

Prekonditionerad CG

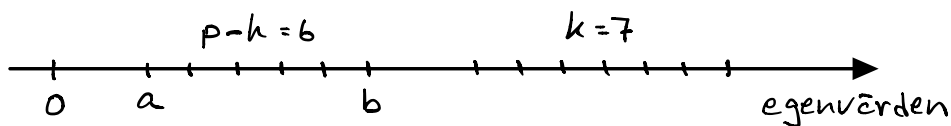
- CG ger konvergens på högst n iterationer
- dock: n mycket stort \rightarrow n iterationer kan vara för beräkningskrävande
- prekonditionering \rightarrow approximativ konvergens på ρ iterationer

$$\text{Problem: } f^* = \min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} x^T Q x - b^T x$$

där Q är symmetrisk och positivt definit, varför f är strikt konvex. Antag att Q har p distinkta egenvärden.

Sets: Antag att Q har $p-k$ distinkta egenvärden inom intervallet $[a, b]$, där $a > 0$, och att resterande k distinkta egenvärden är större än b . För varje x^0 gäller då i CG metoden att

$$f(x^{k+1}) - f^* \leq \left(\frac{b-a}{b+a} \right)^2 (f(x^0) - f^*).$$



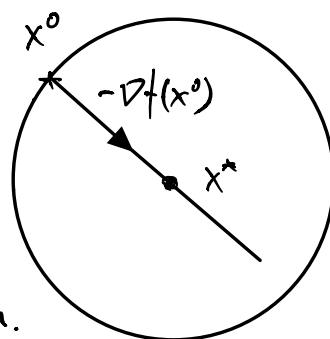
Tolkning: de k första iterationerna optimerar map de k största egenvärdena till Q .

Konsekvens: x^p optimal.

$$[Tg \ a=b=\text{minsta egenvärdet} \Rightarrow p-k=1 \\ \Rightarrow f(x^p) - f^* \leq O \cdot (f(x^0) - f^*) \Rightarrow f(x^p) \leq f^*]$$

Speciellt: alla egenvärden lika, tex om $Q=I$, så får $p=1$, varför x^1 blir optimal.

Alla egenvärden lika
 \Rightarrow sfäriska nivåytor
 $\Rightarrow -\nabla f(x^0)$ pekar
alltid mot optimum.



Ansats: gör en linjär variabeltransformation (basbyte) så att egenvärdestrukturen blir mer fördelaktig!

Målsättning: samla egenvärdena, tex runt ett $\Rightarrow b \approx a \approx 1 \Rightarrow b-a \approx 0$.

Kallas prekonditionering.

Låt $x = Sy$, där $S \in \mathbb{R}^{n \times n}$ är icke-singulär.

Da är

$$\min_{x \in \mathbb{R}^n} f(x) \Leftrightarrow \min_{y \in \mathbb{R}^n} h(y) = f(Sy) = \frac{1}{2} y^T S^T Q S y - b^T S y.$$

Använd CG i y -rummet:

$$y^{k+1} = y^k + \alpha_k \tilde{d}^k$$

$$\tilde{d}^0 = -\nabla_y h(y^0)$$

$$\tilde{d}^k = -\nabla_y h(y^k) + \beta_{k-1} \tilde{d}^{k-1} \quad \text{med} \quad \beta_{k-1} = \frac{\nabla_y h(y^k)^T \nabla_y h(y^k)}{\nabla_y h(y^{k-1})^T \nabla_y h(y^{k-1})}$$

Tolka i x -rummet!

$$\text{Med } x^k = S y^k, \nabla_y h(y^k) = S^T \nabla_x f(x^k) = S^T g^k, d^k = S \tilde{d}^k$$

och $H = S S^T$ för:

$$x^{k+1} = x^k + \alpha_k d^k$$

$$d^0 = S \tilde{d}^0 = S(-\nabla_y h(y^0)) = S(-S^T \nabla_x f(x^0)) = -H g^0$$

$$\begin{aligned} d^k &= S \tilde{d}^k = S(-\nabla_y h(y^k) + \beta_{k-1} \tilde{d}^{k-1}) = -S S^T \nabla_x f(x^k) + \beta_{k-1} S \tilde{d}^{k-1} = \\ &= -H g^k + \beta_{k-1} d^{k-1} \end{aligned}$$

$$\beta_{k-1} = \frac{(S^T \nabla_x f(x^k))^T (S^T \nabla_x f(x^k))}{(S^T \nabla_x f(x^{k-1}))^T (S^T \nabla_x f(x^{k-1}))} = \frac{g^k{}^T H g^k}{g^{k-1}{}^T H g^{k-1}}$$

α_k löser $\min_{\alpha \geq 0} f(x^k + \alpha d^k)$

Riktningarna \tilde{d}^k är konjugerade
mop $S^T Q S$, dvs $\tilde{d}^i{}^T S^T Q S \tilde{d}^j = 0, i \neq j$.
Men $d^k = S \tilde{d}^k$, varför $d^i{}^T Q d^j = 0, i \neq j$,

hos riktningarna d^k är konjugerade
med Q .

Idealt väljs $S = Q^{-1/2} \Rightarrow S^T Q S = I \Rightarrow$
konvergens på en iteration.
Praktiskt eftersträvas $S^T Q S \approx I$.
Finns många sätt och de är ofta
anpassade till strukturer i
specifika tillämpningar.

Varabelbytet görs vanligen inte
explicit, utan beräkningarna
görs fortforande i x -rummet.

Vanlig prekonditionering:
ofullständig Cholesky-faktorisering,
som ger en undertriangulär
matris L sådan att $LL^T \approx Q$.
Med $S = L^{-T}$ får $S^T Q S = L^{-1} Q L^{-T} \approx I$.
Beräkningarna görs i x -rummet
och L ger upphov till framåt-
och bakåtsubstitutioner i
beräkningarna.