IDEA:

KEEP THE STEP LENGTH $l_k = \| t_k p^{(k)} \|$

AND COMPUTE X_{k+1} AS:

$$X_{k+1} - X_k = (A_k + \lambda_k I)^{-1} (\nabla f(x^{(k)}))$$

WHERE $\| X_{k+1} - X_k \| = l_k$

IS λ_k RUINING THE SOLUTION ???

THEOREM 5.5.1

LET f HAS CONTINUOUS SECOND PARTIAL DERIVATIVES, $X_k \in \mathbb{R}^n$ AND

$$Q_k(X) = f(X_k) + \nabla f(X_k)^T (X - X_k) + \frac{1}{2} (X - X_k)^T A_k (X - X_k)$$

THERE EXISTS $\lambda_k \geq 0$ SUCH THAT MINIMIZER X_{k+1} OF $Q_k(X)$ SUBJECT TO

$\| X - X_k \| \leq l_k$ IS A SOLUTION TO

$$(A_k + \lambda_k I)(X - X_k) = -\nabla f(X_k)$$

IF $\| -A_k^{-1} (\nabla f(X_k)) \| \leq l_k$, PICK $\lambda_k = 0$, OTHERWISE $\lambda_k > 0$ S.T. $\| X_{k+1} - X_k \| = l_k$

PROOF APPLY KKT

$$(3)\begin{cases} \min Q_k(x) \\ \text{s.t. } \|x - x_k\|^2 - l_k^2 \le 0 \end{cases}$$

$$L(x, \mu_k) = f(x_k) + \nabla f(x_k)^T(x - x_k) + \frac{1}{2}(x-x_k)^T A_k (x - x_k) +$$
$$+ \mu_k (\|x - x_k\|^2 - l_k^2)$$

WHERE μ_k IS KKT MULTIPLIER (λ^* FROM KKT)

x_{k+1} : $L(x^*, \mu_k) \le L(x, \mu_k)$ OR GRADIENT CONDITION

$$0 = \nabla f(x_k) + A_k(x_{k+1} - x_k) + 2\mu_k(x_{k+1} - x_k)$$

SO

$$(A_k + 2\mu_k)(x^{k+1} - x^k) = -\nabla f(x_k)$$

LET $\lambda_k = 2\mu_k$ $\square$

λ_k CAN BE CHOSEN TO PRESERVE OPT. SOLUTION

$\rightarrow$ GOOD APPROXIMATION TO $f(x)$

HOPEFULLY

l_k .. PLACE WE TRUST THE APPROXIMATION IS GOOD.

TRUST REGION METHOD

MIN $f(x)$

- APPROXIMATE $f(x)$ BY

$$Q_k(x) = f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{1}{2}(x - x_k)^T A_k (x - x_k)$$

HOPE THAT $Q_k(x)$ IS A GOOD APROXIMATION OF $f(x)$ FOR $x \in B(x_k, l_k)$

SOLVE x_{k+1} AS

$$\min_{x \in B(x_k, l_k)} Q_k(x)$$

WE ARE HAPPY WITH x_{k+1} IF

FOR FIXED γ

$$f(x_{k+1}) \le f(x_k) + \gamma \big(\underbrace{Q_k(x_{k+1}) - f(x_k)}_{<0} \big)$$

(IF NOT SATISFIED, DECREASE l_k

SIMILAR TO CONDITION 4)

3.4 BROYDEN'S METHOD

(QUASI-NEWTON METHODS)

RECALL NEWTON'S METHOD:

$g : \mathbb{R}^n \to \mathbb{R}^n$

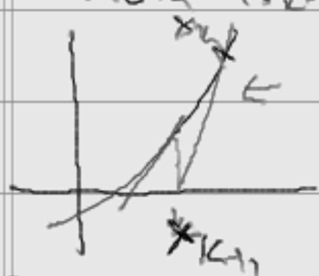
FIND x S.T. $g(x) = 0$

(USED ON ∇f FOR MINIMIZATION)

LINEAR APPROXIMATION:

$L_k(x) = g(X_k) + \nabla g(X_k)(X - X_k)$

PICK X_{k+1} S.T. $L_k(X_{k+1}) = 0$



$\leftarrow$ TANGENTS. $L_k(X)$

[JACOBIAN $\nabla g(X_k) = Hf(X_k)$ - COSTLY TO COMPUTE]

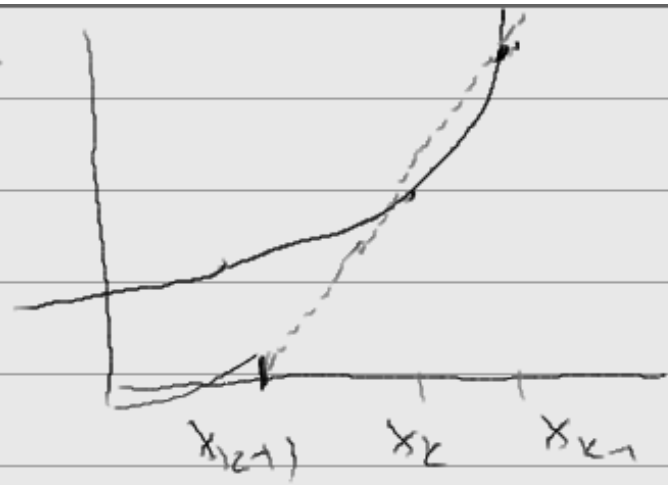
TRY $l_k(X) = g(X_k) + D_k(X - X_k)$

$$X_{k+1} = X_k - D_k^{-1} g(X_k)$$

D_k EASIER TO UPDATE THAN $\nabla g(X_k)$

IDEA IN 2D:

SECANT METHOD



MAKE D_k S.T. $l_k(x_k) = g(x_k)$

$$l_k(x_{k-1}) = g(x_{k-1})$$

SO $D_k(x_k - x_{k-1}) = g(x_k) - g(x_{k-1})$

FIXED IMAGE OF 1 VECTOR $n-1$ FREEDOM

IN ITERATION: x_{k+1} & D_{k+1}

D UPDATE: U_k .. UPDATE MATRIX

$$D_{k+1} = D_k + (D_{k+1} - D_k)$$

RECALL LINEAR ALGEBRA:

$a, b \in \mathbb{R}^n$, TENSOR PRODUCT $a \otimes b$ IS $ab^T \in \mathbb{R}^{n \times n}$

OUTER PRODUCT $\boxed{(a \otimes b)x = (b \cdot x)a}$

$\therefore$ RANK $(a \otimes b) = 1$ IF $a^T b \neq 0$ THEN $a^T b$ IS

THE ONLY NON ZERO EIGENVALUE

Δ POS SEMIDEF

$(a \otimes b)^T = (b \otimes a)$, $\Rightarrow$ $(a \otimes a)$ IS SYMMETRIC

WISH $U_k = a_k \otimes b_k$ $\quad$ D_{k+1} AFTER x_{k+1}

$$D_{k+1} = D_k + a_k \otimes b_k$$

$$D_{k+1}(x_{k+1} - x_k) = g(x_{k+1}) - g(x_k)$$

$$D_k(x_{k+1} - x_k) + a \otimes b\,(x_{k+1} - x_k) = g(x_{k+1}) - g(x_k)$$

$$\left(b_k^T (x_{k+1} - x_k)\right) a_k = g(x_{k+1}) - g(x_k) - D_k(x_{k+1} - x_k)$$

$$a_k = \frac{g(x_{k+1}) - g(x_k) - D_k(x_{k+1} - x_k)}{b_k^T (x_{k+1} - x_k)}$$

↳ PICK OF b_k DETERMINES a_k

INFLUENCE OF b_k

$$l_{k+1}(x) - l_k(x) = g(x_{k+1}) + D_{k+1}(x - x_{k+1}) - g(x_k) - D_k(x - x_k)$$

$$= \quad (x - x_k + x_k - x_{k+1})$$

$$= \underbrace{g(x_{k+1}) - g(x_k) - D_{k+1}(x_{k+1} - x_k)}_{= 0 \text{ BY SECANT CONDITION}} + U_k(x - x_k)$$

$$= \left(b_k^T (x - x_k)\right) a_k$$

IF x LINE $x_k - x_{k+1}$ .. CLEAR

HOW TO UPDATE IF NOT?

FOR $(x - x_k)$ ORTHOGONAL TO $(x_{k+1} - x_k)$

WE WANT $l_{k+1}(x) = l(x)_k$

TO AVOID "UNKNOWN" CHANGES

SO $0 = \ell_{k+1}(x) - \ell_k(x) = (b^T(x - x_k))a_k$

$\forall x$ S.S. $(x - x_k)^T(x_{k+1} - x_k) = 0$

SO PICK

$$b = (x_{k+1} - x_k)$$

AND HENCE $a_k = \dfrac{g(x_{k+1}) - g(x_k) - D_k b_k}{b_k^T b_k}$

METHOD: BROYDEN RANK-ONE UPDATE

$g: \mathbb{R}^n \to \mathbb{R}^n$, x_0 & D_0 INITIALIZED

THEN

$$x_{k+1} = x_k - D_k^{-1} g(x_k)$$

WHERE

1) SOLVE $D_k(x - x_k) = -g(x_k)$

SET x_{k+1} = SOLUTION

2) SET $d_k = (x_{k+1} - x_k)$

3) SET

$$D_{k+1} = D_k + \overbrace{\frac{1}{d_k^T d_k} \cdot (g(x_{k+1}) - g(x_k) - D_k d_k)}^{a_k} \otimes \overbrace{d_k}^{b_k}$$

PICK D_0 -- $\nabla g(x_0)$ OR ANYTHING ELSE YOU LIKE

FEATURE:

A, B MATRICES, DISTANCE:

$$d(A, B) = \max\{\|Ax - Bx\| : \|x\| \leq 1\}$$

THEOREM 3.4.5

D IS $m \times n$ SATISFYING SECANT CONDITION $Dd_k = g(x_{k+1}) - g(x_k)$

THEN

$$d(D_{k+1}, D_k) \leq d(D, D_k)$$