Gradientenverfahren & Co

$$\min_{x \in \mathbb{R}^n} f(x) \qquad x^{k+1} = x^k - \alpha_k \nabla f(x^k)$$

z.B. $f(x) = \frac{1}{2} x^T A x - b^T x$ mit A spd

$f'(x) = x^T A - b^T$

Minimum: $f'(\hat{x}) = 0 \iff Ax = b$

Gradientenverfahren: $\nabla f(x) = f'(x)^T = Ax - b$

$$x^{k+1} = x^k - \alpha_k (Ax^k - b)$$

Schrittweitenwahl:

- $\alpha_k = \alpha$ fest wählen

$$\frac{1}{\alpha} x^{k+1} = \frac{1}{\alpha} x^k - \quad (Ax^k - b)$$

$$= \left(\frac{1}{\alpha} I - A\right) x^k + b$$

$M = \frac{1}{\alpha} I$: $\quad M x^{k+1} = (M - A) x^k - b$

(Richardson-Verfahren, einfachstes Matrixzerlegungsverfahren)

- optimale Wahl von α_k.

$$\alpha_k = \arg\min_{\alpha > 0} f(x^k - \alpha \nabla f(x^k))$$

$$f'(x^k - \alpha \nabla f(x^k)) \cdot \nabla f(x^k) = 0 \quad \text{(Kettenregel, Richtungsableitung)}$$

hier: $\nabla f(x^k) = Ax^k - b$

$$\left[(x^k - \alpha \nabla f(x^k))^T A - b^T \right] \cdot (Ax^k - b) = 0$$

$$\Leftrightarrow \left[(x^k - \alpha (Ax^k - b))^T A - b^T \right] \cdot (Ax^k - b) = 0$$

$r^k := Ax^k - b$

$$\Leftrightarrow \left[(x^k - \alpha r^k)^T A - b^T \right] r^k = 0$$

$$\Leftrightarrow \quad \alpha = \frac{r^T r}{r^T A r}$$

also $$x^{k+1} = x^k - \frac{r^{k^T} r^k}{r^{k^T} A r^k} r^k$$

Konvergenz: Monotonie in f

Konvergenzgeschwindigkeit:

$$\| \hat{x} - x^k \|_A \leq \left(\frac{\kappa - 1}{\kappa + 1} \right)^k \| \hat{x} - x^0 \|_A$$

mit $\|x\|_A^2 = x^T A x$

$$\kappa = \kappa(A) = \frac{\lambda_{max}}{\lambda_{min}}$$

Höhenlinien von f

$\|\cdot\|_A = \text{const}$

$\longrightarrow$ Gradientenverfahren ist am besten geeignet für annähernd kreisförmige Höhenlinien

Gradient: Richtung steilsten Anstiegs:

hängt von der Norm ab!

$$\nabla f(x) = \underset{y \neq 0}{\arg\max} \frac{f'(x)\,y}{\|y\|_B} \quad \text{mit } \|y\|_B^2 = y^T B y, \quad B \text{ spd}$$

$$g(y) = \frac{f'(x)\,y}{\sqrt{y^T B y}} \stackrel{!}{=} \max$$

$$g'(y) = 0$$

$$g'(y) = \frac{f'(x)[\,]\sqrt{y^T B y} - f'(x)y \cdot \frac{1}{\cancel{2}\sqrt{y^T B y}} \cdot \cancel{2} y^T B}{y^T B y} = 0$$

$$\Leftrightarrow \quad (y^T B y) \cdot f'(x) - f'(x) y \; y^T B = 0$$

$$\Leftrightarrow \quad y^T B = \beta \cdot f'(x) \qquad \beta = \frac{y^T B y}{f'(x) y}$$

$$\Rightarrow \quad y = B^{-1} f'(x)^T$$

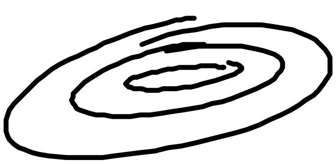
$$= B^{-1}(Ax - b)$$

Gradientenverfahren mit modifizierter Norm

$$x^{k+1} = x^k - \alpha_k B^{-1} f'(x^k)^T$$

$$= x^k - \alpha_k B^{-1}(Ax^k - b)$$

Wahl der Norm B:



Höhenlinien von f

$$f(x) \approx f(\hat{x}) + \overbrace{f'(\hat{x})}^{=0}(x - \hat{x}) + \frac{1}{2} f''(\hat{x})(x - \hat{x})^2 + \ldots$$

$$f(x) = \text{const} \quad \Leftrightarrow \quad (x - \hat{x})^T f''(\hat{x})(x - \hat{x}) = \text{const}$$

$\iff \quad \|x-\hat{x}\|_{f''(\hat{x})} = \text{const}$

$\iff$ Höhenlinien sind ungefähr Kreise in $\|\cdot\|_{f''(\hat{x})}$

Optimale Wahl $B = f''(\hat{x})$

approximativ optimale Wahl: $B = f''(x^k)$

Gradientenverfahren:

$$x^{k+1} = x^k - \alpha_k f''(x^k)^{-1} f'(x^k)^T$$

ist das Newtonverfahren für $f'(x) = 0$ (bei optimaler Wahl der Norm)

Hier: $f(x) = \frac{1}{2} x^T A x - b^T x$

$$f''(x) = A$$

$$x^{k+1} = x^k - \alpha_k A^{-1} (Ax^k - b)$$

$$= x^k - \alpha_k (x^k - \hat{x})$$

exakte Lösung nach einem Schritt bei $\alpha_k = 1$.

nicht realisierbar wegen A^{-1}

aber vielleicht für $B \approx A \longrightarrow$ Vorkonditionierung

---

$$x^{k+1} - x^0 \in \operatorname{span}\{x^k - x^0, r^k\}$$

$$= \operatorname{span}\{x^k - x^0, Ax^k - b\}$$

$$= \operatorname{span}\{x^k - x^0, Ax^k - Ax^0 + Ax^0 - b\}$$

$$= \operatorname{span}\{x^k - x^0, A(x^k - x^0) + r^0\}$$

Mit $K_0 = \operatorname{span}\{r^0\}$

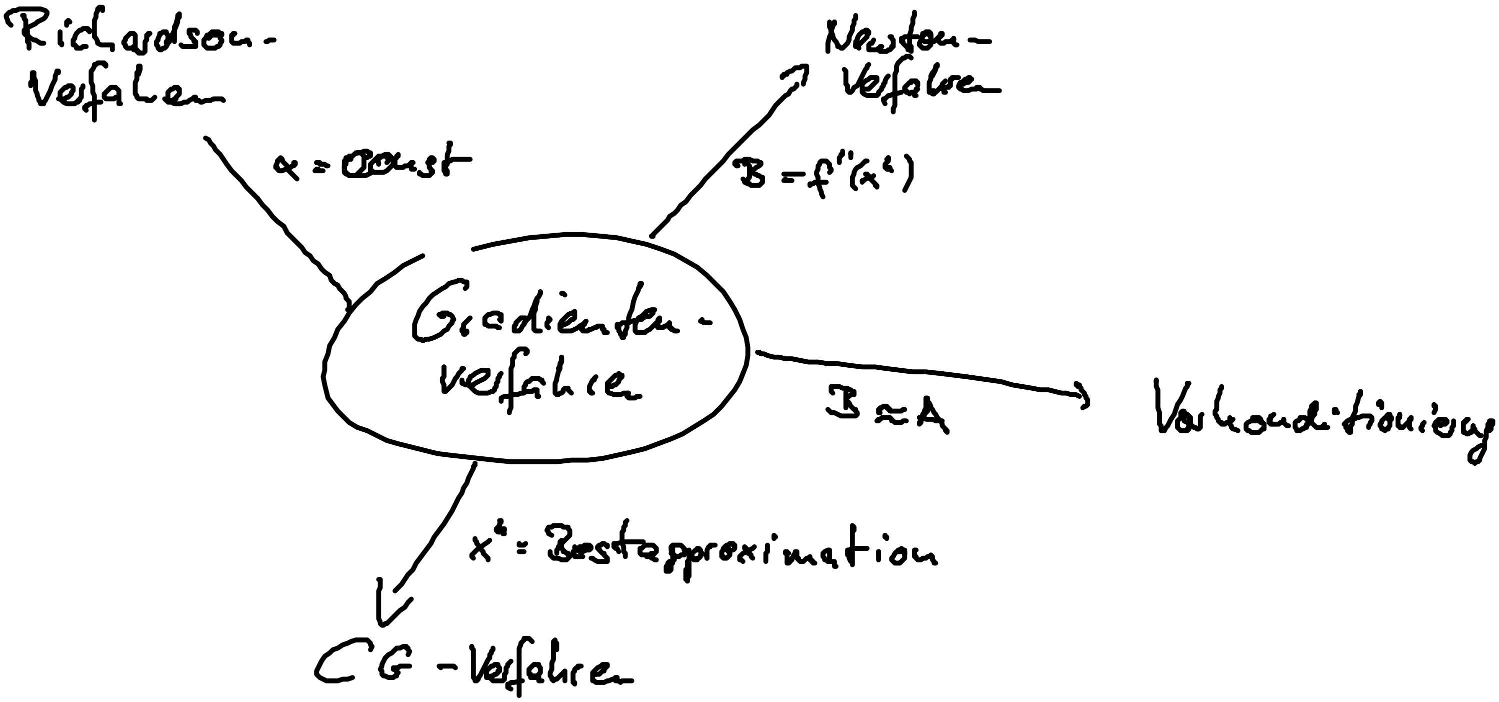
und $K_{k+1} = K_k + A K_k = \{ y = v + Aw \mid v \in K_k, w \in K_k \}$

gilt daher $x^{k+1} - x_0 \in K_k$

optimale Wahl: $x^{k+1} - x^0 = \underset{x \in K_k}{\arg\min} \, \| \hat{x} - (x + x^0) \|_A$

wird vom CG-Verfahren realisiert

Konvergenzgeschwindigkeit: $\| x^k - \hat{x} \|_A \le 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k \| x^0 - \hat{x} \|_A$



Aufwandsabschätzungen

$$A = \begin{bmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}$$

Diskretisierung von $-u'' = b$

$\longrightarrow \kappa(A) \sim n^2$

$\rho(D^{-1}(A - D)) = 1 - \frac{1}{n^2} =: q_J$

Jacobi: $\frac{\| x^k - \hat{x} \|_A}{\| x^0 - \hat{x} \|_A} \le \rho^k$

Gradientenv.: $\le \left(\frac{\kappa - 1}{\kappa + 1} \right)^k =: q_G^k$

$$CG \quad : \quad \le \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k =: q_{CG}$$

$$q_G = \frac{\kappa-1}{\kappa+1} = \frac{1-\frac{1}{\kappa}}{1+\frac{1}{\kappa}} \approx 1-\frac{2}{\kappa} \approx 1-\frac{2}{n^2}$$

$$q_{CG} \approx 1-\frac{2}{n} \quad << q_G$$

Anzahl der Iterationen

$$\frac{\|\hat{x}-x^k\|_A}{\|\hat{x}-x^0\|_A} = TOL$$

$$\Rightarrow \quad \log TOL \approx k \log q$$

$$\Rightarrow \quad k \approx \frac{\log TOL}{\log\left(1-\frac{2}{n^2}\right)} \approx \frac{\log TOL}{-\frac{2}{n^2}} = n^2 \frac{-\log TOL}{2}$$

$$CG: \quad k \approx n \frac{-\log TOL}{2}$$

$$\frac{\#\text{ it Grad}}{\#\text{ it CG}} \approx n$$