$\Rightarrow \quad a_k = y_k \qquad b_k = D_k d_k$

SOLVE $\alpha_k (y_k^T d_k) = 1$

$\beta_k ((D_k d_k)^T d_k) = -1$

SO

$$D_{k+1} = D_k + \frac{y_k \otimes y_k}{y_k^T d_k} - \frac{(D_k d_k) \otimes (D_k d_k)}{(D_k d_k)^T d_k} \quad (*)$$

THEOREM 35.2

D_k POS. DEF. X_k GIVEN

IF $t_k > 0$ SELECTED SO THAT

$$X_{k+1} = X_k - t_k \underbrace{D_k^{-1} (\nabla f(X_k))}_{= p_k},$$

SATISFY

3) $\exists\ 0 < \beta < 1$ S.T. $p_k^T \nabla f(X_{k+1}) > \beta p_k^T \nabla f(X_k)$

THEN D_{k+1} DEFINED BY $(*)$ IS POSITIVE DEFINITE

PROOF: 3) $\Rightarrow d_k^T y_k > 0 \quad y_k = A d_k$

$D_k = D_k^{1/2} D_k^{1/2}$ AND THEN FROM DEFINITION

$x^T D_{k+1} x \geq 0$?

BFGS METHOD

x_k D_k COMPUTED

1, FIND t_k S.T.

$$x_{k+1} = x_k - t_k D_k^{-1} \nabla f(x_k)$$

SATISFY 1, .. 4,

2, SET

$$D_{k+1} = D_k + \frac{y_k \otimes y_k}{d_k^T y_k} - \frac{D_k d_k \otimes D_k d_k}{d_k^T D_k d_k}$$

WHERE $d_k = x_{k+1} - x_k$

$$y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$$

NOTE:

NICE THEORETICALLY, AS WELL AS IN PRACTICE.

FOR POS DEF: CHOLESKY FACTORIZATION:

$D_{k+1} = L_{k+1} L_{k+1}^T$ $\quad D_k = L_k L_k^T$

LOWER TRIANGULAR

CLAIM: $\|L_{k+1} - L_k\| \le \|L - L_k\|$ $\forall L$ S.T. $LL^T d_k = y_k$

DFP DAVIDON - FLETCHER POWELL

BFGS

TRIES TO APPROXIMATE $H_f^{-1}(x)$ INSTEAD OF $H_f(x)$

RECALL $x_{k+1} = x_k - t_k H_f(x_k)^{-1} \nabla f(x_k)$ IN NEWTON

NOW WE WANT:

$$x_{k+1} = x_k - t_k D_k \nabla f(x_k)$$

SUCH THAT

1) t_k S.T. CRIT 1-4 SATISFIED

1) DESCENT $f(x_{k+1}) < f(x_k)$

2) $p_k^T \nabla f(x_k) < 0$ .. $(D_k \nabla f(x_k))^T \nabla f(x_k) > 0$

WAS WITH $x_{k+1} = x_k + t_k \cdot p_k$

3) t_k NOT TOO SMALL: $0 < \beta < 1$: $p_k^T \nabla f(x_{k+1}) \geq \beta p_k^T \nabla f(x_k)$

4) GOOD IMPROVEMENT: $0 < \alpha < \beta < 1$:

$$f(x_{k+1}) \leq f(x_k) + \alpha t_k p_k^T \nabla f(x_k)$$

2) D_k POSITIVE DEFINITE

3) INVERSE SECANT CONDITION

$$D_{k+1}^{-1} d_k = y_k \Rightarrow D_{k+1} y_k = d_k$$

$$d_k = x_{k+1} - x_k \quad y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$$

4) UPDATE FROM D_k TO D_{k+1} EASY

SIMILAR TO BFGS:

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

$$d_k = D_{k+1} y_k =$$

$$= D_k y_k + \alpha_k (a_k^T y) . a_k + \beta_k (b_k^T y_k) . b_k$$

$$d_k - D_k y_k = \underbrace{\alpha_k (a_k^T y)}_{\overset{!}{=} 1} \underbrace{a_k}_{\overset{!}{=} d_k} + \underbrace{\beta_k (b_k^T y_k)}_{\overset{!}{=} -1} \underbrace{b_k}_{D_k y_k}$$

$a_k = d_k$

$b_k = D_k y_k$

$$\alpha_k = \frac{1}{d^{kT} y^k} \qquad \beta_k = = \frac{-1}{y^{kT} D_k y^k}$$

$$D_{k+1} = D_k + \frac{d^k \otimes d_k}{y^{kT} d_k} - \frac{D_k y_k \otimes D_k y^k}{y^{kT} D_k y_k}$$

↗ SWAPPED ROLES OF y^k & d_k COMPARED TO BFGS

↪ GIVES DFP METHOD $x_{k+1} = x_k - t_k D_k \nabla f(x_k)$ (1959)

[BFGS MORE STABLE (POS DEF D_k), BUT NOT AS EASY)

← BETTER

CONJUGATE DIRECTIONS METHOD CH 9.
(LIN & NONLIN PROG. LUENBERGER & YE)
MIX OF NEWTON & STEEPEST DESC.

POS. DEF.

$\text{MIN } \frac{1}{2}x^T Q x - b^T x$ ← STUDY FOR QUADRATIC

" AROUND SOLUTION, PROGRAMS CAN BE NICELY APPROXIMATED BY QUADRATIC"

DEF:
Q SYMMETRIC $\in \mathbb{R}^{n\times n}$, $d_1\ d_2 \in \mathbb{R}^n$ ARE Q-ORTHOGONAL (OR CONJUGATE WITH RESPECT TO Q) IF $d_1^T Q d_2 = 0$
[Q POSITIVE DEFINITE LATER]
- $I = Q$ .. CLASSIC ORTHOGONALITY
-> ALSO FOR SET OF Q-ORTHOGONAL VECTORS

POS. DEF.

CLAIM: $d_1 \ldots d_k$ Q-ORTHOGONAL => $d_1 \ldots d_k$ LIN. INDEPENDENT
PROOF: $\alpha_0 d_0 + \ldots \alpha_k d_k = 0$ TIMES $d_i Q$ => ↯ □

MIN $\frac{1}{2}x^TQx - b^Tx = f(x)$ Q POS DEF

$\rightarrow$ SOLUTION FROM $\nabla f(x) = 0 \Rightarrow Qx = b$

$\exists x^*$ SOL

LET $d_1 \dots d_n$ Q-ORTHOGONAL

CLAIM $\Rightarrow x^* = \alpha_0 d_0 + \dots \alpha_{n-1} d_{n-1}$ $\quad / d_i^T Q$

$$d_i^T \underbrace{Qx^*}_{b} = \alpha_i d_i^T Q d_i$$

$$\Rightarrow \alpha_i = \frac{d_i^T b}{d_i^T Q d_i}$$

VANISHING DUE ORTOG

REMOVING x^*

AND $x^* = \sum_{i=0}^{n-1} \frac{d_i^T b}{d_i^T Q d_i} d_i$

AS ITERATIVE PROCESS:

FOR ANY x_0 SEQUENCE $\{x_k\}$ S.T.

$$x_{k+1} = x_k + \alpha_k d_k \quad k \geq 0$$

WITH

$$\alpha_k = -\frac{g_k^T d_k}{d_k^T Q d_k}$$

GRADIENT $\boxed{\nabla f(x_k) = Q x_k - b}$

AND $g_k = Qx_k - b$ CONVERGES TO UNIQUE SOLUTION x^* OF $Qx = b$ ST $x_n = x^*$

PROOF: SINCE d_k INDEP:

$$x^* - x_0 = \alpha_0 d_0 + \dots + \alpha_{n-1} d_{n-1} \quad / d_k^T Q$$

$$d_k^T Q (x^* - x_0) = d_k^T Q d_k \cdot \alpha_k$$

$$\alpha_k = \frac{d_k^T Q (x^* - x_0)}{d_k^T Q d_k}$$

$$x_k - x_0 = \alpha_0 d_0 + \dots \alpha_{k-1} d_{k-1} \quad / d_k^T Q$$

$$d_k^T Q (x_k - x_0) = 0$$

$$\alpha_k = \frac{d_k^T Q (\overbrace{x^* - x_k}^{b} + \overbrace{x_k - x_0}^{=0})}{d_k^T Q d_k} = \frac{g_k^T d_k}{d_k^T Q d_k}$$

$$Q(x^* - x_k) = b - Q x_k = g_k$$

$\square$

QUESTIONS:

- FOR NONLINEAR PROGRAMS?
- HOW TO GET $d_0 \dots d_{n-1}$?