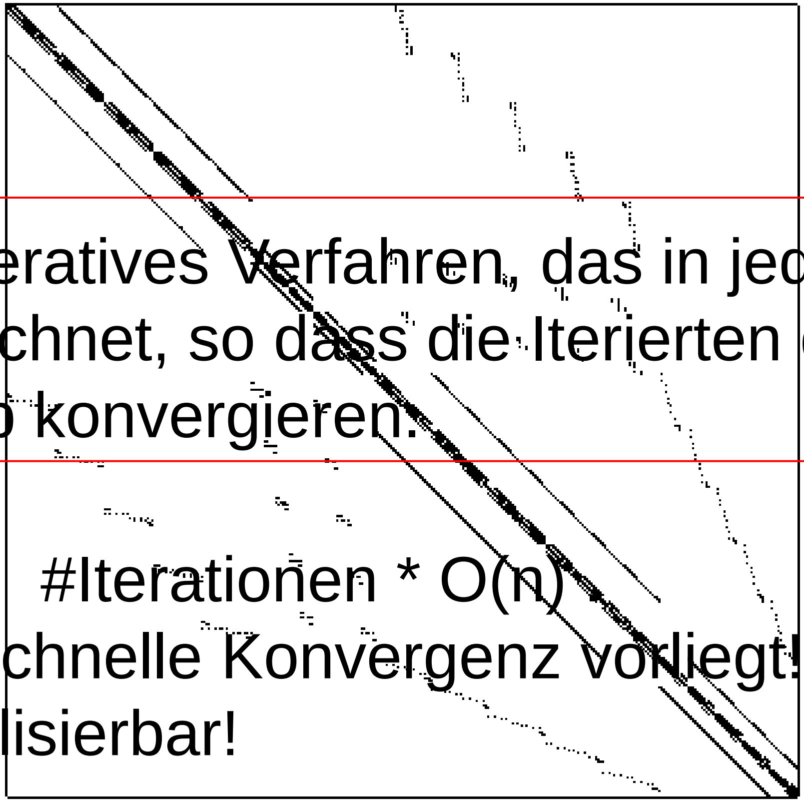
6.3. Iterative Lösung linearer Gleichungssysteme

Großes lineares dünnbesetztes Gleichungssystem $Ax = b$

Gauss-Elimination nutzt in der Regel die Dünnbesetztheit nicht aus und führt meist auf Kosten $O(n^3)$;
Im Gegensatz dazu ist oft nur $O(n)$ Speicherbedarf für Matrix A



Idee: Formuliere iteratives Verfahren, das in jedem Schritt nur Matrix*Vektor berechnet, so dass die Iterierten gegen die Lösung von $Ax = b$ konvergieren.

Gesamtkosten: #Iterationen * $O(n)$
Effizient, falls schnelle Konvergenz vorliegt!
Einfach parallelisierbar!

6.3.1. Stationäre Methoden:

Richardson-Verfahren:

Formulierung eines passenden Fixpunktproblems

$$b = Ax = (A - I)x + x \Leftrightarrow \boxed{x = b + (I - A)x =: \Phi(x)}$$

Daraus ergibt sich die Iteration:

$$x_{k+1} = b + (I - A)x_k = x_k + (b - Ax_k) = x_k + r_k = \Phi(x_k)$$

mit Residuum r_k .

Frage:

Wann konvergiert die ausgehend von einem x_0 definierte Folge x_k gegen die gesuchte Lösung $\bar{x} = A^{-1}b$?

Betrachte dazu den Fehler im k-ten Schritt:

$$\bar{x} - x_{k+1} = \bar{x} - x_k - (b - Ax_k) = \bar{x} - x_k - (A\bar{x} - Ax_k) = (I - A)(\bar{x} - x_k)$$

Ergibt in der Norm $\|\bar{x} - x_{k+1}\|_2 = \|(I - A)(\bar{x} - x_k)\|_2 \leq \|I - A\|_2 \cdot \|\bar{x} - x_k\|_2$

und daher $\|\bar{x} - x_k\|_2 \leq \|I - A\|_2^k \cdot \|\bar{x} - x_0\|_2$

Daher liegt Konvergenz vor für $\| I - A \|_2 < 1$.

Dies entspricht der Kontraktionsbedingung für die Iterations-funktion $\Phi(x)$:

$$\|\Phi(x) - \Phi(y)\|_2 \leq \|I - A\|_2 \cdot \|x - y\|_2$$

Richardsonverfahren ist daher nur sinnvoll, wenn $A \approx I$!

Geschickter:

Wende das Verfahren auf das modifizierte Problem an:

$$D = diag(A); \qquad \boxed{\left(D^{-1}A\right)x = \left(D^{-1}b\right)} \qquad \Leftrightarrow \qquad Ax = b$$

Die Bedingung $D^{-1}A \approx I$ ist besser erfüllt!

Bezeichnung: $\tilde{A} = D^{-1}A$ und $\tilde{b} = D^{-1}b$

ergibt neues Gleichungssystem $\tilde{A}x = \tilde{b}$

Richardson für das tilde-System liefert dann:

$$x_{k+1} = x_k + (\tilde{b} - \tilde{A}x_k) = x_k + D^{-1}(b - Ax_k)$$ oder

$$Dx_{k+1} = Dx_k + (b - Ax_k) = b + (D - A)x_k$$

Dies ist das Jacobiverfahren zur iterativen Lösung von Ax = b

Wesentliche Kosten in jedem Schritt: Ax_k

Konvergent, falls $\left\| I - D^{-1} A \right\|_2 < 1$

Notwendig: Diagonalmatrix D regulär!

Allgemeinere Idee zur Formulierung einer Iterationsfunktion:

Matrix-splitting

$$L: \quad Unterdiagonalteil \quad von \quad A$$

$$U: \quad Oberdiagonalteil \quad von \quad A$$

$$b = Ax = (L + D + U)x = (L + D)x + Ux$$

führt auf Iteration

$$(L + D)x_{k+1} = b - Ux_k = b - (A - L - D)x_k$$
$$= (L + D)x_k + (b - Ax_k)$$

Dies ist das Gauss-Seidel-Verfahren

$$x_{k+1} = x_k + (L+D)^{-1}(b - Ax_k)$$
$$= (L+D)^{-1} b + \left(I - (L+D)^{-1} A\right) x_k$$

In jedem Schritt ist dabei nur ein Dreiecksgleichungssystem zu lösen.

Das Verfahren ist konvergent, falls $\left\| I - (L+D)^{-1} A \right\|_2 < 1$

Gauss-Seidel ist äquivalent zu Richardson angewendet auf

$$\left((L+D)^{-1} A\right) x = \left((L+D)^{-1} b\right) \quad \Leftrightarrow \quad Ax = b$$

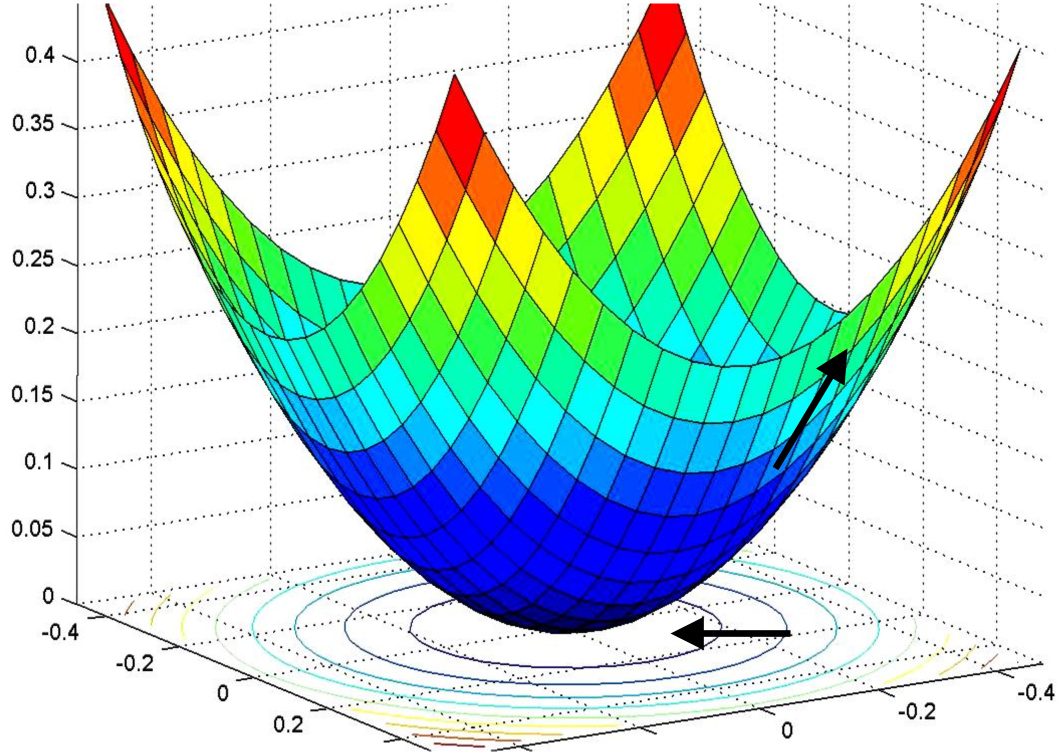
NEU: Komponentenweise Formel für Jacobi und GS!

6.3.2. Nichtstationäre Verfahren

Das Gradientenverfahren für symmetrisch positiv definite Matrix A:

Betrachte Funktion F(x): $R^n \rightarrow R$, $$F(x) = \frac{1}{2} x^T A x - b^T x$$

F beschreibt einen Paraboloid im R^n:



x-Ebene

-grad(F)

Minimum dieser Funktion ist wieder der Punkt mit waagrechter Tangente, also die Stelle mit Gradient gleich Null:

$$\boxed{\nabla F(x) = Ax - b = 0 \Leftrightarrow Ax = b}$$

Stelle x, an der der Paraboloid sein Minimum annimmt ist gleich der gesuchten Lösung des Gleichungssystems!

Betrachte Minimierungsaufgabe!

Von aktueller Stelle x_k aus soll die nächste Iterierte x_{k+1} so gewählt werden, dass sie näher am Minimum liegt.

$$x_{k+1} = x_k + \alpha_k d_k$$

mit Suchrichtung d_k und Schrittweite α_k.
Finde Suchrichtung so, dass Funktionswerte kleiner werden:

Abstiegsrichung ist gegeben durch Richtung des negativen Gradienten!

Denn Richtungsableitung in Richtung n ist gleich $\nabla F \cdot n$, und wird am betragsgrößten für $n = \nabla F$ nämlich $\left\|\nabla F\right\|^2$

Daher ist $n = -\nabla F$ lokal die Richtung des steilsten Abstiegs zum Minimum.

Daher verkleinern sich die Funktionswerte auf jeden Fall, wenn man in diese Abstiegsrichtung geht.

Also wähle $$\boxed{x_{k+1} = x_k - \alpha_k \nabla F(x_k) = x_k + \alpha_k (b - A x_k)}$$

Noch zu bestimmten ist optimale Schrittweite α_k, die am nähesten ans Minimum führt!

Betrachte dazu ein-dimensionale Minimierung:

$$\min_\alpha g(\alpha) := \min_\alpha \left(F(x_k + \alpha(b - Ax_k)) \right) =$$

$$\min_\alpha \left(\frac{1}{2}(x_k + \alpha d_k)^T A(x_k + \alpha d_k) - b^T (x_k + \alpha d_k) \right) =$$

$$\min_\alpha \left(\frac{1}{2}\alpha^2 d_k^T A d_k - \alpha d_k^T d_k + \frac{1}{2} x_k^T A x_k - x_k^T b \right)$$

mit Lösung $\alpha_k = d_k^T d_k \big/ d_k^T A d_k$, $d_k := b - Ax_k$

Insgesamt: $$\boxed{x_{k+1} = x_k + \frac{\|b - Ax_k\|_2^2}{(b - Ax_k)^T A(b - Ax_k)} \cdot (b - Ax_k)}$$

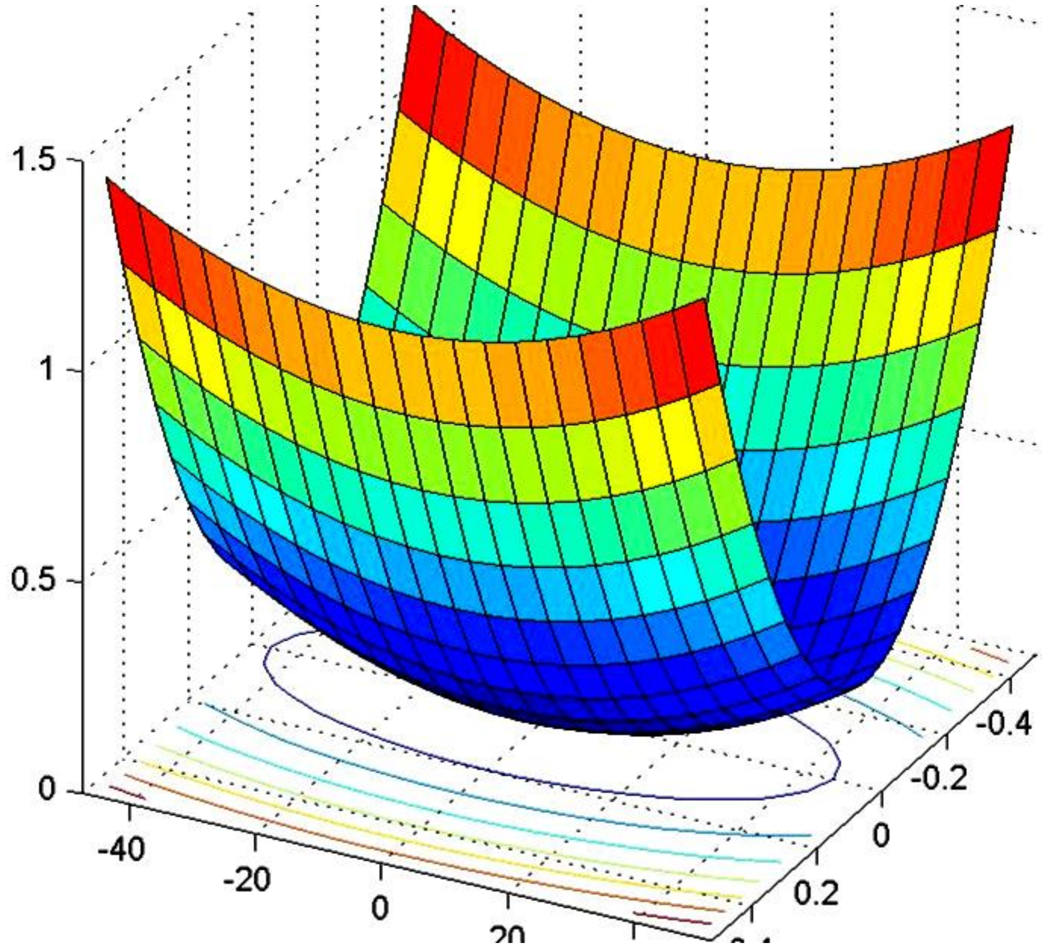
Dazugehörige Fixpunktgleichung ist

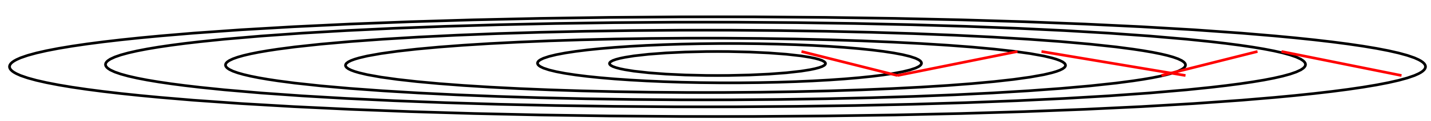
$$x = \Phi(x) := x + \frac{\left\| b - Ax \right\|_2^2}{(b - Ax)^T A (b - Ax)} \cdot (b - Ax)$$

mit Fixpunkt $\bar{x} = A^{-1} b$

Ergibt Verfahren des steilsten Abstiegs (‚steepest descent') oder Gradientenverfahren

Nachteil:
Bei stark verzerrtem Paraboloiden ergibt sich sehr langsame Konvergenz.





Für A ≈ I ist der Paraboloid unverzerrt, die Höhenlinien fast kreisförmig → schnelle Konvergenz!

Daher versucht man, Ax=b zu präkonditionieren:
Ersetze Ax=b durch $M^{-1}Ax = M^{-1}b$ mit $M^{-1}A \approx I$

Bessere Variante eines Gradientenverfahren:
Verfahren der konjugierten Gradienten, kurz cg-Verfahren.

Suchrichtung nicht der negative Gradient selbst, sondern die Projektion des Gradienten, so dass alle Suchrichtungen in gewisser Weise orthogonal zueinander sind
Genauer: Suchrichtungen seien A-konjugiert, d.h.

$$d_k^T A d_j = 0 \qquad \text{für} \quad j \neq k$$

Damit ergibt sich iteratives Verfahren, das nach k Schritten jeweils die beste Näherung an die Lösung in einem k-dimensionalen Unterraum liefert, und daher nach n Schritten fertig ist (in exakter Arithmetik).

NEU: Norm-Minimierung, Unterraum Krylov