac{|Ax|_p}{|x|_p}$, d.h.

$$\|A\|_1 = \max_{n=1,\dots,N} \sum_{m=1}^{M} |A[m,n]|\,, \quad \|A\|_2 = \sqrt{\rho(A^T A)}\,, \quad \|A\|_\infty = \max_{m=1,\dots,M} \sum_{n=1}^{N} |A[m,n]|$$

mit *Spekralradius* $\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$ und Spektrum $\sigma(A)$.

3 Ausgleichsrechnung

Sei $Q \in \mathbb{R}^{N\times N}$ orthogonal, d.h. $Q^T Q = I_N$. Dann ist $\kappa_2(Q) = 1$.

(3.1) Zu $v \in \mathbb{R}^N$ und $k \neq n$ mit $v[k]^2 + v[n]^2 > 0$ existiert eine *Givens-Rotation* $G \in \mathbb{R}^{N\times N}$ mit

$$\begin{pmatrix} G[k,k] & G[k,n] \\ G[n,k] & G[n,n] \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}, \qquad c^2 + s^2 = 1,$$

und $G[j][j] = 1$ für $j \neq k, n$ und $G[i][j] = 0$ sonst, so dass $e_n^T G v = 0$ gilt:
Für $|v[n]| > |v[k]|$ setze $\tau = \frac{v[k]}{v[n]}$, $s = \frac{1}{\sqrt{1+\tau^2}}$, $c = s\tau$, sonst setze $\tau = \frac{v[n]}{v[k]}$, $c = \frac{1}{\sqrt{1+\tau^2}}$, $s = c\tau$.
Es gilt $G = c(e_k e_k^T + e_n e_n^T) + s(e_k e_n^T - e_n e_k^T) + \sum_{j\neq k,n} e_j e_j^T$.

(2.10) Zu $v \in \mathbb{R}^N$, $v \neq 0_N$, existiert eine *Householder-Spiegelung* $H = I_N - \frac{2}{w^T w} w w^T \in \mathbb{R}^{N\times N}$ mit $w \in \mathbb{R}^N$, $w[1] = 1$, so dass $Hv = \sigma e_1$ mit $\sigma \in \mathbb{R}$ und $Hw = -w$ gilt:
Falls $v[1] > 0$, setze $\sigma = -|v|_2$, sonst setze $\sigma = |v|_2$. Dann definierte $w = \frac{1}{v[1]-\sigma}(v - \sigma e_1)$.

(3.3) Zu $A \in \mathbb{R}^{K\times N}$ existiert eine QR-Zerlegung $A = QR$ mit einer orthogonalen Matrix $Q \in \mathbb{R}^{K\times K}$ und eine oberen Dreiecksmatrix $R \in \mathbb{R}^{K\times N}$, d.h. $QQ^T = I_K$ und $R[k+1:K, k] = 0_{K-k}$ für $k = 1, ..., K$.

(3.4) Sei $A \in \mathbb{R}^{K\times N}$ und $b \in \mathbb{R}^K$. Dann gilt:

$$x \in \mathbb{R}^N \text{ minimiert } |Ax - b|_2 \qquad \Longleftrightarrow \qquad A^T A x = A^T b.$$

Berechnung der Householder-Vektoren

```
function [v,beta]=householder(y)
  N = length(y);
  s = y(2:N)' * y(2:N);
  if N == 1
    s = 0;
  end;
  v = [1;y(2:N)];
  if s == 0
    beta = 0;
  else
    mu = sqrt(y(1)^2 + s);
    if y(1) <= 0
      v(1) = y(1) - mu;
    else
      v(1) = -s/(y(1) + mu);
    end
    beta = 2*v(1)^2/(s + v(1)^2);
    v = v / v(1);
  end
return
```

```
function x = qr_solve(A,b)
  [M,N] = size(A);
  for m = 1:min(N,M-1)
    [v,beta] = householder(A(m:M,m));
    if beta ~= 0
      w = beta * v' * A(m:M,m:N);
      A(m:M,m:N) = A(m:M,m:N) - v * w;
      A(m+1:M,m) = v(2:M-m+1);
    end
  end
  for m = 1:min(N,M-1)
    v = [1;A(m+1:M,m)];
    beta = 2 / (v' * v);
    if beta ~= 2
       b(m:M) = b(m:M)-beta*(v'*b(m:M))*v;
    end
  end
  for n=min(N,M):-1:1
    x(n) = (b(n)-A(n,n+1:N)*x(n+1:N))/A(n,n);
  end
return
```

3 Ausgleichsrechnung

(3.5) Zu $A \in \mathbb{R}^{K\times N}$ mit $R = \operatorname{rang}(A)$ existiert eine Singulärwertzerlegung

$$A = V\Sigma U^T$$

mit Matizen $V = (v_1|\cdots|v_R) \in \mathbb{R}^{K\times R}$, $U = (u_1,|\cdots|u_R) \in \mathbb{R}^{N\times R}$, $\Sigma = \operatorname{diag}(\sigma_1,...,\sigma_R) \in \mathbb{R}^{R\times R}$
mit $V^TV = U^TU = I_R$ und den Singulärwerten $\sigma_1,...,\sigma_R > 0$.
Es gilt $A = \sum_{r=1}^{R} \sigma_r u_r v_r^T$ und $Ax = \sum_{r=1}^{R} \sigma_r (v_r^T x) u_r$.

(2.22) $A^+ = U\Sigma^{-1}V^T$ ist die *Pseudo-Inverse*. Es gilt $A^+ = \sum_{r=1}^{R} \sigma_r^{-1} v_r u_r^T$ und $Ax = \sum_{r=1}^{R} \sigma_r^{-1} (u_r^T x) v_r$.

(2.22) $x = A^+b$ löst die Normalengleichung $A^TAx = A^Tb$.

(2.23) Sei $A \in \mathbb{R}^{K\times N}$ und $b \in \mathbb{R}^M$. Dann gilt für die Tikhonov-Regularisierung mit $\alpha > 0$:

$$x \in \mathbb{R}^N \text{ minimiert } |Ax-b|_2^2 + \alpha|x|_2^2 \quad \Longleftrightarrow \quad (A^TA+\alpha I_N)x = A^Tb.$$

(2.24) Es gilt $\lim_{\alpha\longrightarrow 0}(A^TA+\alpha I_N)^{-1}A^Tb = A^+b$.

Diskrepanzprinzip: Sei $b \in \text{Bild}(A)$, $x = A^+b$ und b^δ eine Störung mit $|b-b^\delta| < \delta < |b|_2$.
Dann existiert ein $\alpha = \alpha(\delta) > 0$ mit $|Ax^\alpha - b^\delta|_2 = \delta$ für $x^\alpha = (A^TA+\alpha I_N)^{-1}A^Tb^\delta$.
Es gilt $\alpha(\delta) \longrightarrow 0$ für $\delta \longrightarrow 0$.

4 Eigenwertberechnung

(4.3) $H \in \mathbb{R}^{N\times N}$ heißt *Hessenberg*-Matrix, wenn $H[n+2:N,n] = 0_{N-n-1}$ für $n = 1,...,N-2$.

(4.4) Sei $A \in \mathbb{R}^{N\times N}$. Dann existiert eine orthogonale Matrix $Q \in \mathbb{R}^{N\times N}$, so dass $H = QAQ^\top$ eine Hessenberg-Matrix ist. Die Berechnung benötigt $O(N^3)$ Operationen.

Wenn A symmetrisch ist, dann ist H eine Tridiagonalmatrix.

(4.8) Sei $A \in \mathbb{R}^{N\times N}$ symmetrisch, tridiagonal, und irreduzibel, d.h. $A[n-1,n] = A[n,n-1] \neq 0$ und $A[n+2:N,n] = A[n,n+2:N]^\top = 0_{N-n-1}$.
Die charakteristischen Polynome $P_n(t) = \det\big(A[1:n,1:n] - tI_n\big)$ der Hauptuntermatrizen lassen sich durch eine Dreitermrekursion berechnen:
Setze $P_0 \equiv 1$. Dann gilt $P_1(t) = A[1,1] - t$ und

$$P_n(t) = (A[n,n] - t)P_{n-1}(t) - A[n-1,n]^2 P_{n-2}(t)\,.$$

Sie bilden eine *Sturmsche Kette*: Für die Nullstellen $\lambda_1^n \le \lambda_2^n \le \cdots \le \lambda_n^n$ von P_n gilt

$$\lambda_{k-1}^{n-1} < \lambda_k^n < \lambda_k^{n-1}\,, \qquad k = 1,...,n$$

(mit $\lambda_0^n = -2\|A\|_\infty$ und $\lambda_{n+1}^n = 2\|A\|_\infty$), und es gilt für $t \in (-\|A\|_\infty, \|A\|_\infty)$

$$\lambda_k^n < t \le \lambda_{k+1}^n$$

mit $k = W_n(t)$ und $W_n(t) = \#\big\{j \in \{1,...,n\} : P_j(t)P_{j-1}(t) < 0 \text{ oder } P_j(t) = 0\big\}$.

4 Eigenwertberechnung

Sei $A \in \mathbb{R}^{N\times N}$ symmetrisch mit Eigenwerten $\lambda_1, ..., \lambda_N$ und ONB aus Eigenvektoren $v_1, ..., v_N$. Dann gilt

$$A = \sum_n \lambda_n v_n (v_n)^\top \qquad \text{und} \qquad Ax = \sum_n \lambda_n (v_n^\top x) v_n .$$

(4.9) Der *Rayleigh-Quotient* ist

$$r(A,x) = \frac{x^\top Ax}{x^\top x}, \qquad x \in \mathbb{R}^N,\ x \neq 0_N .$$

(4.10) Sei $|\lambda_1| = \rho(A)$ und $|\lambda_n| < |\lambda_1|$ für $n = 2, ..., N$. Dann gilt für alle $w \in \mathbb{R}^N$ mit $w^\top v_1 > 0$

$$\lim_{k\longrightarrow\infty} r(A, A^k w) = \lambda_1 , \qquad \lim_{k\longrightarrow\infty} \frac{1}{|A^k w|_2} A^k w = v_1 .$$

(4.11) Sei $|w|_2 = 1$ und $\mu = r(A,w)$. Dann gilt

$$\min_{n=1,...,N} |\lambda_n - \mu| \leq |Aw - \mu w|_2 .$$

Eine konvergente Folge (d^k) in $\mathbb{R}$ mit Grenzwert d^* konvergiert

a) *linear*, wenn $c \in (0,1)$ und $k_0 > 0$ existieren mit
$$|d^{k+1} - d^*| \leq c|d^k - d^*| \qquad \text{für } k \geq k_0$$

b) *superlinear*, wenn zu jedem $\varepsilon > 0$ ein $k_0 > 0$ existiert mit
$$|d^{k+1} - d^*| \leq \varepsilon |d^k - d^*| \qquad \text{für } k \geq k_0$$

c) *von der Ordnung* $p > 1$, wenn wenn $C > 0$ existiert mit
$$|d^{k+1} - d^*| \leq C|d^k - d^*|^p .$$

4 Eigenwertberechnung

(4.12) Inverse Iteration mit variablem shift

S0) Wähle $z^0 \in \mathbb{R}^N$, $z^0 \neq 0_N$, $\varepsilon \geq 0$. Setze $k = 0$.
S1) Setze $w^k = \frac{1}{|z^k|_2} z_k$, $\mu_k = r(A, w^k)$.
S2) Falls $|Aw^k - \mu_k w^k|_2 \leq \varepsilon$ STOP.
S3) Berechne $z^{k+1} = (A - \mu_k I_N)^{-1} w^k$.
S4) Setze $k := k+1$, gehe zu S1).

Wenn der Startvektor z^0 hinreichend nahe bei einem Eigenvektor v_m mit isoliertem Eigenwert λ_m liegt, konvergiert die Iteration kubisch (d.h. von der Ordnung $p = 3$).

(4.13) QR-Iteration mit shift ($A \in \mathbb{R}^{N \times N}$ symmetrisch)

S0) Berechne $A_0 = QAQ^\top$ tridiagonal (Hessenberg-Transformation).
Wähle $\varepsilon \geq 0$. Setze $k = 0$.

S1) Falls $|A_k[n+1, n]| \leq \varepsilon$ für ein n:
getrennte Eigenwertberechnung für $A_k[1:n, 1:n]$ und $A_k[n+1:N, n+1:N]$.

S2) Berechne $d_k = \frac{1}{2}(A_k[N-1, N-1] - A_k[N, N])$ und

$$s_k = A_k[N, N] + d_k - \operatorname{sgn}(d_k)\sqrt{d_k^2 + A_k[N-1, N]^2}\,.$$

S3) Berechne QR-Zerlegung $Q_k R_k = A_k - s_k I_N$ und setze $A_{k+1} = R_k Q_k + s_k I_N$.
S4) Setze $k := k+1$, gehe zu S1).

Es gilt $A_{k+1} = Q_k^\top A_k Q_k$. Falls der shift $\mu_k = A_k[N, N]$ gewählt wird, entspricht die QR-Iteration der Inversen Iteration mit variablem shift und Startvektor $z^0 = e_N$.

4 Eigenwertberechnung

(4.14) Gershgorin

Zu $A \in \mathbb{R}^{N\times N}$ sind die Gershgorin-Kreise durch

$$K_n = \left\{\lambda \in \mathbb{C}\colon |\lambda - A[n,n]| \leq \sum_{k\neq n} |A[n,k]|\right\}, \qquad n = 1,\ldots,N$$

defininiert. Dann gilt

$$\sigma(A) \subset \bigcup_{n=1}^{N} K_n\,.$$

(4.15) Sei $A \in \mathbb{R}^{N\times N}$ symmetrisch mit Eigenwerten $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$.

$$\begin{aligned} \lambda_n &= \max_{\dim S = n} \min_{0_N \neq x \in S} r(A,x)\,, \\ \lambda_{N+1-n} &= \min_{\dim S = n} \max_{0_N \neq x \in S} r(A,x)\,. \end{aligned}$$

5 Iterative Lösungsverfahren für lineare Gleichungen

(5.1) Sei $A, B \in \mathbb{R}^{N\times N}$ mit $\rho(I_N - BA) < 1$.
Dann ist A invertierbar, und für alle $b \in \mathbb{R}^N$ und alle Startvektoren $x^0 \in \mathbb{R}^N$ konvergiert die Iteration

$$x^{k+1} = x^k + B(b - Ax^k), \qquad k = 0, 1, 2, \ldots$$

gegen $\lim_{k\longrightarrow\infty} x^k = A^{-1}b$.

(5.2) Sei $K \in \mathbb{R}^{N\times N}$ und $\varepsilon > 0$
Dann existiert eine Vektor-Norm $|\cdot|$ und eine zugeordnete Matrix-Norm $\|\cdot\|$, so dass $\|K\| \leq \rho(K) + \varepsilon$ gilt.

Anwendung: Es gilt $|x - x^k| \leq \|I_N - BA\|^k |x - x^0|$ (lineare Konvergenz).

(5.3) Konvergenz des Gauß-Seidel-Verfahrens
Sei $A = L + D + R \in \mathbb{R}^{N\times N}$ symmetrisch positiv definit und sei $B = (L + D)^{-1}$.
Dann gilt bezüglich der Energienorm $|x|_A = \sqrt{x^T Ax}$ und $\|K\|_A = \sup_{x\neq 0_N} \frac{|Kx|_A}{|x|_A}$

$$\|I_N - BA\|_A < 1\,.$$

Anwendung der Neumannschen Reihe ergibt dann $A^{-1} = \sum_{k=0}^{\infty} (I_N - BA)^k B$.

5 Iterative Lösungsverfahren für lineare Gleichungen

(5.4) Eine Matrix $A \in \mathbb{R}^{N\times N}$ heißt stark diagonal-dominant, wenn

$$\sum_{\substack{k=1\\k\neq n}}^{N} |A[n,k]| \leq |A[n,n]|\,, \qquad n = 1,...,N\,,$$

und wenn ein $j \in \{1,\ldots,N\}$ existiert $\sum_{\substack{k=1\\k\neq n}}^{N} |A[j,k]| < |A[j,j]|$.

(5.5) Eine Matrix $A \in \mathbb{R}^{N\times N}$ sei irreduzibel. Dann existiert zu jedem Paar $j \neq n$ eine Folge $j = j_0, j_1, j_2, \ldots, j_R = n$ mit $A[j_1, j_0] \neq 0, ..., A[j_R, j_{R-1}] \neq 0$.

(5.6) Eine Matrix $A \in \mathbb{R}^{N\times N}$ sei stark diagonal-dominant und irreduzibel. Dann gilt:

a) A ist regulär und das Jacobi-Verfahren $x^{k+1} = x^k + \mathrm{diag}(A)^{-1}(b - Ax^k)$ konvergiert.

b) Sei $A[n,n] > 0$ für alle n. Dann ist A positiv definit.

c) Sei $A[n,n] > 0$ und $A[n,k] \leq 0$ für $n \neq k$. Dann ist $A^{-1}[n,k] \geq 0$ für alle n,k.

5 Iterative Lösungsverfahren: Krylov-Verfahren

(5.7) Zu $C \in \mathbb{R}^{N\times N}$ und $d \in \mathbb{R}^N$ ist k-te *Krylov-Raum*

$$\mathscr{K}_k(C,d) = \operatorname{span}\{d, Cd, ..C^{k-1}d\} = \{P(C)d \colon P \in \mathbb{P}_{k-1}\}.$$

(5.8) Zu einer regulären Matrix $A \in \mathbb{R}^{N\times N}$, einer rechten Seite $b \in \mathbb{R}^N$ und einem Startwert $x^0 \in \mathbb{R}^N$ sei $x \in \mathbb{R}^N$ die Lösung von $Ax = b$ und $r^0 = b - Ax^0$.
Sei $B \in \mathbb{R}^{N\times N}$ eine reguläre Matrix (*Vorkonditionierer*).
Wenn $\dim \mathscr{K}_k(AB, r^0) < k$ für ein k gilt, dann ist $x \in x^0 + \mathscr{K}_{k-1}(BA, Br^0)$.

Sei $\langle\cdot,\cdot\rangle$ ein Skalarprodukt in $\mathbb{R}^N$.

Gram-Schmidt-Verfahren zur Berechnung einer Orthonormalbasis $v^1, \dots, v^k$ von

$$\mathscr{K}_k(BA, Br^0) = \operatorname{span}\{Br^0, BABr^0, ..., (BA)^{k-1}Br^0\} = \{V_k y \colon y \in \mathbb{R}^k\}, \quad V_k = \left(v^1 | \cdots | v^k\right).$$

S0) Wähle $x^0 \in \mathbb{R}^N$, setze $r^0 = b - Ax^0$, $z^1 = Br^0$, $h_{10} = |z^1|_V$ und $v^1 = \frac{1}{h_{10}} z^1$.

S1) Für $k = 1, 2, 3, ...$ berechne $w^k = BAv^k$,

$$\begin{aligned} z^{k+1} &= w^k - \sum_{j=1}^{k} h_{jk} v^j \text{ mit } h_{jk} = \langle v^j, w^k \rangle_V \\ v^{k+1} &= \frac{1}{h_{k+1,k}} z^{k+1} \text{ mit } h_{k+1,k} = |z^{k+1}|_V \end{aligned}$$

Dann gilt $BAv^k = \sum_{j=1}^{k+1} h_{jk} v^j$, also $BAV_k = V_{k+1} H_k$ mit $H_k = (h_{jm}) \in \mathbb{R}^{k+1,k}$.

5 Iterative Lösungsverfahren: GMRES-Verfahren

S0) Wähle $x^0 \in \mathbb{R}^N$, $\varepsilon > 0$.
Berechne $r^0 = b - Ax^0$, $z^1 = Br^0$, $h_{10} = |z^1|_2$ und $v^1 = \frac{1}{h_{10}} z^1$. Setze $k = 1$.

S1) Berechne $w^k = BAv^k$ und

$$z^{k+1} = w^k - \sum_{j=1}^{k} h_{jk} v^j \text{ mit } h_{jk} = (v^j)^T w^k$$

$$v^{k+1} = \frac{1}{h_{k+1,k}} z^{k+1} \text{ mit } h_{k+1,k} = |z^{k+1}|_2$$

S2) Berechne $y^k \in \mathbb{R}^k$ mit $\rho_k = |H_k y_k - h_{10} e^1|_2 = \min!$
Dabei ist $H_k = (h_{jm})_{j=1,\dots,k+1,\ m=1,\dots,k} \in \mathbb{R}^{k+1,k}$.

S3) Wenn $\rho_k < \varepsilon$, setze $x^k = x^0 + \sum_{j=1}^{k} y_j^k v^j$ STOP.

S4) Setze $k := k+1$ und gehe zu S1).

(5.4) Es gilt $\rho_k = \min_{z \in x^0 + \text{span}\{v^1,\dots,v^k\}} |B(b - Az)|_2$.

Das GMRES-Verfahren ist wohldefiniert, und wenn es abbricht, gilt $|x - x^k|_2 \le \|(BA)^{-1}\|_2 \varepsilon$.

(5.5) Für $C \ge \alpha > 0$ gelte $z^T BAz \ge \alpha z^T z$ und $\|BA\|_2 \le C$.

Dann gilt für das GMRES-Verfahren $|x^k - x|_2 \le \kappa_2(BA) \left(1 - \frac{\alpha^2}{C^2}\right)^{k/2} |x^0 - x|_2$.

5 Iterative Lösungsverfahren: CG-Verfahren

S0) Wähle $x^0 \in \mathbb{R}^N$, $\varepsilon > 0$.
Berechne $r^0 = b - Ax^0$, $y^0 = Br^0$, $\rho_0 = (y^0)^T r^0$ und $d^1 = y^0$. Setze $k = 0$.

S1) Falls $\rho_k \leq \varepsilon$ STOP

S2) Setze $k := k+1$ und berechne

$$\begin{aligned} u^k &= Ad^k \\ \alpha_k &= \frac{\rho_{k-1}}{(u^k)^T d^k} \\ x^k &= x^{k-1} + \alpha_k d^k \\ r^k &= r^{k-1} - \alpha_k u^k \\ y^k &= Br^k \\ \rho_k &= (y^k)^T r^k \\ d^{k+1} &= y^k + \frac{\rho_k}{\rho_{k-1}} d^k \end{aligned}$$

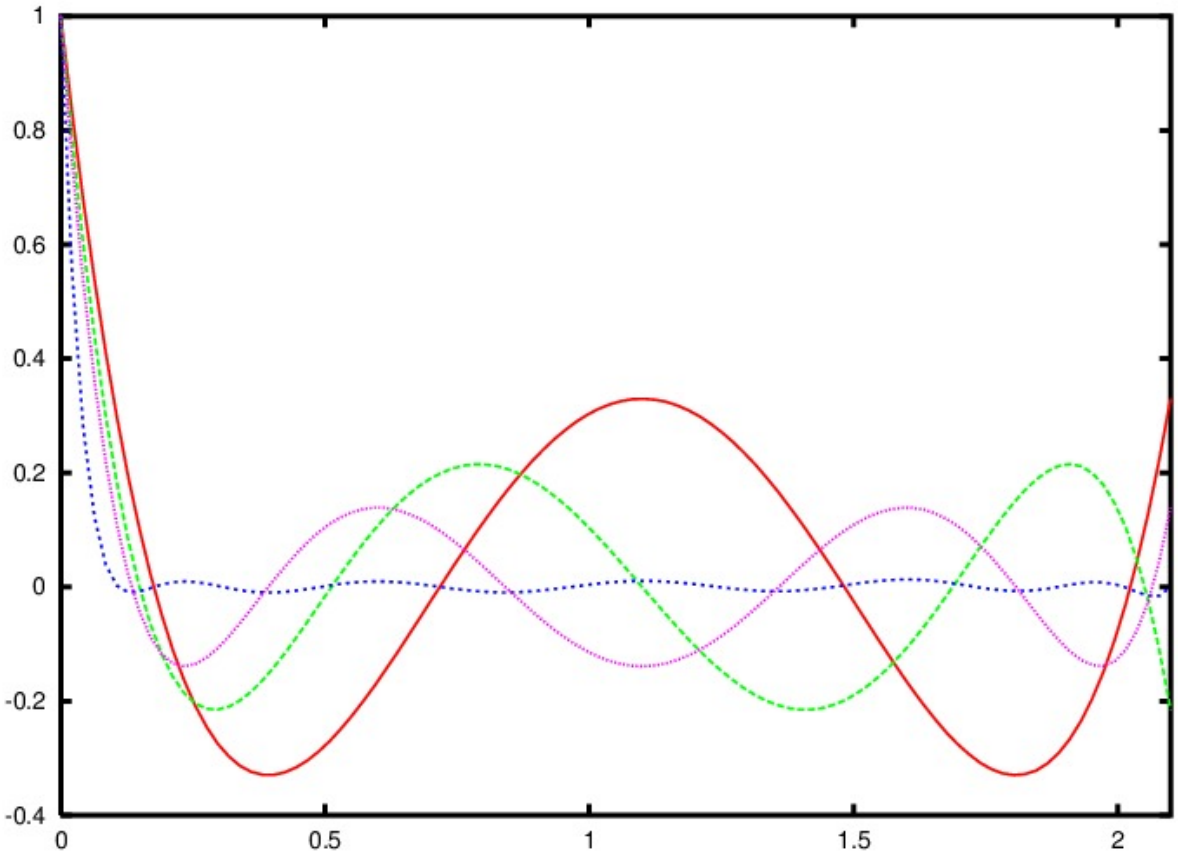
Gehe zu S1).

(5.6)
$$\begin{aligned} |x^k - x|_A &= \min_{z \in x^0 + \operatorname{span}\{d^1,\ldots,d^k\}} |z - x|_A \\ &\leq \min_{P \in \mathbb{P}_k,\ P(0)=1} \max_{\lambda \in \sigma(BA)} |P(\lambda)| |x^0 - x|_A \leq 2\left(\frac{\sqrt{\kappa(BA)}-1}{\sqrt{\kappa(BA)}+1}\right)^k |x^0 - x|_A . \end{aligned}$$

5 Transformierte Chebychev-Polynome

(5.6)

$$\min_{P\in\mathbb{P}_k,\ P(0)=1}\ \max_{t\in[a,b]} |P(t)| \le 2\left(\frac{\sqrt{\frac{b}{a}}-1}{\sqrt{\frac{b}{a}}+1}\right)^k$$



P_4, P_5, P_6, P_{12} mit $P_k(0) = 1$ zu $[a,b] = [0.1, 2.1]$

6 Lösungsverfahren für nichtlineare Gleichungen

(6.1) Sei $\mathscr{D} \subset \mathbb{R}^N$ konvex und $F\colon \mathscr{D} \longrightarrow \mathbb{R}^N$ ein stetig differenzierbares Vektorfeld.
Sei $x^* \in \mathscr{D}$ eine Nullstelle von F, und sei $B \in \mathbb{R}^{N\times N}$.
Wenn $\rho\big(I_N - B\mathrm{D}F(x^*)\big) < 1$ gilt, dann existiert ein $\delta > 0$, so dass für alle $x^0 \in \overline{\mathscr{U}}(x^*, \delta)$ die Fixpunktiteration $x^{k+1} = \Phi(x^k)$ mit $\Phi(x) = x - BF(x)$ linear gegen x^* konvergiert.

(6.2) Banachscher Fixpunktsatz
Sei $\mathscr{D} \subset \mathbb{R}^N$ konvex und $\phi\colon \mathscr{D} \longrightarrow \mathscr{D}$ kontrahierend mit $q \in (0.1)$, d.h.

$$|\phi(y) - \phi(z)| \le q|y - z| \quad \text{für} \quad y, z \in \mathscr{D}.$$

Dann existiert genau ein Fixpunkt $x^* \in \mathscr{D}$ von ϕ, d.h. $\phi(x^*) = x^*$, und es gelten für die Fixpunktiteration $x^{k+1} = \Phi(x^k)$ die Abschätzungen

$$|x^k - x^*| \le \frac{q^k}{1-q}\,|x^0 - x^1| \qquad \text{und} \qquad |x^k - x^*| \le \frac{q}{1-q}\,|x^k - x^{k-1}|.$$

(6.3) Seien $\alpha, \beta, \gamma > 0$ mit $2\alpha\gamma < \beta^2$ und $P(t) = \alpha - \beta t + \frac{\gamma}{2}t^2$.
Dann konvergiert für $t_0 = 0$ das Newton-Verfahren

$$t_{k+1} = t_k - P'(t_k)^{-1} P(t_k)$$

quadratisch gegen die kleinste Nullstelle t^* von P, und $\{t_k\}$ ist monoton steigend mit

$$t_0 = 0 < t_1 < \cdots < t_k < t_{k+1} = t_k + \frac{\gamma}{2}\,\frac{(t_k - t_{k-1})^2}{\beta - \gamma t_k} \le t^* = \frac{2\alpha}{\beta + \sqrt{\beta^2 - 2\alpha\gamma}}.$$

6 Lösungsverfahren für nichtlineare Gleichungen

(6.3) Sei $\mathscr{D} \subset \mathbb{R}^N$ offen und $F\colon \mathscr{D} \longrightarrow \mathbb{R}^N$ ein stetig differenzierbares Vektorfeld. Sei $x^0 \in \mathscr{D}$ mit

a) $|F(x^0)| \leq \alpha$

b) $|\mathrm{D}F(x^0)y| \geq \beta\,|y|$ für $y \in \mathbb{R}^N$

c) $\overline{\mathscr{U}}(x^0, 2\alpha/\beta) = \{z \in \mathbb{R}^N\colon |z - x^0| \leq 2\alpha/\beta\} \subset \mathscr{D}$

d) $\|\mathrm{D}F(y) - \mathrm{D}F(z)\| \leq \gamma |y - z|$ für $y, z \in B(x^0, 2\alpha/\beta)$

e) $2\alpha\gamma < \beta^2$

mit $\alpha, \beta, \gamma > 0$. Dann ist das Newton-Verfahren $x^{k+1} = x^k - \mathrm{D}F(x^k)^{-1} F(x^k)$ wohldefiniert und konvergiert quadratisch gegen $x^* \in \mathscr{D}$ mit

$$|x^* - x^0| \leq \frac{2\alpha}{\beta + \sqrt{\beta^2 - 2\alpha\gamma}}.$$

Gedämpftes Newton-Verfahren

S0) Wähle $x^0 \in \mathscr{D}$, $\varepsilon > 0$, $\theta \in (0,1)$. Setze $k = 0$.

S1) Falls $|F(x^k)| \leq \varepsilon$ STOP

S2) Löse $\mathrm{D}F(x^k)d^k = -F(x^k)$.

S3) Bestimme $s_k \in \{1, \theta, \theta^2, ..., \theta^r\}$ mit $x^k + s_k d^k \in \mathscr{D}$ und $|F(x^k + s_k d^k)| < |F(x^k)|$.

S4) Setze $x^{k+1} = x^k + s_k d^k$, $k := k + 1$ und gehe zu S1).