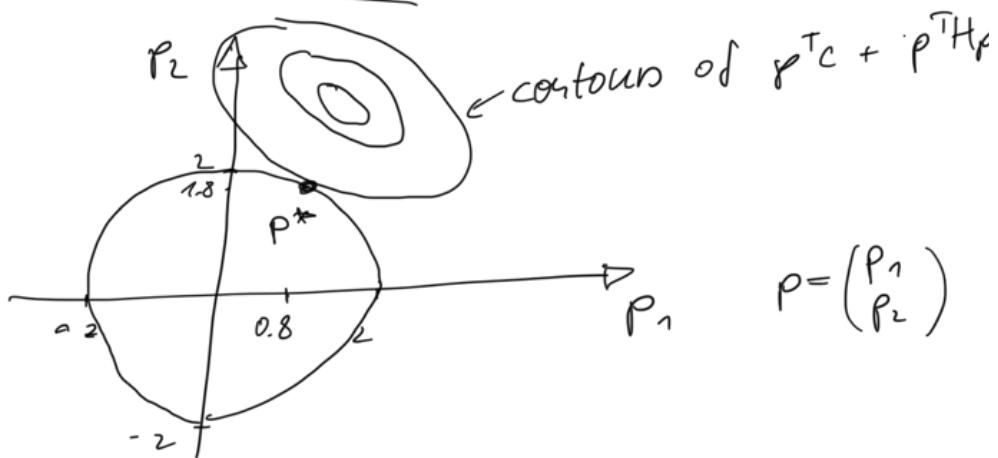


-1      2    0    1



The above theorem can be generalized for a positive semidefinite or indefinite matrix  $H$ . In these cases, the solution need not be unique.

Theorem : Let  $H \in \mathbb{R}^{n,n}$  be symmetric and  $c \in \mathbb{R}^n$ . Any global minimizer  $\rho^*$  of (2) satisfies  $(H + \lambda^* I)\rho^* = -c$ , where  $\lambda^*$  is s.t.  $(H + \lambda^* H)$  is positive semidefinite,  $\lambda^* \geq 0$ , and  $\lambda^*(\|\rho^*\| - 1) = 0$ .

Proof : See D.C. Sorensen : Newton Method with a model trust-region modification, SIAM J. Numerical Analysis 19, pp. 404-426, 1982. ■

Example : We apply the trust region method to the function  $f(x) = x_1^2 + x_2^2$

$$f(x) = x_1^4 + x_1^2 + x_1 x_2 + (1+x_2).$$

We get:  $\nabla f(x) = \begin{pmatrix} 4x_1^3 + 2x_1 + x_2 \\ x_1 + 2(1+x_2) \end{pmatrix}$

$$\nabla^2 f(x) = \begin{pmatrix} 12x_1^2 + 2 & 1 \\ 1 & 2 \end{pmatrix},$$

where  $\nabla^2 f(x)$  is positive definite for all  $x$ . We choose  $x^0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ,  $\Delta_0 = 1$ ,  $\eta_1 = 0.2$ ,  $\eta_2 = 0.9$ ,  $\sigma_1 = 0.5$ ,  $\delta_2 = 2$ .

$k=0$ : We get

$$H_0 = \nabla^2 f(x^0) = \begin{pmatrix} 14 & 1 \\ 1 & 2 \end{pmatrix}$$

$$c^0 = \nabla f(x^0) = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$$

$$\text{Since } |H_0^{-1}c^0| = \left| \begin{pmatrix} 2/27 & -1/27 \\ -1/27 & 14/27 \end{pmatrix} \cdot \begin{pmatrix} 7 \\ 5 \end{pmatrix} \right| = \frac{\sqrt{50}}{3} > \Delta_0,$$

we calculate the solution by means of (3):

From the diagonalization  $H_0 = Y D Y^\top$

we get  $\lambda_1 = 14.08$ ,  $\lambda_2 = 1.917$ ,

$$y^1 \approx \begin{pmatrix} 0.9966 \\ 0.08248 \end{pmatrix}, \quad y^2 \approx \begin{pmatrix} 0.08248 \\ -0.9966 \end{pmatrix},$$

$$\text{thus } \alpha_1 = c^{0\top} y^1 \approx 7.389,$$

$$\alpha_2 = c^{0\top} y^2 \approx -4.406.$$

Formula (4) yields

$$\Delta_0^2 = 1 \approx \frac{7.389^2}{(14.08 + x^*)^2} + \frac{4.406^2}{(1.917 + x^*)^2}$$

$$\Rightarrow x^* \approx 2.972.$$

By (3), we get

$$\rho^0 = - (H_0 + x^* I)^{-1} c^0$$

$$\approx - \begin{pmatrix} 16.97 & 1 \\ 1 & 4.972 \end{pmatrix} \cdot \begin{pmatrix} 7 \\ 5 \end{pmatrix}$$

$$\approx \begin{pmatrix} -0.3575 \\ -0.9377 \end{pmatrix}.$$

We evaluate the decrease:

$$S_0 = \frac{f(x^0) - f(\underline{x^0 + \rho^0})}{f(x^0) - m_0(p^0)}$$

$$\approx \frac{7 - 1.763}{7 - 1.929} \approx 1.033 \geq \eta_2, \eta$$

The decrease is better than predicted by the model ( $S_0 \geq \eta_1$ ), and so,

$$x^{*+} = \hat{x}^0 = x^0 + \rho^0 \approx \begin{pmatrix} 0.6425 \\ 0.0663 \end{pmatrix}$$

Because  $S_0 > \eta_2$ , we choose

$\Delta_1 \in [\Delta_0, \bar{\delta}_2 \Delta_0]$ . Here, we increase the trust region to  $\Delta_1 = \bar{\delta}_2 \Delta_0 = 2$ .

$k=1$ : We get

$$H_1 = \nabla^2 f(x^*) \approx \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

$$C^* = \nabla f(x^*) \approx \begin{pmatrix} 2.412 \\ 2.775 \end{pmatrix}.$$

Now, the point  $-H_1^{-1}C^* \approx -\begin{pmatrix} 0.1587 \\ 1.308 \end{pmatrix}$   
 is in the interior of the trust region  
 i.e., we get

$$\rho^* \approx -\begin{pmatrix} 0.1587 \\ 1.308 \end{pmatrix}.$$

$$\text{since } S_1 = \frac{f(x^*) - f(x^* + \rho^*)}{f(x^*) - m_1(\rho^*)}$$

$$\approx \frac{1.763 - (-0.2535)}{1.763 - (-0.2435)}$$

$$\approx 1.005 \geq \eta_2,$$

$$\text{we set } x^2 = x^* + \rho^* \approx \begin{pmatrix} 0.4838 \\ -1.242 \end{pmatrix}$$

$$\Delta_2 = \bar{\sigma}_2 \cdot \Delta_1 = 4.$$

$k=2$  : We get

$$H_2 = \nabla^2 f(x^2) \approx \begin{pmatrix} 4.809 & 1 \\ 1 & 2 \end{pmatrix},$$

$$C^2 = \nabla f(x^2) \approx \begin{pmatrix} 0.1786 \\ -0.0002 \end{pmatrix}.$$

$$\text{since } -H_2^{-1} \cdot C^2 \approx \begin{pmatrix} -0.04147 \\ 0.02084 \end{pmatrix}$$

is in the interior of the trust region,  
 we get .

$$\rho^2 \approx \begin{pmatrix} -0.04147 \\ 0.02084 \end{pmatrix}.$$

We have

$$s_2 \approx \frac{-0.2535 - (-0.2572)}{-0.2535 - (-0.2572)} >$$

i.e.,  $x^3 = x^2 + p^2 \approx \begin{pmatrix} 0.4423 \\ -1.221 \end{pmatrix}$ ,

This point is already "close" to the minimizer  $x^* \approx \begin{pmatrix} 0.4398 \\ -1.220 \end{pmatrix}$  of  $\min f(x)$ .

The above example shows that the two-region method can cause considerable computational work per iteration if an exact solution of the quadratic subproblem is determined  $\Rightarrow$  Resolution of large-scale problems is not practical. However, there are some approaches which solve the subproblem only approximately  $\Rightarrow$  Determine a point that is not worse than the so-called Cauchy-point.

Definition: The Cauchy point  $p^c$  of problem (2) is obtained by one iteration of the gradient method, starting with  $p = 0$ ; i.e.,

$$-c - -x^* \nabla m(\eta) = -x^{*c}$$

$P = \min_{\alpha} m(-\alpha c)$ ,  
 where  $m$  denotes the objective  
 function of (2) and  $\alpha^*$  is the  
 minimize of

$$\min_{\alpha} m(-\alpha c) \quad \text{s.t. } \|\alpha c\| \leq \Delta$$

Remark: Observe that  $\alpha^*$  is the minimiz  
 of the one-dimensional problem pos. o

$$\min_{\alpha} g(\alpha) := -\alpha c^T c + \frac{1}{2} \alpha^2 c^T H c \quad (5)$$

s.t.  $\|\alpha c\| \leq \Delta$

$$\text{We have : } \nabla_{\alpha} m(-\alpha c) = -c^T c + \alpha c^T H c \stackrel{!}{=} 0$$

$$\Rightarrow \alpha c^T H c = c^T c$$

$$\Rightarrow \bar{\alpha} = \frac{\overbrace{c^T c}^{> 0}}{\underbrace{c^T H c}_{> 0}} \geq 0.$$

The function  $g(\alpha)$  is a convex  
 parabola, thus, if  $\bar{\alpha} \leq \frac{\Delta}{\|c\|}$

and  $\bar{\alpha} > 0$ , then  $\bar{\alpha} = \alpha^*$ , i.e.,  
 $\bar{\alpha}$  is the minimize of (5).

otherwise, the minimum is attained  
 at the boundary :  $\bar{\alpha} = \frac{\Delta}{\|c\|}$ .

Hence, the Cauchy point is

$$x^C = -\alpha^* c \quad \text{with } \alpha^* = \min \left\{ \frac{c^T c}{\|c\|^2} \right\}$$

$$\frac{\Delta}{\|C\|} \quad \boxed{J}$$

Example: For the first subproblem in the above example, we obtain

$$H = \begin{pmatrix} 14 & 1 \\ 1 & 2 \end{pmatrix}, \quad c = \begin{pmatrix} 7 \\ 5 \end{pmatrix}, \quad \Delta = 1.$$

Therefore,

$$\frac{c^T C}{c^T H c} = \frac{37}{403} \approx 0.09181 \Rightarrow \alpha^*$$

$$\frac{\Delta}{\|C\|} = \frac{1}{\sqrt{74}} \approx 0.1162$$

$$\Rightarrow p^c = - \frac{37}{403} \cdot \begin{pmatrix} 7 \\ 5 \end{pmatrix} \approx - \begin{pmatrix} 0.6427 \\ 0.459 \end{pmatrix}$$

The objective values are  $m(p^c) \approx -3.3$  and  $m(p^*) \approx -5.071$ . ■

We finally present a global convergence theorem:

Theorem: Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  be twice continuously differentiable. If the solution of the subproblem is not worse<sup>\*</sup> than the Cauchy point and every Hessian matrix  $\nabla^2 f(x^k)$  is bounded, then the trust region

method satisfies:

$$\lim_{k \rightarrow \infty} f(x^k) = -\infty$$

or  $\lim_{k \rightarrow \infty} \nabla f(x^k) = 0$ .

(\*) in terms of the function  $m$  i.e.,  $m(p^k) \not\geq m(p^C)$

---

### Linear Problems

Introduction: objective function:

$$Q = c^T x \quad (c \in \mathbb{R}^n)$$

constraint:

$$x \geq 0$$

$$Ax \leq b \quad \text{or} \quad Ax = b$$

(inequality or equality constraints)

( $A \in \mathbb{R}^{m,n}$ )  $\Rightarrow$   $m$  inequ./  
equ.  
 $n$  unknown.

Problem:  $\min c^T x$

s.t.  $Ax \leq b$   
 $x \geq 0$



Example: Two products A and B shall be produced.