$$d_k = \sum_{i=0}^{k-1} \underbrace{\alpha_i^k}_{\frac{\|\nabla F(u_k)\|^2}{\|\nabla F(u_i)\|^2}} \nabla F(u_i) + \nabla F(u_k)$$

α_i^{k-1}

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

ξ_k

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

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

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

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

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

$\boxed{CG}$ ~~d_k~~ start u_0 (any)

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

$u_1 = u_0 - \rho_0 d_0$

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

Thm. If B $n \times n$ has only r distinct eigenvalues then CG converges at r iterations at most.

Thm. (Estimate of the error over)

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

~~X~~ if all eigenvalues almost the same - fast convergence

Preconditioning: $Bu = c$

Define $v = Cu$ $\quad u = C^{-1} v$

such that $C^{-T} B C^{-1}$ has "good" eigenvalues

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

CG for nonquadratic functions

- Fletcher - Reeves CG (proof for conv.)

- Polak Ribière

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

faster (no theoretical conv.)

2.2.4 *Quasi - Newton methods*

? $\nabla F(\hat{u}) = 0$ $\qquad G(u_{k+1})$

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

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

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

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

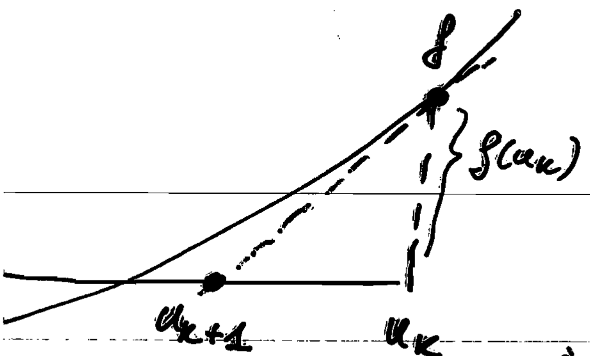
$$u_{k+1} = u_k - (\nabla^2 F(u_k))^{-1} \nabla F(u_k).$$

Quasi-Newton:

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

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

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



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

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

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

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

$\nabla F(u) = Bu - c \qquad \nabla^2 F(u) = B$

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

$B u_{k+1} = c$ ~~then~~ convergence at 1 it. !? invert B...

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

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

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

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

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

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

$u_k \in B \quad \forall k$

$u_k \to \hat{u}$ 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

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

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

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

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

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

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

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

apply to $u \to \nabla F(u) - \mathcal{H}_{K-1}(u_{K-1})\, u$

$$\|\nabla F(u_K)\| \leq \underbrace{\sup_{u \in B} \|\nabla^2 F(u) - \mathcal{H}_{K-1}(u)\|}_{\leq \gamma / M} \|u_K - u_{K-1}\|$$

$$\leq \frac{\gamma}{M} \|u_K - u_{K-1}\|$$

introduce this in $(\rightarrow\bullet)$

$$\|u_{K+1} - u_K\| \leq M \|\nabla F(u_K)\| \leq \gamma \|u_K - u_{K-1}\|$$

$$\leq \gamma^2 \|u_{K-1} - u_{K-2}\| \ldots \leq \gamma^K \|u_1 - u_0\|$$

$$\Rightarrow \|u_{K+1} - u_0\| = \|u_{K+1} - u_K + u_K - u_{K-1} - \ldots - u_0\|$$

$$\leq \sum_{i=0}^{K} \|u_{i+1} - u_i\| \leq \|u_1 - u_0\| \sum_{i=0}^{K} \gamma^i$$

$$\leq \frac{\|u_1 - u_0\|}{1-\gamma} \Rightarrow u_{K+1} \in B.$$

Existence of a zero in B

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

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

$\left.\begin{array}{l} (u_K) \text{ is Cauchy} \\ \bar{B} \text{ complete} \end{array}\right) \Rightarrow \exists\, \hat{u} \in \bar{B}$ such that $\lim_{K\to\infty} u_K = \hat{u}$

$\hookrightarrow \nabla F(u)$ continuous

$$\|\nabla F(\hat{u})\| = \lim_{K\to\infty} \|\nabla F(u_K)\| \leq \frac{\gamma}{M} \underbrace{\lim_{K\to\infty} \|u_K - u_{K-1}\|}_{=0}$$

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

$\hookrightarrow j \to \infty \quad \|\hat{u} - u_K\| \leq \dfrac{\gamma^K}{1-\gamma} \|u_1 - u_0\|.$

Uniqueness

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

$$u^* - \hat{u} = -H_0^{-1}(u_0)\Big(\underbrace{\nabla F(u^*)}_{0} - \underbrace{\nabla F(\hat{u})}_{0} - H_0(u_0)(u^* - \hat{u})\Big)$$

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

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

$$\gamma < 1$$

Hence $u^* = \hat{u}$.

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 $G(v, u)$ = approx. from above for $F(u)$:

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

$H(u)$ continuous $\quad \lambda_{\min}(H(u)) \geq \mu > 0 \quad \forall u$

$F(v) \leq G(v, u) \quad \forall v$

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

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

independent of u

$$\nabla F(u) = H(u)\, u + z \quad \stackrel{?}{=} 0$$

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

2.2.5 **Half-quadratic regularization HQ**

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

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

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

φ - in image proc. - "edge preserving"

$\varphi' \simeq$ bounded eg. $\varphi(t) = \sqrt{t^2 + \alpha}$ $\quad \alpha > 0$

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

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

HQ - 2 forms : * "and "+" Geman Reynolds, Geman Yang

introduce an auxiliary variable b

Augmented objective $\mathcal{F}(u, b)$ - quadratic in u (for b fixed), separable in b (for u fixed) Then

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

Alternate minimization

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

(a) Multiplicative form

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

$t \to \varphi(\sqrt{t})$ $\underline{\text{concave}}$ + φ continuous increasing

increases less fast than t^2 on $\mathbb{R}_+$ $\varphi(0) = 0$

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

f convex conjugate is

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

Thm

$$\boxed{\begin{array}{l} f \text{ convex lower semi continuous } f \not\equiv \infty \\ \text{then } f^{**}(u) = f(u) \end{array}}$$

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

Rockafellard 70 Convex an.

$$\theta^*(b) = \sup_{t \geq 0} \left(bt - \theta(t) \right)$$

$b \in \mathbb{R}$

$b > 0$ $\theta^*(b) = +\infty$

Define $\boxed{\Psi(b) = \theta^*(-\tfrac{1}{2} b)}$ which means

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

t^2 in place of t

$\theta(t^2) = -\varphi(t)$

we have

$$\theta^{**} = \theta \qquad \theta(t) = \sup_{b}\left(bt - \theta^{*}(b)\right) = \theta^{**}(t)$$

Calculate θ^{**} at t^2:

$$\varphi(t) = \theta(t^2) = \sup_{b \le 0}\left(bt^2 - \theta^{*}(b)\right)$$

replace b by $-\frac{1}{2}b$

(if $b>0$ $\theta^{*}(b) = +\infty$)

$$= \sup_{b \ge 0}\left(-\frac{1}{2}bt^2 - \psi(b)\right)$$

$$\boxed{\begin{aligned}\varphi(t) &= \inf_{b \ge 0}\left(\frac{1}{2}bt^2 + \psi(b)\right)\\ \text{where}\quad \psi(b) &= \sup_{t \ge 0}\left(-\frac{1}{2}bt^2 + \varphi(t)\right)\end{aligned}}$$

$$\inf_{b \in \mathbb{R}^2} \mathcal{F}(u,b) = \|Au - v\|^2 + \beta \sum_{i=1}^{r} \underbrace{\inf_{b[i]}\left(\frac{b[i]}{2}\|D_i u\|^2 + \psi(b[i])\right)}_{\varphi(\|D_i u\|)}$$

$$= F(u)$$

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