

# EINFÜHRUNG IN DIE GRUNDLAGEN DER NUMERIK

Institut für Numerische Simulation  
Rheinische Friedrich-Wilhelms-Universität Bonn

Wintersemester 2014/2015

# SCHNELLER LÖSER FÜR LINEARE GLEICHUNGSSYSTEME

## DIREKTE LÖSER (TEUER)

Basieren auf multiplikativen Zerlegungen / Faktorisierungen

$$A = B \cdot C$$

## KLASSISCHE ITERATIVE VERFAHREN (LANGSAME KONVERGENZ)

Basieren auf einem Fixpunktansatz

$$Ax = b \leftrightarrow F(x) = x + b - Ax \rightarrow \text{Fixpunkt}$$

## OPTIMIERUNGSVERFAHREN

Ein weiterer Zugang/Umformulierung

$$Ax = b \leftrightarrow F(x) = F(x, b, A) \rightarrow \min$$

als Optimierungs- oder Minimierungsproblem.

$$F(x) := \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle$$

- Wir nehmen an das  $A \in \mathbb{R}^{n \times n}$  sym.pos.def. dann gilt

$$x_{\star} := \operatorname{argmin} F(x) \Leftrightarrow Ax_{\star} = b$$

- Optimierungsverfahren bestehen immer aus zwei elementaren Schritten:
  - Bestimmung der Suchrichtung, des Suchraums.
  - Bestimmung der Schrittweite entlang der Suchrichtung.

## STEILSTER ABSTIEG

- Setze  $x_0 := y$ ;  $r_0 := b - Ax_0$ ;  $m := 0$ .
- Solange  $\|r_m\| > \epsilon$ 
  - Setze  $\lambda_m := \frac{\langle r_m, r_m \rangle}{\langle Ar_m, r_m \rangle}$
  - Setze  $x_{m+1} := x_m + \lambda_m r_m$
  - Setze  $r_{m+1} := r_m - \lambda_m Ar_m$
  - Setze  $m = m + 1$

- Lokal optimale Suchrichtung - der Gradient von  $F$ .
- Das ist ein **nicht-lineares** Verfahren, denn

$$x_{m+1} := x_m + \lambda_m r_m$$

enthält Produkt von Unbekannten.

- Iterationsupdate hat besondere Struktur (Krylovraum).
- Für die Residuen gilt (Projektionsverfahren)

$$\langle r_{m+1}, r_m \rangle = 0$$

## FOLGE DER ITERIERTEN

$$x_{m+1} \in x_0 + \text{span}\langle r_0, \dots, r_m \rangle, \quad r_{m+1} \in \text{span}\langle r_0, Ar_0, \dots, A^m r_0 \rangle$$

## KRYLOVRAUM

Sei  $A \in \mathbb{K}^{n \times n}$  und  $v \in \mathbb{K}^n$  gegeben. Dann definiert

$$\mathcal{K}_m(A, v) := \text{span}\langle v, Av, \dots, A^{m-1}v \rangle$$

einen Unterraum des  $\mathbb{K}^n$  und wird **Krylovraum** der Ordnung  $m$  zu  $(A, v)$  genannt.

Die Residuen des Gradientenverfahrens liegen alle in dem Krylovraum  $\mathcal{K}_m(A, r_0)$ . Die Iterierten im affinen Raum  $x_0 + \mathcal{K}_m(A, r_0)$ .

# EIGENSCHAFTEN VON $\mathcal{K}_m(A, v)$

- Jedes Element  $w \in \mathcal{K}_m(A, v)$  ist ein Polynom in  $A$  angewendet auf  $v$ .

$$w = \alpha_0 v + \alpha_1 A v + \alpha_2 A^2 v + \cdots + \alpha_{m-1} A^{m-1} v = \pi(A)v$$

- Die Krylovräume sind geschachtelt  $\mathcal{K}_m(A, v) \subset \mathcal{K}_{m+1}(A, v)$ .
- Die Dimension wächst maximal um eins wenn die Ordnung um eins erhöht wird.

## GRAD VON $v$ BZGL. $A$

$$\text{grad}_A(v) := \min\{j : \dim(\mathcal{K}_{j+1}(A, v)) < j + 1\}$$

## GRADIENTENVERFAHREN ALS MINIMIERUNG

Es gilt nach Konstruktion von oben  $\lambda_m := \frac{\langle r_m, r_m \rangle}{\langle Ar_m, r_m \rangle}$  und mit

$$r_{m+1} := b - Ax_{m+1}, \quad x_{m+1} := x_m + \lambda_m r_m.$$

dann auch  $\langle r_{m+1}, r_m \rangle = 0$ .

## GRADIENTENVERFAHREN ALS PROJEKTIONSMETHODE

Wir können das aber auch umformulieren: Finde  $\lambda_m$ , so dass für

$$x_{m+1} := x_m + \lambda_m r_m, \quad r_{m+1} := b - Ax_{m+1}$$

eben  $\langle r_{m+1}, r_m \rangle = 0$  gilt.

Wieder zwei Schritte:

- Wähle einen Suchraum  $\mathcal{V}$  (hier  $\mathcal{V} = \text{span}\langle r_m \rangle$ ).
- Bestimme  $x_{m+1} \in x_m + \mathcal{V}$  so, dass das resultierende Residuum  $r_{m+1} := b - Ax_{m+1}$  senkrecht auf einem gewählten Testraum  $\mathcal{W}$  (hier  $\mathcal{W} = \mathcal{V} = \text{span}\langle r_m \rangle$ ) steht.

## IN WORTEN

Konstruktion der Iterierten so, dass die Projektionen der Residuen auf den Testraum verschwinden.

## IN MATRIX-SCHREIBWEISE

Sei  $V$  die Matrix der Basisvektoren von  $\mathcal{V}$ ,  $W$  entsprechend für  $\mathcal{W}$ .  
Dann suchen wir  $y \in \mathbb{R}^{\dim(\mathcal{V})}$  mit

$$x_{m+1} := x_m + Vy, \quad \text{mit} \quad W^T r_{m+1} = W^T (b - A(x_m + Vy)) = 0$$

oder einfacher

$$W^T AVy = W^T r_m.$$



# NOCH EIN PROJEKTIONSVERFAHREN

## GAUß-SEIDEL-VERFAHREN

Wähle sukzessive in jedem  $i$ ten (Sub-)Iterationsschritt

$$\mathcal{W} = \mathcal{V} = \text{span}\langle e_i \rangle$$

den  $i$ ten Einheitsvektor. Dann erhalten wir

$$y := (e_i^T A e_i)^{-1} (e_i^T r_{m+\frac{i}{n}}) = \frac{1}{a_{ii}} \langle b - A x_{m+\frac{i}{n}}, e_i \rangle.$$

Mit ein paar Umformungen erhalten wir hieraus die Iterationsvorschrift des Gauß-Seidel-Verfahrens.

## BEMERKUNG

Für das Gauß-Seidel-Verfahren wurde nie die Symmetrie oder Definitheit gefordert, wie oben beim Gradientenverfahren. Der Projektionszugang ist also allgemeiner als der Minimierungszugang. Es muß nur die Matrix  $W^T A V$  invertierbar sein.

## SATZ

Es gelte entweder

- 1  $A$  ist positiv definit und  $\mathcal{V} = \mathcal{W}$ , oder
- 2  $A$  ist nicht singulär und  $\mathcal{W} = A\mathcal{V}$ .

Dann ist  $W^TAV$  nicht-singulär für beliebige Basen von  $\mathcal{W}, \mathcal{V}$ .

- Für  $A$  (sym.)pos.def. macht also  $\mathcal{V} = \mathcal{W}$  und  $\mathcal{W} = A\mathcal{V}$  Sinn.
- Sowohl das Gradientenverfahren als auch das Gauß–Seidel-Verfahren verwenden eindimensionale Such- und Testräume.
- Bisher haben wir noch kein Verfahren mit höher dimensionalen Such- und Testräumen kennengelernt.
- Noch ein letztes Beispiel mit eindimensionalen  $\mathcal{W}$  und  $\mathcal{V}$ .

# MINIMAL RESIDUAL ALS PROJKTIONSVERFAHREN

Nehmen wir an, dass  $A$  positiv definit ist aber nicht symmetrisch und wählen  $\mathcal{W} = A\mathcal{V}$  mit  $\mathcal{V} = \text{span}\langle r_m \rangle$ .

## EIN MINIMAL RESIDUAL VERFAHREN

$$\alpha_m = (W^T A V)^{-1} W^T r_m = \frac{\langle A r_m, r_m \rangle}{\langle A r_m, A r_m \rangle}, \quad x_{m+1} = x_m + \alpha_m r_m$$

- Das Verfahren minimiert in jedem Iterationsschritt

$$f(x) := \|b - Ax\|_2^2 = \langle b - Ax, b - Ax \rangle$$

für alle  $x = x_m + \alpha_m r_m$  (man bestimmt also  $\alpha_m \in \mathbb{R}$ ).

- Das Verfahren mit obigen Voraussetzungen konvergiert, aber wird so selten verwendet.
- Das eigentliche MINRES-Verfahren kommt gleich!

## GRUNDSÄTZLICH

Wie immer: Eine Basis von orthonormalen Vektoren ist wünschenswert.

## WAHL DER RÄUME

Bisher haben wir immer nur die eindimensionalen Räume  $\text{span}\langle r_m \rangle$  verwendet. Jetzt wollen wir mehr-dimensionale Krylovräume verwenden.

## AUFGABE

Wir müssen also versuchen eine orthonormale Basis für unsere Krylovräume zu konstruieren und wollen dies mit möglichst geringen Kosten!

## ARNOLDI-PROZESS

- Gegeben sei eine Matrix  $A \in \mathbb{R}^{n \times n}$  und  $v \in \mathbb{R}^n$
- Setze  $w_1 := \frac{v}{\|v\|}$ .
- Iteriere für  $k = 1, 2, \dots, n$

$$r_k = Aw_k, \quad h_{i,k} = w_i^* r_k \text{ für } i = 1, 2, \dots, k,$$

und

$$r_k = r_k - \sum_{i=1}^k h_{i,k} w_i, \quad h_{k+1,k} = \|r_k\|_2,$$

setze für  $h_{k+1,k} \neq 0$

$$w_{k+1} = \frac{1}{h_{k+1,k}} r_k$$

- Die Vektoren  $w_k$  bilden eine Orthonormalbasis.
- Die Matrix  $H = (h_{i,k}) = W_k^* A W_k$  ist Hessenberg-Matrix.

## LANCZOS-PROZESS

Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch und  $v_1 = v$  ein beliebiger normierter Vektor. Ferner sei  $v_0 = 0$  im  $\mathbb{R}^n$  und  $\beta_0 = 0$ . Dann bilden die Vektoren  $\{v_i\}_{i=1}^k$  aus der dreistufigen Rekursionsformel

$$r_{i+1} := (A - \alpha_i \mathbb{I})v_i - \beta_{i-1}v_{i-1}, \quad v_{i+1} := \frac{r_{i+1}}{\beta_i},$$

mit

$$\beta_i := \|r_{i+1}\|, \quad \alpha_i := v_i^* A v_i$$

und  $i = 1, \dots, k-1$  eine Orthonormalbasis von  $\mathcal{K}(A, v)$ , falls alle  $\beta_i \neq 0$ .

# FULL ORTHOGONALIZATION METHOD

- Gegeben  $A$ ,  $x_0$ ,  $b$  und  $m$ .
- Setze  $r_0 := b - Ax_0$ ,  $\beta := \|r_0\|$ ,  $v_1 := r_0/\beta$  und  $H_m = (h_{i,j}) = 0$ .
- Für  $j = 1, \dots, m$ 
  - Setze  $w_j := AV_j$
  - Für  $i = 1, \dots, j$ 
    - Setze  $h_{i,j} := \langle w_j, v_i \rangle$ .
    - Setze  $w_j := w_j - h_{i,j}v_i$
  - Setze  $h_{j+1,j} := \|w_j\|$ . Falls  $h_{j+1,j} = 0$ , setze  $m := j$ . Sonst  $v_{j+1} := w_j/h_{j+1,j}$ .
- Setze  $y_m = H_m^{-1}(\beta e_1)$  und  $x_m = x_0 + V_m y_m$ .

## ALGORITHMUS

- Gegeben eine sym. Matrix  $A \in \mathbb{R}^{n \times n}$ ,  $b \in \mathbb{R}^n$  und  $x_0 \in \mathbb{R}^n$ .
- Setze  $d_{-1} = 0$ ,  $d_0 = r_0 = b - Ax_0$ .
- Iteriere für  $k = 0, 1, \dots$  bis  $\|r_{k+1}\|_2 \leq \epsilon$

$$\alpha_k = \frac{d_k^* A r_k}{d_k^* A^2 d_k}, \quad x_{k+1} = x_k + \alpha_k d_k, \quad r_{k+1} = r_k - \alpha_k A d_k$$

und

$$\beta_k = \frac{d_k^* A^3 d_k}{d_k^* A^2 d_k}, \quad \gamma_k = \frac{d_k^* A^3 d_{k-1}}{d_{k-1}^* A^2 d_{k-1}}, \quad d_{k+1} = A d_k - \beta_k d_k - \gamma_k d_{k-1}$$

- Das Verfahren ist für indefinite symmetrische Matrizen entworfen ( $A^2$  ist dann sym.pos.def.).



- Gegeben sei  $A \in \mathbb{R}^{n \times n}$ , rechte Seite  $b \in \mathbb{R}$  und  $x_0$ .
- Setze  $r_0 = b - Ax_0$ ,  $h_{1,0} = \|r_0\|_2$  und  $k = 0$ .
- Iteriere  $k = 1, 2, \dots$  bis  $h_{k+1,k} \leq \epsilon$  gilt: Berechne

$$w_k = \frac{1}{h_{k,k-1}} r_{k-1}, \quad r_k = Aw_k, \quad h_{i,k} = w_i^* r_k \text{ für } i = 1, 2, \dots, k$$

und

$$r_k = r_k - \sum_{i=1}^k h_{i,k} w_i, \quad h_{k+1,k} = \|r_k\|_2$$

- Löse das Ausgleichsproblem

$$z = \operatorname{argmin}_{y \in \mathbb{R}^k} \|e_1 h_{1,0} - H_k y\|$$

mit  $H_k = (h_{i,j}) \in \mathbb{R}^{(k+1) \times k}$  (Hessenberg) und setze mit  $W_k = [w_1, w_2, \dots, w_k]$

$$x_k = x_0 + W_k z$$

Zu  $A \in \mathbb{R}^{n \times n}$  sym.pos.def. definieren wir das quadratische Funktional

$$\Phi(x) = \frac{1}{2}x^*Ax - x^*b.$$

Dies hat ein eindeutiges Minimum. Für  $\hat{x} = A^{-1}b$  gilt

$$\Phi(x) - \Phi(\hat{x}) = \frac{1}{2}(x - \hat{x})^*A(x - \hat{x}) \geq 0$$

also ist  $\hat{x}$  das eindeutige Minimum von  $\Phi$ . Um also  $Ax = b$  zu lösen, können wir auch  $\operatorname{argmin} \Phi(x)$  bestimmen.

$$A\hat{x} = b \iff \hat{x} = \operatorname{argmin}_{x \in \mathbb{R}^n} \Phi(x)$$

Die Abweichung des Funktionals von seinem Minimum

$$\Phi(\mathbf{x}) - \Phi(\hat{\mathbf{x}}) = \frac{1}{2}(\mathbf{x} - \hat{\mathbf{x}})^* A(\mathbf{x} - \hat{\mathbf{x}})$$

ist ein vernünftiges Maß, um den Fehler  $\mathbf{x} - \hat{\mathbf{x}}$  zu messen.

$$\frac{1}{2}(\mathbf{x} - \hat{\mathbf{x}})^* A(\mathbf{x} - \hat{\mathbf{x}}) =: \frac{1}{2}\|\mathbf{x} - \hat{\mathbf{x}}\|_A$$

ist die Energienorm zur Matrix  $A$ . Wenn wir das Funktional nun sukzessive minimieren, reduzieren wir auch den Fehler in der Energienorm.

## ALGORITHMUS

- Gegeben sym.pos.def.  $A \in \mathbb{R}^{n \times n}$ ,  $b \in \mathbb{R}^n$  und  $x_0 \in \mathbb{R}^n$ .
- Setze  $k = 0$ , und  $d_0 = r_0 = b - Ax_0$ .
- Iteriere  $k = 0, 1, 2, \dots$  bis  $\|r_{k+1}\|_2 \leq \epsilon$

$$\alpha_k = \frac{\|r_k\|_2^2}{d_k^* A d_k}, \quad x_{k+1} = x_k + \alpha_k d_k, \quad r_{k+1} = r_k - \alpha_k d_k$$

und

$$\beta_k = \frac{\|r_{k+1}\|_2^2}{\|r_k\|_2^2}, \quad d_{k+1} = r_{k+1} + \beta_k d_k$$

## EIGENSCHAFTEN

- In exakter Arithmetik direktes, in der Praxis iteratives Verfahren.
- Konvergenzverhalten durch Kondition der Matrix  $A$  bestimmt

$$\|x - x_k\|_A \leq 2 \left( \frac{\sqrt{\kappa_2(A)} - 1}{\sqrt{\kappa_2(A)} + 1} \right)^k \|x - x_0\|_A$$

Wichtiges Hilfsmittel für die Konvergenzbeweise:

## KANTOROVICH UNGLEICHUNG

$B$  sei symmetrisch positiv definite reelle Matrix mit den extremalen Eigenwerten  $\lambda_{\max}$  und  $\lambda_{\min}$ . Dann gilt

$$\frac{\langle Bx, x \rangle \langle B^{-1}x, x \rangle}{\langle x, x \rangle^2} \leq \frac{(\lambda_{\max} + \lambda_{\min})^2}{4\lambda_{\max}\lambda_{\min}}$$

## GMRES

Sei  $A = X\Lambda X^{-1}$  diagonalisierbar. Dann gilt für die Iterierten des GMRES

$$\|b - Ax_k\|_2 \leq \kappa_2(X)\epsilon_k \|b - Ax_0\|$$

mit  $\epsilon_k = \min_{p \in \tilde{\Pi}_k} \max_{i=1, \dots, n} |p(\lambda_i)|$ .

$$Ax = b \iff W^{-1}Ax = W^{-1}b$$

- $W$  muss regulär sein.
- $A$  ist sym.pos.def,  $W^{-1}A$  i.A. nicht.
- Sei  $W$  sym.pos.def, dann existiert die Cholesky-Zerlegung

$$W = LL^T \text{ und mit } \tilde{A} := L^{-1}AL^{-T}, \quad \tilde{x} := L^T x, \quad \tilde{b} := L^{-1}b \\ \text{gilt } \tilde{A}\tilde{x} = \tilde{b} \text{ wobei}$$

- $\tilde{A}$  sym.pos.def ist.
- Kann man CG-Verfahren auf  $\tilde{A}$  anwenden, ohne  $\tilde{A}$ ,  $L$ ,  $L^{-1}$  explizit zu kennen?

## ALTERNATIVES SKALARPRODUKT

$$x^*y = \langle x, y \rangle_2, \quad \langle x, y \rangle_W := \langle x, Wy \rangle_2 = x^*Wy$$

$W^{-1}A$  nicht symmetrisch bzgl  $\langle \cdot, \cdot \rangle_2$ , aber symmetrisch bzgl  $\langle \cdot, \cdot \rangle_W$

## ALGORITHMUS

- Gegeben sym.pos.def.  $A, W \in \mathbb{R}^{n \times n}$ ,  $b \in \mathbb{R}^n$  und  $x_0 \in \mathbb{R}^n$ .
- Setze  $k = 0$ ,  $s_0 = b - Ax_0$  und  $d_0 = r_0 = W^{-1}s_0$ .
- Iteriere  $k = 0, 1, 2, \dots$  bis  $\|s_{k+1}\|_2 \leq \epsilon$

$$\alpha_k = \frac{s_k^* r_k}{d_k^* A d_k}, \quad x_{k+1} = x_k + \alpha_k d_k, \quad s_{k+1} = s_k - \alpha_k A d_k$$

und

$$r_{k+1} = W^{-1}s_k, \quad \beta_k = \frac{s_{k+1}^* r_{k+1}}{s_k^* r_k}, \quad d_{k+1} = r_{k+1} + \beta_k d_k$$

## WAHL DES VORKONDITIONERS $W$

- Optimaler Vorkonditionierer  $W = A$
- Quasi-optimale Vorkonditionierer  $\kappa_2(W^{-1}A) = O(1) \quad W \approx A$
- Berechnung von  $W^{-1}x$  muss schnell sein



## UNVOLLSTÄNDIGE $LR$ -ZERLEGUNG

- Gegeben sei  $A \in \mathbb{R}^{n \times n}$  und ein Besetzungsmuster  $E \in \{1, \dots, n\} \times \{1, \dots, n\}$ .
- Setze  $\tilde{L} = \mathbb{I}$  und  $\tilde{R} = 0$  auf  $\mathbb{R}^n$
- Iteriere für  $i = 1, 2, \dots, n$ 
  - Iteriere für  $k = 1, 2, \dots, i - 1$
  - Falls  $(i, k) \in E$ , setze  $\tilde{L}_{i,k} = \frac{1}{\tilde{R}_{k,k}} (A_{i,k} - \sum_{j=1}^{k-1} \tilde{L}_{i,j} \tilde{R}_{j,k})$ .
  - Iteriere für  $k = i, i + 1, \dots, n$
  - Falls  $(i, k) \in E$ , setze  $\tilde{R}_{i,k} = A_{i,k} - \sum_{j=1}^{i-1} \tilde{L}_{i,j} \tilde{R}_{j,k}$
- Es gilt für  $(i, j) \in E$ , dass  $A_{i,j} = (\tilde{L}\tilde{R})_{i,j}$ .
- Es gilt für  $(i, j) \notin E$ , dass  $\tilde{L}_{i,j} = 0 = \tilde{R}_{i,j}$
- Für symmetrisches  $A$  gilt  $\tilde{R} = \tilde{D}\tilde{L}^T$  mit  $\tilde{D} = \text{diag}(\tilde{R})$ .
- Wir erhalten dann  $\tilde{L}\tilde{D}\tilde{L}^T$  als unvollständige Cholesky-Zerlegung von  $A$ . Ein PCG-Verfahren mit diesem Vorkonditionierer heißt auch oft ICCG-Verfahren.

# VORKONDITIONIERER ÜBER LINEARE ITERATIONSVERFAHREN

## 1. NORMALFORM

$$\text{Lineare Iteration: } \Phi_1(x, b) = Mx + Nb$$

$$\text{Konsistenz: } (Ax_* = b \Rightarrow \Phi(x_*, b) = x_*) \Rightarrow M + NA = \mathbb{I}$$

## 2. NORMALFORM $M = \mathbb{I} - NA$

$$\Phi_2(x, b) = x - N(Ax - b)$$

## 3. NORMALFORM $N$ sei regulär

$$\Phi_2(x_m, b) = x_m - N(Ax_m - b) = x_{m+1}, \quad x_m - x_{m+1} = N(Ax_m - b)$$

Implizite Definition für  $x_{m+1}$

$$N^{-1}(x_m - x_{m+1}) = Ax_m - b, \quad W(x_m - x_{m+1}) = Ax_m - b$$

# KLASSISCHE ITERATIONSVERFAHREN

## MATRIZEN DER NORMALFORMEN

- $M$  Iterationsmatrix  $M = \mathbb{I} - NA$
- $N$  Matrix der 2. Normalform (Näherung für  $A^{-1}$ )
- $W = N^{-1}$  Matrix der 3. Normalform (Näherung für  $A$ )

## ADDITIVE ZERLEGUNG VON $A$

$$A = D - L - U$$

## KLASSISCHE ITERATIONSVERFAHREN

RICHARDSON  $W = \frac{1}{\omega}\mathbb{I},$   $M = \mathbb{I} - \omega A$

JACOBI  $W = D,$   $M = \mathbb{I} - D^{-1}A$

GAUSS-SEIDEL  $W = D - L,$   $M = \mathbb{I} - (D - L)^{-1}A$

SOR  $W = \frac{1}{\omega}(D - \omega L),$   $M = \mathbb{I} - \omega(D - \omega L)^{-1}A$

# KLASSISCHE ITERATIONSVERFAHREN ALS VORKONDITIONIERER

- Die Matrix der 3. Normalform kann als Vorkonditionierer genutzt werden.
- Man erhält die Anwendung von  $W^{-1}$  über einen Iterationsschritt mit Startwert 0.