

$$d_k = \sum_{i=0}^{k-1} \alpha_i^k \nabla F(u_i) + \nabla F(u_k)$$

$$\frac{\|\nabla F(u_k)\|^2}{\|\nabla F(u_{k-1})\|^2}$$

$$= \nabla F(u_k) + \frac{\|\nabla F(u_k)\|^2}{\|\nabla F(u_{k-1})\|^2} \left(\sum_{i=0}^{k-2} \frac{\|\nabla F(u_{k-1})\|^2}{\|\nabla F(u_i)\|^2} \nabla F(u_i) + \nabla F(u_{k-1}) \right)$$

$$d_k = \nabla F(u_k) + \frac{\|\nabla F(u_k)\|^2}{\|\nabla F(u_{k-1})\|^2} d_{k-1} \quad (1)$$

$$f(\rho) = F(u_k - \rho d_k) \quad (\rho \in \mathbb{R}_+)$$

$$= \frac{1}{2} \langle B(u_k - \rho d_k), u_k - \rho d_k \rangle - \langle c, u_k - \rho d_k \rangle$$

the optimal ρ_k is found by solving $f'(\rho) = 0$

$$\rho_k = \frac{\langle \nabla F(u_k), d_k \rangle}{\langle B d_k, d_k \rangle} \quad (2)$$

CG

start u_0 (any)

$$d_0 = \nabla F(u_0) \quad \rho_0 \text{ by (2)}$$

$$u_1 = u_0 - \rho_0 d_0$$

Then $k > 1$

$$\begin{cases} d_k & \text{by (1)} \\ \rho_k & \text{by (2)} \\ u_{k+1} & = u_k - \rho_k d_k \end{cases}$$

$n \times n$

Thm. If B has only r distinct eigenvalues then CG converges at r iterations at most.

Thm. (Estimate of the error)

$$\|u_k - \hat{u}\|_B \leq \left(\frac{\sqrt{\lambda_1/\lambda_n} - 1}{\sqrt{\lambda_1/\lambda_n} + 1} \right)^{2k} \|u_0 - \hat{u}\|_B$$

λ_1/λ_n if all eigenvalues almost the same - fast convergence

Preconditioning:

$$Bu = c$$

Define $v = Cu$ $u = C^{-1}v$
such that $C^{-T}BC^{-1}$ has "good" eigenvalues

$$\text{solve } \underbrace{C^{-T}BC^{-1}}_{=(C^{-1})^T} v = C^{-T}c$$

CG for nonquadratic functions

Fletcher-Reeves CG (proof for conv.)

Polak Ribiere

$$\beta_k = \frac{\langle \nabla F(u_k) - \nabla F(u_{k-1}), \nabla F(u_k) \rangle}{\|\nabla F(u_{k-1})\|^2}$$

faster (no theoretical conv.)

2.2.4 Quasi-Newton methods

? $\nabla F(\hat{u}) = 0$

$G(u_{k+1})$

Newton: $F(u_{k+1}) - F(u_k) \approx \langle \nabla F(u_k), u_{k+1} - u_k \rangle$

$+ \frac{1}{2} \langle \nabla^2 F(u_k) (u_{k+1} - u_k), u_{k+1} - u_k \rangle$

u_{k+1} such that $\nabla G(u_{k+1}) = 0$

$\nabla F(u_k) + \nabla^2 F(u_k) (u_{k+1} - u_k) = 0$

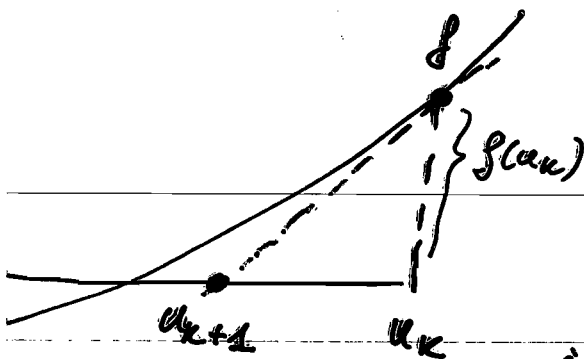
$u_{k+1} = u_k - (\nabla^2 F(u_k))^{-1} \nabla F(u_k)$

Quasi-Newton:

$$u_{k+1} = u_k - H_k^{-1}(u_k) \nabla F(u_k)$$

$H_k^{-1}(u_k) \approx (\nabla^2 F(u_k))^{-1}$

we wish to solve $f(u) = 0 \quad u \in \mathbb{R}$



$\frac{f(u_k)}{u_k - u_{k-1}} = f'(u_k)$

$u_{k+1} = u_k - \frac{f(u_k)}{f'(u_k)}$

identify $f(u)$ with $\nabla F(u)$

For $F(u) = \frac{1}{2} \langle Bu, u \rangle - \langle cu \rangle$

$\nabla F(u) = Bu - c$

$\nabla^2 F(u) = B$

$u_{k+1} = u_k - B^{-1} (Bu_k - c)$

$Bu_{k+1} = c$ then convergence at 1 if !? invert B...

Thm. $F: U \subset V \rightarrow \mathbb{R}$ two times differentiable on a neigh. of U . Suppose \exists constants $\tau > 0$, $M > 0$ and $\gamma \in (0, 1)$ such that

$$B = \overline{B_r(u_0)} \subset U \text{ and}$$

$$(i) \quad \sup_{k \geq 0} \sup_{u \in B} \|H_k^{-1}(u)\| \leq M$$

$$(ii) \quad \sup_{k \geq 0} \sup_{u, v \in B} \|\nabla^2 F(u) - H_k(v)\| \leq \frac{\gamma}{M}$$

$$(iii) \quad \|\nabla F(u_0)\| \leq \frac{\tau}{M} (1 - \gamma)$$

Then

$$\boxed{u_{k+1} = u_k - H_k^{-1}(u_k) \nabla F(u_k)} \quad \text{(*)}$$

$$u_k \in B \quad \forall k$$

$u_k \rightarrow \hat{u}$ which is the unique zero of ∇F on B and

$$\|u_k - \hat{u}\| \leq \frac{\gamma^k}{1 - \gamma} \|u_1 - u_0\|$$

Proof

$$u_{k+1} - u_k = H_k^{-1}(u_k) \nabla F(u_k)$$

$$\Rightarrow \|u_{k+1} - u_k\| \leq M \|\nabla F(u_k)\|$$

$$\|u_1 - u_0\| \stackrel{(iii)}{\leq} M \frac{\tau}{M} (1 - \gamma) < \tau \Rightarrow u_1 \in B$$

Suppose $u_k \in B$, let us show that $u_{k+1} \in B$.

$$\nabla F(u_{k-1}) + H_{k-1}^{-1}(u_{k-1}) (u_k - u_{k-1}) = 0$$

$$\nabla F(u_k) = \nabla F(u_k) - \nabla F(u_{k-1}) - H_{k-1}^{-1}(u_{k-1}) (u_k - u_{k-1})$$

$$|h(x_1) - h(x_0)| \leq \sup |\nabla h(x)| \|x_1 - x_0\|$$

apply to $u \rightarrow \nabla F(u) - H_{k-1}(u_{k-1})u$

$$\|\nabla F(u_k)\| \leq \sup_{u \in B} \|\nabla^2 F(u) - H_{k-1}(u)\| \|u_k - u_{k-1}\|$$
$$\leq \gamma/M$$

$$\leq \frac{\gamma}{M} \|u_k - u_{k-1}\|$$

introduce this in (\rightarrow)

$$\|u_{k+1} - u_k\| \leq M \|\nabla F(u_k)\| \leq \gamma \|u_k - u_{k-1}\|$$
$$\leq \gamma^2 \|u_{k-1} - u_{k-2}\| \dots \leq \gamma^k \|u_1 - u_0\|$$

$$\Rightarrow \|u_{k+1} - u_0\| = \|u_{k+1} - u_k + u_k - u_{k-1} - \dots - u_0\|$$
$$\leq \sum_{i=0}^k \|u_{i+1} - u_i\| \leq \|u_1 - u_0\| \sum_{i=0}^k \gamma^i$$
$$\leq \frac{\|u_1 - u_0\|}{1-\gamma} \Rightarrow u_{k+1} \in B.$$

Existence of a zero in B

$$\|u_{k+j} - u_k\| \leq \sum_{i=0}^{j-1} \|u_{k+i+1} - u_{k+i}\|$$

$$\leq \gamma^k \left(\sum_{i=0}^{j-1} \gamma^i \right) \|u_1 - u_0\| \leq \frac{\gamma^k}{1-\gamma} \|u_1 - u_0\|$$

(u_k) is Cauchy
 B complete $\Rightarrow \exists \hat{u} \in \bar{B}$ such that
 $\lim_{k \rightarrow \infty} u_k = \hat{u}$

$\Rightarrow \nabla F(u)$ continuous

$$\|\nabla F(\hat{u})\| = \lim_{k \rightarrow \infty} \|\nabla F(u_k)\| \leq \frac{\gamma}{M} \lim_{k \rightarrow \infty} \|u_k - u_{k-1}\|$$
$$= 0$$

hence $\nabla F(\hat{u}) = 0$

$$\rightarrow j \rightarrow \infty \quad \|\hat{u} - u_k\| \leq \frac{\gamma^k}{1-\gamma} \|u_1 - u_0\|.$$

Uniqueness

Assume \exists another solution $u^* \in B$

$$u^* - \hat{u} = -H_0^{-1}(u_0) \left(\underbrace{\nabla F(u^*)}_0 - \underbrace{\nabla F(\hat{u})}_0 - H_0(u_0)(u^* - \hat{u}) \right)$$

$$\|u^* - \hat{u}\| \leq \|H_0^{-1}(u_0)\| \sup_{u \in B} \|\nabla^2 F(u) - H_0(u_0)\| \|u^* - \hat{u}\|$$

$$\leq M \frac{\gamma}{M} \|u^* - \hat{u}\| < \|u^* - \hat{u}\|$$

Hence $u^* = \hat{u}$. $\gamma < 1$

The iteration Q-N is general enough.

Proofs under \neq conditions...

Gradient descent corresp. to $H_k = \frac{1}{\rho_k} I$

cite DFP, BFGS - construct an approx of $\nabla^2 F$ based only on the variations of the grad. during the iterations

Generalized Weiszfeld method (1937?)
construct convex quadratic $G(v, u) = \text{approx.}$
from above for $F(u)$:

$$G(v, u) = F(u) + \langle v - u, \nabla F(u) \rangle + \frac{1}{2} \langle v - u, H(u)(v - u) \rangle$$

$$H(u) \text{ continuous } \lambda_{\min}(H(u)) \geq \mu > 0 \quad \forall u$$
$$F(v) \leq G(v, u) \quad \forall v$$

$$u^{k+1} = \arg \min_v G(v, u^k)$$

This min exists and is given by

$$\nabla F(u^k) + H(u^k)(u^{k+1} - u^k) = 0$$

A particular case :

put $\nabla F(u)$ in the form \leftarrow independent of u

$$\nabla F(u) = H(u)u + z = 0$$

$$u^{k+1} = -H(u^k)^{-1}z \dots$$

2.2.5 Half-quadratic regularization HQ

$$F(u) = \|Au - v\|^2 + \beta \sum_i \varphi(\|D_i u\|)$$

$$D_i : \mathbb{R}^p \rightarrow \mathbb{R}^r \text{ e.g. } u_i - u_j \text{ or } \begin{pmatrix} u_{ij} - u_{i,j-1} \\ u_{ij} - u_{i-1,j} \end{pmatrix}$$

$\|\cdot\|$ can be $\|\cdot\|_2$

φ - in image proc. - "edge preserving"

φ' is bounded e.g. $\varphi(t) = \sqrt{t^2 + \alpha}$ $\alpha > 0$

Newton: $u_{k+1} = u_k - H(u_k)^{-1} \nabla F(u_k)$

$$\nabla F(u) = 2A^T(Au - v) + \dots$$

HQ - 2 forms \gg "*" and "+" Geman Reynolds
Geman Yang

introduce an auxiliary variable b

Augmented objective $F(u, b)$ - quadratic in u (for b fixed), separable in b (for u fixed)

Then

$$(\hat{u}, \hat{b}) = \arg \min F(u, b)$$

Alternate minimization

$$\forall k \geq 0 \begin{cases} u_k = \arg \min_u F(u, b_{k-1}) \\ b_k = \arg \min_b F(u_k, b) \end{cases}$$

(a) Multiplicative form

$$F(u, b) = \|Au - v\|^2 + \beta \sum_{i=1}^n \left(\frac{b^{[i]}}{2} \|D_i u\|^2 + \underbrace{\Psi(b^{[i]})}_{?} \right)$$

$t \rightarrow \Psi(\sqrt{t})$ concave + Ψ continuous increasing
 increases less fast than t^2 on \mathbb{R}_+ $\Psi(0) = 0$

Define $\underline{\Theta}(t) = \begin{cases} -\Psi(\sqrt{t}) & t \geq 0 \\ \infty & \text{else} \end{cases} \Rightarrow$ Convex continuous

f convex conjugate is

$$f^*(x) = \sup_u (\langle u, x \rangle - f(u))$$

Thm

f convex lower semi continuous $f \not\equiv \infty$
 then $f^{**}(u) = f(u)$

$$f^{**}(u) = \sup_x (\langle u, x \rangle - f^*(x))$$

Rockafellar
 \Rightarrow convex an.

$$\Theta^*(b) = \sup_{t \geq 0} (bt - \Theta(t)) \quad \begin{matrix} b \in \mathbb{R} \\ b > 0 \Rightarrow \Theta^*(b) = +\infty \end{matrix}$$

Define $\Psi(b) = \Theta^*(-\frac{1}{2}b)$ which means

$$\Psi(b) = \sup_{t \geq 0} \left(\frac{1}{2}bt - \Theta(t) \right) = \sup_{t \geq 0} \left(-\frac{1}{2}bt^2 + \varphi(t) \right)$$

t^2 in place of t
 $\Theta(t^2) = -\varphi(t)$

we have

$$\Theta^{**} = \Theta$$

$$\Theta(t) = \sup_b (bt - \Theta^*(b)) = \Theta^{**}(t)$$

calculate Θ^{**} at t^2 :

$$\begin{aligned} \varphi(t) = \Theta(t^2) &= \sup_{b \leq 0} (bt^2 - \Theta^*(b)) \\ &\quad \uparrow \text{replace by } -\frac{1}{2}b \\ &\quad \uparrow \text{if } b > 0 \quad \Theta^*(b) = +\infty \\ &= \sup_{b \geq 0} \left(-\frac{1}{2}bt^2 - \Psi(b) \right) \end{aligned}$$

$$\varphi(t) = \inf_{b \geq 0} \left(\frac{1}{2}bt^2 + \Psi(b) \right)$$

$$\text{where } \Psi(b) = \sup_{t \geq 0} \left(-\frac{1}{2}bt^2 + \varphi(t) \right)$$

$$\begin{aligned} \inf_{b \in \mathbb{R}^2} F(u, b) &= \|Au - v\|^2 + \beta \sum_{i=1}^2 \inf_{b[i]} \left(\frac{b[i]}{2} \|D_i u\|^2 + \Psi(b[i]) \right) \\ &= F(u) \qquad \qquad \qquad \varphi(\|D_i u\|) \end{aligned}$$

$$\hat{u}, \hat{b} \Rightarrow \hat{u} \text{ minimizes } F(u)$$