Konjugerade gradient metoden

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

där Q är symmetrisk och positivt definit, varför f är strikt konvex.

Initiering: välj $x^0 \in \mathbb{R}^n$ och låt $\beta_{-1} = 0$

Iteration: $x^{k+1} = x^k + \alpha_k d^k$ där $d^k = -g^k + \beta_k d^{k-1}$ med $g^k = Qx^k - b$, $\beta_k = \dfrac{g^{kT} g^k}{g^{k-1\,T} g^{k-1}}$ och

$\alpha_k = -\dfrac{g^{kT} d^k}{d^{kT} Q d^k} > 0$ löser $\min\limits_{\alpha \geq 0} f(x^k + \alpha d^k)$.

Alternativt: $\beta_k = \dfrac{g^{kT}(g^k - g^{k-1})}{g^{k-1\,T} g^{k-1}}$, vilket är

ekvivalent eftersom $g^k \perp g^{k-1}$ gäller.

$\beta_{-1} = 0 \Rightarrow d^0 = -g^0$, dvs första iterationen är brantaste lutning.
Därefter korrigeras brantaste-lutningsriktningen med den förra sökriktningen.

Observera att varje sökriktning därför innehåller komponenter från alla tidigare sökriktningar.

För icke-kvadratiska $f(x)$:

Steglängden α_k bestäms genom den exakta linjesökningen $\min_{\alpha \geq 0} f(x^k + \alpha d^k)$. Denna görs vanligen numeriskt och kan vara beräkningskrävande.

$$d^0 = -\nabla f(x^0)$$

$$\beta_{k-1} = \frac{\nabla f(x^k)^T \nabla f(x^k)}{\nabla f(x^{k-1})^T \nabla f(x^{k-1})} \quad \left[= \frac{\|\nabla f(x^k)\|^2}{\|\nabla f(x^{k-1})\|^2} \right]$$

Kallas Fletcher-Reeves.

Alternativt: $\beta_{k-1} = \dfrac{\nabla f(x^k)^T \left(\nabla f(x^k) - \nabla f(x^{k-1})\right)}{\nabla f(x^{k-1})^T \nabla f(x^{k-1})}$

Kallas Polak-Ribière. Obs: är inte ekvivalent då $f(x)$ är icke-kvadratisk. Anses vara numeriskt mer robust.

Konjugerade gradient metoder kräver:

- exakta linjesökningar, ty annars tappas konjugeringen hos sökriktningarna
- omstart efter n iterationer med $d^n = -\nabla f(x^n)$, ty högst n konjugerade riktningar kan skapas

En stor fördel med konjugerade gradient metoder är att de är mycket minnessnåla, och därför kan användas på mycket stora problem.