MODELOS QUADRÁTICOS (CONT.)

- REGIÕES DE CONFIANÇA (MIN. IRRESTRITA):

$$\min \ \frac{1}{2} d^t B_k d + \nabla f(x^k)^t d$$

- SQP:

$$\min \ \frac{1}{2} d^t B_k d + \nabla f(x^k)^t d$$
$$\text{s.a.} \ \nabla h(x^k)^t d + h(x^k) = 0.$$

→ $B_k \approx \nabla^2 f(x^k)$.

→ TAYLOR: $\nabla f(x^{k+1}) \approx \nabla f(x^k) + \nabla^2 f(x^k)(x^{k+1} - x^k)$

EQ. SECANTE $\boxed{B_{k+1} s = y}$, $s = x^{k+1} - x^k$, $y = \nabla f(x^{k+1}) - \nabla f(x^k)$

→ A EQ. SECANTE TEM SOLUÇÃO EM B (AULA PASSADA).

HÁ VÁRIAS SOLUÇÕES DA EQUAÇÃO SECANTE. UMA DELAS É REALIZAR UMA "CORREÇÃO DE POSTO 2":

$$B_{k+1} = B_k + \alpha u u^t + \beta v v^t .$$

OBS.: O POSTO DE $\alpha u u^t + \beta v v^t$ É NO MÁXIMO 2.

ESCOLHEMOS $u = y$ E $v = B_k s$. DAÍ,

$$B_{k+1} = B_k + \alpha y y^t + \beta B_k s s^t B_k \quad (B_k \text{ É SIMÉTRICA})$$

QUEREMOS $y = B_{k+1} s$ (EQ. SECANTE). ASSIM,

$$y = B_{k+1} s = B_k s + \alpha y (y^t s) + \beta B_k s (s^t B_k s).$$

UMA SOLUÇÃO É

$$\alpha = \frac{1}{y^t s}, \qquad \beta = - \frac{1}{s^t B_k s}.$$

OU SEJA,

$$\boxed{B_{k+1} = B_k + \frac{y y^t}{y^t s} - \frac{B_k s s^t B_k}{s^t B_k s}}$$

ATUALIZAÇÃO BFGS

BROYDEN FLETCHER GOLDFARB SHANNO

• BFGS É UMA DAS MAIS UTILIZADAS: ALÉM DAS PROPRIEDADES TEÓRICAS (SIMETRIA / POSITIVIDADE), TEM BOM DESEMPENHO NUMÉRICO EM PROBLEMAS GERAIS.

TEOREMA: SUPONHA QUE B_k SEJA SIMÉTRICA E DEFINIDA POSITIVA. ENTÃO, SE $\underline{y^t s > 0}$,

- B_{k+1} ESTÁ BEM DEFINIDA;
- B_{k+1} É SIMÉTRICA;
- B_{k+1} É DEF. POSITIVA.

PROVA: SE B_{k+1} É BEM DEFINIDA, É SIMÉTRICA POIS É SOMA DE MATRIZES SIMÉTRICAS. SUPONHA QUE B_k ESTEJA BEM DEFINIDA, E $y^t s > 0$ (NA ITERAÇÃO K-1). MULTIPLICANDO B_k POR s EM AMBOS OS LADOS,

$$s^t B_k s = \cancel{s^t B_{k-1} s^t} + s^t y - \cancel{s^t B_{k-1} s} = s^t y > 0 .$$

ASSIM A CONTA PARA A MATRIZ B_{k+1} É POSSÍVEL, E A ITERAÇÃO BFGS É BEM DEFINIDA.

AGORA, CONSIDERE O PRODUTO $x^t B_{k+1} x$:

$$x^t B_{k+1} x = x^t B_k x + \frac{(x^t y)^2}{y^t s} - \frac{(x^t B_k s)^2}{s^t B_k s}$$

$$= \frac{(x^t y)^2}{y^t s} + \frac{(s^t B_k s)(x^t B_k x) - (x^t B_k s)^2}{s^t B_k s}$$

COMO B_k É DEFINIDA POSITIVA, B_k POSSUI DECOMPOSIÇÃO DE CHOLESKY

$$B_k = G G^t .$$

ASSIM,

$$s^t B_k s = (s^t G)(G^t s) = \| \underline{G^t s} \|^2 ,$$

$$x^t B_k x = (x^t G)(G^t x) = \| \underline{G^t x} \|^2 \quad \text{E}$$

$$x^t B_k s = (x^t G)(G^t s) = (\underline{G^t x})^t (\underline{G^t s})$$

DE CAUCHY-SCHWARTZ OBTEMOS

$$(G^t x)^t (G^t s) \leqslant \|G^t x\|^2 \cdot \|G^t s\|^2 .$$

OU SEJA,

$$(x^t B_k x)(s^t B s) - (x^t B_k s)^2 \geqslant 0 . \quad (*)$$

ASSIM,

$$x^t B_{k+1} x \geqslant 0 \qquad (B_{k+1} \text{ É SEMI-DEF. POSITIVA}).$$

AGORA, $\underline{x^t B_{k+1} x = 0} \Longrightarrow x^t y = 0$ <u>E</u> $(*) = 0$.

SABEMOS QUE $(*)$ SÓ SE REALIZA COMO IGUALDADE SE $G^t x$ E $G^t s$ FOREM COLINEARES. ISTO É, $G^t x = \mu G^t s$. COMO G^t É INVERSÍVEL (POIS É O FATOR DE CHOLESKY DE B_k), ENTÃO $x = \mu s$.

DAÍ, $\underbrace{x^t y}_{=0} = \mu \underbrace{s^t y}_{>0 \text{ (HIP)}} \Longrightarrow \mu = 0 \Rightarrow x = 0.$

OU SEJA,

$$x^t B_{k+1} x > 0, \quad \forall x \neq 0.$$

(B_{k+1} É DEF. POSITIVA). ∎

<u>OBS.</u>: O TEOREMA DIZ QUE B_k DEF. POSIT. / SIMÉTRICA $\Longrightarrow B_{k+1}$ DEF. POSIT. / SIMÉTRICA (SE $y^t s > 0$).

INICIAMOS ENTÃO $B_0 = I$ (IDENTIDADE).

CASO ACONTEÇA DE $y^t s \not> 0$, A ATUALIZAÇÃO BFGS PODE NÃO SER BEM DEFINIDA OU NÃO SER DEF. POSITIVA. SOLUÇÃO: GAMBIARRA:

TROCAR y POR

$$\hat{y} := \theta y + (1-\theta) B_k s$$

PARA UM $\theta \in (0,1]$ ADEQUADO.

- $\theta = 1 \Rightarrow$ BFGS USUAL
- $\theta \approx 0 \Rightarrow \hat{y}^t s = \theta (y^t s) + (1-\theta) \underbrace{(s^t B_k s)}_{>0}$ TENDE A FICAR POSITIVO.

PROBLEMA: SE $\theta < 1$, A EQUAÇÃO SECANTE PODE FALHAR $(B_{k+1} s \neq \hat{y})$.