

EINFÜHRUNG IN DIE GRUNDLAGEN DER NUMERIK

Institut für Numerische Simulation
Rheinische Friedrich-Wilhelms-Universität Bonn

Wintersemester 2014/2015

EIGENWERTPROBLEM

Zu einer Matrix $A \in \mathbb{R}^{n \times n}$ sind gesucht $\lambda \in \mathbb{C}$ und $v \in \mathbb{C}^n$ mit $\|v\| = 1$, so dass

$$Av = \lambda v.$$

λ heißt Eigenwert zum Eigenvektor v und (λ, v) Eigenpaar von A .

- Das Problem ist **nichtlinear**, denn λ und v sind gesucht und das Produkt λv taucht in der Problemstellung auf.

Erinnerung: Brückenproblem/Tragwerk

- Gesucht war bisher statisches Gleichgewicht unter äußeren Kräften
- Nun fragen wir nach dem dynamischen Verhalten: Die Positionen der Gelenke z_i sind nun zeitabhängig ($i = 1, \dots, 8$).

Position	z_i	$: [0, T] \rightarrow \mathbb{R}^2,$
Geschwindigkeit	\dot{z}_i	$: [0, T] \rightarrow \mathbb{R}^2,$
Beschleunigung	\ddot{z}_i	$: [0, T] \rightarrow \mathbb{R}^2,$
Auslenkung :	$x_i(t)$	$:= z_i(t) - z_i(0)$

- Schwingungsverhalten der Brücke
- Schaukelt sich die Auslenkung auf, Resonanzen, Einsturz der Brücke

MOTIVATION: SCHWINGUNGEN

- Auslenkung aus Gleichgewicht bewirkt Rückstellkraft

$$F_R = -Ax(t) \in \mathbb{R}^{16}, \quad A \text{ Steifigkeitsmatrix}$$

- Newtonsche Gesetze fordern (M Massen der Gelenke)

$$F_R = -Ax(t) = Mx''(t)$$

- Gewöhnliche Differentialgleichung zweiter Ordnung

$$Mx''(t) + Ax(t) = 0$$

- Betrachte spezielle Lösungen (Vereinfachung $M = \mathbb{I}$)

$$x''(t) + Ax(t) = 0$$

- A ist sym.pos.def., Eigenvektoren v sind VONS, Eigenwerte $\lambda > 0$

MOTIVATION: SCHWINGUNGEN

- Ansatz

$$x(t) = \cos(\sqrt{\lambda}t)v$$

- Erfüllt

$$x'(t) = -\sqrt{\lambda} \sin(\sqrt{\lambda}t)v, \quad x''(t) = -\lambda \cos(\sqrt{\lambda}t)v$$

Außerdem gilt $Av = \lambda v$, somit $x''(t) + Ax(t) = 0$.

- Die Funktion

$$z(t) = z(0) + \cos(\sqrt{\lambda}t)v$$

beschreibt eine cos-Schwingung des Tragwerks (Positionen der Gelenke) um die Ausgangslage $z(0)$.

- Je größer der Eigenwert λ , um so schneller schwingt die Brücke.
- Keine Dämpfung im System, periodischer Lösungsverlauf, wenig realistisch.

MOTIVATION: SCHWINGUNGEN

- Reibung bewirkt eine weitere Kraft neben der Rückstellkraft

$$F_F = -Dx'(t)$$

- Insgesamt erhalten wir das Modell

$$Mx''(t) + Ax(t) + Dx'(t) = 0$$

- Lösung dieses Problems ($M = \mathbb{I}$) sind

$$x(t) = \exp\left(-\frac{Dt}{2}\right) \cos(\omega t)v, \quad \omega := \sqrt{\lambda - \frac{D^2}{4}}.$$

- Ein einmal ausgelenktes/angeregtes System kommt also nach einer gewissen Zeit wieder zur Ruhe.

- Wenn wir aber eine dauerhafte Anregung haben

$$Mx''(t) + Ax(t) + Dx'(t) = \cos(\omega_0 t)v$$

in Richtung eines Eigenvektors v , können wieder periodische und sogar anwachsende Lösungen auftreten.

- Für $\omega_0 = \sqrt{\lambda}$, wobei λ der Eigenwert zu v ist, ist die Amplitude der Schwingung am größten und abhängig von der Dämpfung D kann sich die Schwingung aufschaukeln und die Brücke wird instabil.

EIGENWERTE UND BAUWERKE

Eine Spektralanalyse (Eigenwertanalyse) der Steifigkeitsmatrix von Bauwerken wesentliche Aufgabe für die Sicherheit. Besonders wichtig das Prüfen, ob mögliche/wahrscheinliche Anregungen (Erdbeben, Wind) zu Resonanzen führen können.

WAS SIND EIGENWERTE?

EIGENWERTPROBLEM

Zu einer Matrix $A \in \mathbb{R}^{n \times n}$ sind gesucht $\lambda \in \mathbb{C}$ und $v \in \mathbb{C}^n$ mit $\|v\| = 1$, so dass

$$Av = \lambda v. \quad (1)$$

λ heißt Eigenwert zum Eigenvektor v und (λ, v) Eigenpaar von A .

CHARAKTERISTISCHES POLYNOM

Eine Zahl λ ist genau dann ein Eigenwert von A , wenn gilt

$$\det(A - \lambda I) = 0$$

SPEKTRUM

$$\sigma(A) := \{\lambda \in \mathbb{C} : \det(A - \lambda I) = 0\}$$

ÄHNLICHE MATRIZEN

Zwei Matrizen A und B heißen ähnlich, wenn für ein beliebiges nicht-singuläres T gilt

$$B = T^{-1}AT$$

INVARIANZ DES SPEKTRUM

Ähnliche Matrizen haben dasselbe charakteristische Polynom und damit auch Spektrum

$$\sigma(A) = \sigma(T^{-1}AT), \quad T \text{ beliebig nicht-singulär}$$

DIAGONALISIERBARE MATRIZEN

Eine Matrix A heißt diagonalisierbar, wenn A zu einer Diagonalmatrix ähnlich ist. A ist genau dann diagonalisierbar, wenn A genau n linear unabhängige Eigenvektoren hat. Falls alle n Eigenwerte von A verschieden sind, dann ist A diagonalisierbar.

SCHUR-FAKTORISIERUNG

Zu jeder Matrix $A \in \mathbb{C}^{n \times n}$ gibt es eine unitäre Matrix $Q, R \in \mathbb{C}^{n \times n}$, so dass

$$Q^H A Q = R, \quad \{\text{diag}(R)\} = \sigma(A).$$

Selbst, wenn $A \in \mathbb{R}^{n \times n}$ gilt im Allgemeinen $Q, R \in \mathbb{C}^{n \times n}$.

REELLE SYMMETRISCHE MATRIZEN

Jede reelle symmetrische Matrix A lässt sich mittels einer orthogonalen Matrix $Q \in \mathbb{R}^{n \times n}$ ähnlich auf eine reelle Diagonalmatrix D transformieren.

$$Q^T A Q = Q^{-1} A Q = D, \quad \text{diag}(\sigma(A)) = D$$

A besitzt somit nur reelle Eigenwerte und n linear unabhängige Eigenvektoren (Spalten von Q).

EIGENSCHAFTEN

Für $A, B \in \mathbb{R}^{n \times n}$ gilt

- Für A nicht-singulär: $\lambda \in \sigma(A) \Leftrightarrow \lambda^{-1} \in \sigma(A^{-1})$.
- $\lambda \in \sigma(A) \Rightarrow \bar{\lambda} \in \sigma(A)$
- $\sigma(A) = \sigma(A^T)$
- $\sigma(AB) = \sigma(BA)$

ERSTE ABSCHÄTZUNG - INDUZIERTER MATRIXNORM

Für alle $\lambda \in \sigma(A)$ gilt $|\lambda| \leq \|A\|$.

RAYLEIGH QUOTIENT & WERTEBEREICH

$$\mu_A(x) := \frac{\langle Ax, x \rangle}{\langle x, x \rangle}, \quad W_A := \{\mu_A(x) : x \in \mathbb{C}^n \setminus 0\}$$

Mit einem beliebigen gegebenen Innenprodukt.

EIGENSCHAFTEN

- $\sigma(A) \subset W(A)$.
- Hausdorff: $\text{conv}(\sigma(A)) \subset W(A)$.
- Für normale $A \in \mathbb{C}^{n \times n}$ gilt $\text{conv}(\sigma(A)) = W(A)$
- $W(A)$ ist zusammenhängend.
- Ist A hermitesch, dann ist $W(A)$ das reelle Intervall $[\lambda_{\min}, \lambda_{\max}]$.
- Ist A schiefhermitesch ($A^H = -A$), dann ist $W(A)$ ein rein imaginäres Intervall.

RAYLEIGH

Für $A \in \mathbb{K}^{n \times n}$ hermitesch, die reellen Eigenwerte λ_i seien absteigend sortiert, $\{u_i\}$ seien die orthonormalen Eigenvektoren und

$$E_0 := \{0\}, \quad E_j := \text{span}\langle u_1, \dots, u_j \rangle.$$

Dann gilt

$$\lambda_j := \max_{x \in E_{j-1}^\perp \setminus \{0\}} \mu_A(x) = \min_{x \in E_j \setminus \{0\}} \mu_A(x).$$

COURANT-FISCHER

Für $A \in \mathbb{K}^{n \times n}$ hermitesch und

$$U_j := \{u \in \mathbb{K}^n : \dim(U) = j\}$$

gilt

$$\lambda_j := \min_{U \in U_{n+1-j}} \max_{x \in U \setminus \{0\}} \mu_A(x) = \max_{U \in U_j} \min_{x \in U \setminus \{0\}} \mu_A(x).$$

BAUER & FIKE

Sei $A \in \mathbb{C}^{n \times n}$ eine diagonalisierbare Matrix, d.h.

$$V^{-1}AV = \text{diag}(\{\lambda_i\}), \quad \{\lambda_i\} = \sigma(A)$$

Sei μ ein Eigenwert von der gestörten Matrix $A + E$, dann gilt

$$\min_{1 \leq i \leq n} |\lambda_i - \mu| \leq \|V\|_p \|V^{-1}\|_p \|E\|_p = \kappa_p(V) \|E\|_p$$

mit $p = 1, 2, \infty$.

SYMMETRISCHE MATRIZEN

Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix und μ ein Eigenwert von der gestörten Matrix $A + E$, dann gilt

$$\min_{1 \leq i \leq n} |\lambda_i - \mu| \leq \|E\|_2.$$

WEYL

Seien $A, B \in \mathbb{C}^{n \times n}$ hermitesch, dann gilt

$$|\lambda_j(A) - \lambda_j(B)| \leq \|A - B\|_2$$

KOROLLAR

Sei $A \in \mathbb{C}^{n \times n}$ hermitesch und $\{a'_{ii} \in \mathbb{R}\}$ eine Permutation der reellen Diagonaleinträge von A mit $a'_{ii} \geq a'_{jj}$ für $j > i$. Dann gilt für die absteigend sortierten Eigenwerte von A

$$|\lambda_j(A) - a'_{ii}| \leq \max_{j=1, \dots, n} \sum_{k=1, k \neq j}^n |a_{jk}|, \quad |\lambda_j(A) - a'_{ii}| \leq \sqrt{\sum_{j, k=1, \neq j}^n |a_{jk}|^2}$$

GERSCGORIN-KREISE

Zu einer Matrix $A = (a_{i,j}) \in \mathbb{R}^{n \times n}$ definieren wir für $i = 1, \dots, n$

$$K_i := \{z \in \mathbb{C} : |z - a_{i,i}| \leq \sum_{j \neq i} |a_{i,j}|\},$$

die Gerschgorin-Kreise. Dann gilt, dass alle Eigenwerte von A in der Vereinigung dieser Kreise liegen.

$$\sigma(A) \subseteq \bigcup_{i=1}^n K_i$$

BENDIXSON

Das Spektrum von A ist im Rechteck

$$R_B(A) := W\left(\frac{A + A^H}{2}\right) + W\left(\frac{A - A^H}{2}\right)$$

enthalten.

APPROXIMATION VON EIGENWERTEN

Angenommen es gibt ein vollständiges System von Eigenvektoren v_i von A , dann gilt

$$x = \sum_{i=1}^n \xi_i v_i, \quad \forall x \in \mathbb{R}^n$$

und somit

$$A^k x = \sum_{i=1}^n \xi_i \lambda_i^k v_i, \quad \forall x \in \mathbb{R}^n.$$

Falls jetzt gilt

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|,$$

dann gilt asymptotisch

$$A^k x \approx \xi_1 \lambda_1^k v_1.$$

ALGORITHMUS

Wähle Startvektor y^0 mit $\|y^0\|_2 = 1$. Für $k = 0, 1, 2, \dots$ berechne

$$\tilde{y}^{k+1} = Ay^k, \quad \lambda^{(k)} = (y^k)^H \tilde{y}^{k+1}, \quad y^{k+1} = \frac{\tilde{y}^{k+1}}{\|\tilde{y}^{k+1}\|_2}.$$

y^k approximiert einen Eigenvektor zum betragsgrößten Eigenwert, der durch $\lambda^{(k)}$ approximiert wird

EIGENSCHAFTEN

Falls $y^0 = \sum_{i=1}^n \eta_i v_i$ mit $\eta_1 \neq 0$, dann gilt mit $q := \frac{\lambda_2}{\lambda_1} < 1$

- $\|\tilde{y}^k\| = |\lambda_1| + O(q^k)$
 - Falls $\lambda_1 > 0$, gilt $\|y^k - \text{sign}(\eta_1)v_1\| = O(q^k)$.
 - Falls $\lambda_1 < 0$, gilt $\|(-1)^{k+1}y^k - \text{sign}(\eta_1)v_1\| = O(q^k)$.
-
- Normierung um overflow zu vermeiden.
 - Betrag und Vorzeichen von λ_1 können bestimmt werden.
 - Keine Garantie $\eta_1 \neq 0$ bei der Wahl von y^0 , aber praktisch führt Rundung und zufälliger Startvektor dazu.

INVERSE ITERATION

Ist der betragskleinste Eigenwert $\lambda_n \neq 0$ (und eindeutig), kann man A^{-1} iterieren. Denn die Eigenwerte λ_i^{-1} von A^{-1} sind gegeben durch die Eigenwerte λ_i von A und haben die gleichen Eigenvektoren. Sie erfüllen

$$|\lambda_n^{-1}| > |\lambda_{n-1}^{-1}| \geq \dots > |\lambda_1^{-1}|,$$

so dass die inverse Iteration λ_n^{-1} und v_n approximiert.

GEBROCHENE ITERATION

Anstatt A iteriert man hier $(A - \lambda \mathbb{I})^{-1}$ für $\lambda \notin \sigma(A)$. Diese hat die Eigenwerte $(\lambda_i - \lambda)^{-1}$ und die Iteration approximiert den Eigenvektor v_i von A , der zum Eigenwert λ_i von A gehört, der am nächsten an λ liegt.

Die gebrochene Iteration konvergiert um so schneller, je kleiner der Abstand $|\lambda - \lambda_i|$ ist. Daher kann man versuchen die Schätzung λ in jedem Schritt zu verbessern.

ALGORITHMUS

Gegeben eine Approximation $(\mu^{(0)}, y^0)$ mit $\|y^0\|_2 = 1$. Für $k = 0, 1, 2, \dots$ berechne

$$\mu^{(k+1)} = (y^k)^H A y^k, \quad \tilde{y}^{k+1} = (\mu^{(k+1)} \mathbb{I} - A)^{-1} y^k, \quad y^{k+1} = \frac{\tilde{y}^{k+1}}{\|\tilde{y}^{k+1}\|_2}.$$

$\mu^{(k)}$ approximiert einen Eigenwert von A und y^k den zugehörigen Eigenvektor

ALGORITHMUS

Gegeben eine Matrix $A \in \mathbb{R}^{n \times n}$, definiere $A_0 := A$. Für $k = 0, 1, \dots$ berechne

$$A_k = Q_k R_k, \quad A_{k+1} := R_k Q_k$$

A_k ist eine Approximation an die Schur-Zerlegung von A

EIGENSCHAFTEN

- $A_{k+1} = Q_k^H A_k Q_k$
- $A_{k+1} = (Q_0 \cdots Q_k)^H A (Q_0 \cdots Q_k)$
- $A^{k+1} = (Q_0 \cdots Q_k)(R_k \cdots R_0)$
- Alle Matrizen A_k sind ähnlich zueinander.

Es gilt nach oben

$$A^{k+1} = (Q_0 \cdots Q_k)(R_k \cdots R_0) =: \mathbf{Q}_k \mathbf{R}_k,$$

d.h. Potenzen von A lassen sich als $\mathbf{Q}_k \mathbf{R}_k$ schreiben (eine QR-Zerlegung von A^{k+1}). Betrachte die erste Spalte

$$A^{k+1} e_1 = \mathbf{Q}_k \mathbf{R}_k e_1 = \mathbf{Q}_k r_{1,1}^{(k)} e_1 = r_{1,1}^{(k)} \mathbf{Q}_k e_1 = r_{1,1}^{(k)} \mathbf{q}_1^{(k)},$$

denn \mathbf{R}_k ist obere Dreiecksmatrix. Die Vektoriteration lässt uns erwarten, dass $\mathbf{q}_1^{(k)}$ eine Näherung an den Eigenvektor v_1 von A zum dominanten Eigenwert λ_1 ist. Nach oben gilt auch

$$A_{k+1} = (Q_0 \cdots Q_k)^H A (Q_0 \cdots Q_k) = \mathbf{Q}_k^H A \mathbf{Q}_k$$

und somit (\mathbf{Q}_k ist unitär/orthogonal)

$$A_{k+1} e_1 = \mathbf{Q}_k^H A \mathbf{q}_1^{(k)} \approx \lambda_1 \mathbf{Q}_k^H \mathbf{q}_1^{(k)} = \lambda_1 e_1.$$

FUNKTIONSWEISE DES QR-VERFAHRENS

Ist A invertierbar, so gilt auch

$$A^{(k+1)} = \mathbf{Q}_k \mathbf{R}_k \Rightarrow \mathbf{Q}_k^H = \mathbf{R}_k A^{-(k+1)}.$$

Betrachten wir nun die letzte Spalte bzw. e_n

$$\mathbf{q}_n^{(k)H} = e_n^H \mathbf{Q}_k^H = e_n^H \mathbf{R}_k A^{-(k+1)} = r_{n,n}^{(k)} e_n^H A^{-(k+1)}$$

Analog zu oben erwarten wir jetzt

$$e_n^H A_{k+1} = e_n^H \mathbf{Q}_k^H A \mathbf{Q}_k \approx \lambda_n \mathbf{q}_n^{(k)H} \mathbf{Q}_k = \lambda_n e_n.$$

Insgesamt also

$$A_{k+1} \approx \begin{pmatrix} \lambda_1 & * & & & \\ 0 & * & & & \\ \vdots & * & & & \\ 0 & * & & & \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

DIAGONALISIERBARE MATRIZEN

Sei A diagonalisierbar und habe betragsmässig verschiedene Eigenwerte $|\lambda_i| \neq |\lambda_j|$ und V bezeichne die Matrix der Eigenvektoren $V = [v_1, v_2, \dots, v_n]$, d.h. $A = VDV^{-1}$ mit $D = \text{diag}(\lambda_i)$. Existiert nun eine LR -Zerlegung von V^{-1} , dann sind die Matrizen A_k des QR -Verfahrens (von oben/d.h. ohne Shift) asymptotisch rechte obere Dreiecksmatrizen und ihre Diagonale konvergiert mindestens linear gegen D .

IMPLEMENTIERUNG DES QR -VERFAHRENS

- Konvergenzgeschwindigkeit $\frac{|\lambda_{i+1}|}{|\lambda_i|} \leq 1$
- Jede Iteration erfordert
 - Berechnung einer QR -Zerlegung
 - Produktberechnung RQ
- Im allgemeinen kein effizientes Verfahren!

EFFIZIENTERE FORMULIERUNG DES QR -VERFAHRENS

- Optimierung der Konvergenzgeschwindigkeit
- Reduktion der Kosten einer Iteration

TRANSFORMATION AUF HESSENBERGFORM

Jede Matrix $A \in \mathbb{R}^{n \times n}$ ist ähnlich zu einer Matrix mit oberer Hessenberggestalt. Diese kann mittels Householder-Transformationen bestimmt werden, d.h. $Q^H A Q = H$ mit orthogonalem Q .

$$\text{Kosten} \approx \frac{5}{3}n^3$$

Beispiel

$$A = \begin{pmatrix} 1 & 15 & -6 & 0 \\ 1 & 7 & 3 & 12 \\ 2 & -7 & -3 & 0 \\ 2 & -28 & 15 & 3 \end{pmatrix}$$

IMPLEMENTIERUNG DES QR-VERFAHRENS

QR-ZERLEGUNG EINER HESSENBURG-MATRIX

Mittels Givens-Rotationen kann man die QR-Zerlegung einer Hessenberg-Matrix A mit $O(n^2)$ Operationen bestimmen. Falls A symmetrisch ist, dann sogar mit $O(n)$ Operationen.

ITERATIONSSCHRITT DES QR-VERFAHRENS

Sei A eine obere Hessenberg-Matrix, dann ist mit

$$A = QR \quad \text{auch die Matrix} \quad B := RQ$$

eine obere Hessenberg-Matrix.

QR-VERFAHREN

- Transformiere A auf obere Hessenberg-Form $O(n^3)$
- Alle Iterierten A_k behalten obere Hessenberg-Form
- Bestimme QR-Zerlegung von A_k $O(n^2)$
- Konvergenzgeschwindigkeit über $a_{i+1,i} \rightarrow 0$ ablesbar

SPEKTRALVERSCHIEBUNG/SHIFT

- Liegen $|\lambda_j|$ und $|\lambda_{j+1}|$ nahe aneinander, konvergiert das Verfahren sehr langsam. Mittels einer Spektralverschiebung kann man die Konvergenz beschleunigen.
- Idee: Schiebe die Eigenwerte näher an die 0, so dass der Quotient der verschobenen Eigenwerte wieder groß ist.
- Angenommen wir haben $\mu \approx \lambda_i$, so dass

$$|\mu - \lambda_i| \ll |\mu - \lambda_j| \quad \text{für alle } j \neq i$$

dann konvergiert das Verfahren für $A - \mu\mathbb{I}$ schneller als für A

QR-VERFAHREN MIT SHIFT

Sei A eine (nicht-reduzierbare) Hessenberg-Matrix und setze $A_0 = A$. Iteriere für $k = 1, \dots$:

- Bestimme $\mu_{k-1} \in \mathbb{R}$
- Berechne $A_{k-1} - \mu_{k-1}\mathbb{I} = QR$
- Setze $A_k := RQ + \mu_{k-1}\mathbb{I}$

BESTIMMUNG DER SHIFTS

Für $\mu \approx \lambda_i$ konvergiert der betragskleinste Eigenwert von $A - \mu\mathbb{I}$ am schnellsten. Die Eigenwerte sind auf der Diagonale sortiert, also konvergiert die letzte Zeile am schnellsten. Eine geeignete Wahl für μ ist also $a_{n,n}^{(k)}$ der letzte Diagonaleintrag von A_k .

KONVERGENZGESCHWINDIGKEIT

- Transformation auf Hessenbergform in $O(n^3)$.
- QR-Verfahren mit Shift konvergiert in der Regel quadratisch, für hermitesche Matrizen sogar kubisch.
- Hinreichend genaue Approximation eines Eigenwerts mit konstanter Zahl Iterationen möglich.
- Für alle Eigenwerte also $O(n)$ Iterationen notwendig.
- Jede Iteration kostet $O(n^2)$ (QR-Zerlegung für Hessenberg) bzw. $O(n)$ (QR-Zerlegung für Tridiagonal).
- Insgesamt also $O(n^3)$.

DEFLATION

Das letzte Diagonalelement konvergiert also schnell gegen den Eigenwert und mit der gleichen Rate geht das Subdiagonalelement gegen 0.

$$A_k \rightarrow \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ & \ddots & \ddots & \vdots & \vdots \\ & & * & * & * \\ 0 & \dots & 0 & 0 & \lambda_n \end{pmatrix} = \begin{pmatrix} B & * \\ 0 & \lambda_n \end{pmatrix}$$

Wobei B auch eine Hessenberg-Matrix ist, wir können also mit kleinerem B weitermachen.

DÜNNBESETZTE SYMMETRISCHE MATRIZEN

- QR-Verfahrens für eine allgemeine Matrix $O(n^3)$
- Potenzmethode benötigt nur Matrix-Vektor-Produkt
- Gebrochene Iteration erfordert $(A - \lambda\mathbb{I})^{-1}$ pro Iteration

RAYLEIGH-QUOTIENT

Für eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ mit den Eigenwerte $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{n-1} \geq \lambda_n$ gelten für den Rayleigh-Quotienten

$$RQ_A(x) = \frac{x^H A x}{x^H x}$$

die Identitäten

$$\lambda_1 = \sup_{x \in \mathbb{R}^n \setminus \{0\}} RQ_A(x), \quad \lambda_n = \inf_{x \in \mathbb{R}^n \setminus \{0\}} RQ_A(x),$$

und

$$\lambda_k = \max_{V \subset \mathbb{R}^n, \dim(V)=k} \inf_{x \in V \setminus \{0\}} RQ_A(x).$$

POTENZMETHODE, RAYLEIGH-QUOTIENT, KRYLOVRAUM

Idee: Schöpfe \mathbb{R}^n durch Räume V_j mit $\dim(V_j) = j$ so aus, dass die Eigenwerte immer besser approximiert werden.

- Folge der Iterierten in Potenzmethode y^k .
- Folge der Eigenwertapproximationen $\mu^{(k)} = (y^{k-1})^H A y^{k-1}$.
- Es gilt $\mu^{(k)} \rightarrow \lambda_1$ und $\mu^{(k)} \leq \lambda_1$. (Wertebereich)
- Es gilt auch

$$(y^k)^H A y^k = \mu^{(k)} \leq \mu_1^{(k)} := \max_{y \in Y_k} \frac{y^H A y}{y^H y}$$

mit

$$Y_k := \text{span}\langle y^0, y^1, y^2, \dots, y^{k-1} \rangle$$

- Somit $\mu_1^{(k)}$ mindestens eine sogute Approximation wie $\mu^{(k)}$.
- Kann man $\mu_1^{(k)}$ irgendwie schnell bestimmen?
- Es gilt: $Y_k = \mathcal{K}(A, y^0)$ ist ein Krylovraum

$$\text{span}\langle y^0, y^1, y^2, \dots, y^{k-1} \rangle = \text{span}\langle y^0, A y^0, \dots, A^{k-1} y^0 \rangle = \mathcal{K}(A, y^0)$$

- Seien v_i mit $i = 1, \dots, k$ eine Orthonormalbasis von Y_k .
- Definiere die Matrix $V_k = [v_1, \dots, v_k]$.
- Es gilt $y \in Y_k$ hat eine Darstellung als $y = V_k z$ mit $z \in \mathbb{R}^k$.
- Somit gilt für den Rayleigh-Quotient zu $y \in Y_k$

$$\frac{y^H A y}{y^H y} = \frac{z^H V_k^H A V_k z}{z^H V_k^H V_k z} = \frac{z^H V_k^H A V_k z}{z^H z}$$

- Wir erhalten also

$$\mu_1^{(k)} = \max_{z \in \mathbb{R}^k, z \neq 0} \frac{z^H V_k^H A V_k z}{z^H z} = \lambda_{\max}(V_k^H A V_k)$$

- Auch die weiteren Eigenwerte von $V_k^H A V_k$ können als Näherungen an bestimmte Eigenwerte von A herangezogen werden.

- Orthogonalprojektor

$$P_k := V_k V_k^H : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

auf $\mathcal{K}(A, y^0)$.

- Es gilt die Orthogonalprojektion von A

$$A_k := P_k A|_{\mathcal{K}(A, y^0)}$$

wird durch $V_k^H A V_k$ repräsentiert, denn es gilt die Äquivalenz

$$A_k v_j = \sum_{i=1}^k \alpha_i v_i \Leftrightarrow (V_k^H A V_k) e_j = \sum_{i=1}^k \alpha_i e_i$$

- Je größer der Unterraum $\mathcal{K}(A, y^0)$ ist, um so besser approximiert A_k die Matrix A bzw. um so besser approximieren die Eigenwerte von A_k die von A .

MONOTONIE

Seien $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{n-1} \geq \lambda_n$ und $\mu_1^{(k)} \geq \mu_2^{(k)} \geq \dots \geq \mu_k^{(k)}$ die Eigenwerte der symmetrischen Matrix A bzw. $V_k^H A V_k$ für $k = 1, 2, \dots, n$. Dann gelten für $1 \leq j \leq k$ die Ungleichungsketten

$$\lambda_{n-j+1} \leq \mu_{k+1-j+1}^{(k+1)} \leq \mu_{k-j+1}^{(k)}, \quad \mu_j^{(k)} \leq \mu_j^{(k+1)} \leq \lambda_j$$

- Mit wachsendem k fällt der j -kleinste Eigenwert von $V_k^H A V_k$ monoton von oben gegen den j -kleinsten Eigenwert von A .
- Mit wachsendem k wächst der j -größte Eigenwert von $V_k^H A V_k$ monoton von unten gegen den j -größten Eigenwert von A .

LANZOS PROZESS

Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch und $v_1 = z^{(0)}$ 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

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

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

TRIDIAGONALSYSTEM

Sei $V_k = [v_1, v_2, \dots, v_k]$ mit den Vektoren v_i des Lanczos-Prozesses für $i = 1, \dots, k$. Dann gilt mit α_j und β_j aus dem Lanczos-Prozess

$$T_k := V_k^H A V_k = \begin{pmatrix} \alpha_1 & \beta_1 & 0 & \cdots \\ \beta_1 & \alpha_2 & \ddots & 0 \\ 0 & \ddots & \ddots & \beta_{k-1} \\ 0 & \cdots & \beta_{k-1} & \alpha_k \end{pmatrix}$$

A-POSTERIORI ABSCHÄTZUNG

Sei (μ, w) ein Eigenpaar von T_k mit normiertem w und w_k bezeichne die letzte Komponente von w . Dann besitzt A einen Eigenwert λ , so dass für β_k aus dem Lanczos-Prozess

$$|\lambda - \mu| \leq \beta_k |w_k|$$

- Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch, groß und dünnbesetzt.
- Wähle einen Startvektor $y^{(0)} \in \mathbb{R}^n$.
- Setze $k = 0$, $T_0 \in \mathbb{R}^{k \times k}$, $v_0 = 0$, und $r_1 = y^{(0)}$.
- Iteriere für $k = 1, 2, \dots$
 - Setze $\beta_{k-1} = \|r_{k-1}\|$, $v_k = \frac{r_k}{\beta_{k-1}}$, $\alpha_k = v_k^H A v_k$.
 - Setze $r_{k+1} = (A - \alpha_k \mathbb{I})v_k - \beta_{k-1}v_{k-1}$ und

$$T_k = \begin{pmatrix} T_{k-1} & \beta_{k-1} e_{k-1} \\ \beta_{k-1} e_{k-1}^H & \alpha_k \end{pmatrix}$$

wobei $T_1 = (\alpha_1)$ gilt.

- Bestimme die Eigenwerte von T_k (und die Eigenvektoren).