Ex 1 (Descent direction)

Def $p \in \mathbb{R}^n$ desc. dir $\Leftrightarrow$ $f(x+\alpha p) < f(x)$ $\forall \alpha \in (0,$

$$f(x) = x_1^2 + x_1 x_2 - 4x_2^2 + 10$$

Show that $p = (2 \ \ -1)^T$ is not a descent direction at $x' = (1 \ \ 1)^T$

Taylour expansion:

$$f(x+\alpha p) = f(x) + \alpha \nabla f(x)^T p + O(\alpha^2)$$

If $\nabla f(x)^T p > 0$

$$\frac{f(x+\alpha p) - f(x)}{\alpha} = \nabla f(x)^T p + O(\alpha)$$

$$\Rightarrow \quad \frac{f(x+\alpha p) - f(x)}{\alpha} > 0 \quad \forall \alpha : 0 < \alpha \leq \delta \quad \text{some } \delta$$

i.e. $f(x+\alpha p) > f(x)$ $\quad \alpha \in (0, \delta]$ some δ

We have $\nabla f(x) = \begin{pmatrix} 2x_1 + x_2 \\ -8x_2 + x_1 \end{pmatrix}$

$\nabla f(x^1) = \begin{pmatrix} 3 \\ -7 \end{pmatrix}$

$\nabla f(x^1)^T p = (3 \quad -7) \begin{pmatrix} 2 \\ -1 \end{pmatrix} = 13 > 0$

$\Rightarrow$ p not desc. direction

Ex 2

$\boxed{\text{Descent direction}}$

$f \in C^2$

Assume x^0 is a point s.t. $\nabla f(x^0) = 0$,

$\nabla^2 f(x^0)$ indefinite.

Show that there is a desc. direction at x^0.

$\boxed{\text{2 min}}$

$$f(x+\alpha p) = f(x) + \alpha p^T \nabla f(x) + \alpha^2 p^T \nabla^2 f(x) p + O(\alpha^3)$$

we have $\nabla f(x^0) = 0$ and $p^T \nabla^2 f(x) p < 0$

for some p.

$$\Rightarrow \frac{f(x^0 + \alpha p) - f(x^0)}{\alpha^2} = \underbrace{p^T \nabla^2 f(x^0) p}_{\wedge\, 0} + O(\alpha)$$

$\exists \delta$ s.t. $\forall \alpha \in (0, \delta]$ $\left(|O(\alpha)| < p^T \nabla^2 f(x^0) p \right)$

$$\frac{f(x^0 + \alpha p) - f(x^0)}{\alpha^2} < 0$$

$\Rightarrow$ $f(x^0 + \alpha p) < f(x^0)$ $\Rightarrow p$ desc. direction.

Ex 3

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) = (2x_1^2 - x_2)^2 + 3x_1^2 - x_2$$

a) Perform one iteration of S.D. with exact line search. Start at $x_0 = (1/2, 5/4)^T$

1/ decide direction

$$\nabla f(\mathbf{x}) = \begin{pmatrix} 2(2x_1^2 - x_2)4x_1 + 6x_1 \\ 2(2x_1^2 - x_2)(-1) - 1 \end{pmatrix}$$

$$p = -\nabla f(x^0) \underset{\dots}{=} = -\begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix}$$

2) Line search $\boxed{2 \text{ min}}$

$$\min_{\alpha \geq 0} \varphi(\alpha) = f(x^0 + \alpha p)$$ solve exactly!

$$\varphi(\alpha) = \left(-\frac{3}{4} + \frac{\alpha}{2}\right)^2 - \frac{1}{2} + \frac{\alpha}{2}$$

$$0 = \varphi'(\alpha) = 2\left(-\frac{3}{4} + \frac{\alpha}{2}\right)\frac{1}{2} + \frac{1}{2} = 0$$

$$\Leftrightarrow \alpha = \frac{1}{2}$$

$\varphi''(\alpha) = \frac{1}{2} > 0$ quadratic convex funct.

$\Rightarrow \quad \alpha = \frac{1}{2}$ optimal

$\left(\varphi'(0) < 0 \Rightarrow \alpha = 0 \text{ non-optimal} \right)$

$$x^1 = x^0 + \frac{1}{2} p = \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}$$

b) Is the function convex in a small nbh. of x^1?

$$\nabla f(x) = \begin{pmatrix} 4^2 \underbrace{x_1^2 x_1}_{= x_1^3} - 8x_2 x_1 + 6x_1 \\ -4x_1^2 + 2x_2 - 1 \end{pmatrix}$$

$$\nabla^2 f(x) = \begin{pmatrix} 3 \cdot 4^2 \cdot x_1^2 - 8x_2 + 6 & -8x_1 \\ -8x_1 & 2 \end{pmatrix}$$

$$\nabla^2 f(x^1) = \ldots = \begin{pmatrix} 10 & -4 \\ -4 & 2 \end{pmatrix}$$

eigenvalues?
$$\begin{vmatrix} 10-\lambda & -4 \\ -4 & 2-\lambda \end{vmatrix} = (10-\lambda)(2-\lambda) + 4^2 = 0$$

$\Leftrightarrow \quad \lambda^2 - 12\lambda + 36 = 0 \quad \Rightarrow \quad \lambda_{1,2} = 6$

$\nabla^2 f(x)$ is cont $\Rightarrow$ $\nabla^2 f(x)$ is pos def.
$\forall x \in nbh(x^1)$

c) Will the method conv. to a global min?

p. 279 thm 11.4: Method will converge to stationary points.

Which stationary points exist?

$\nabla f(x) = 0$

$$\begin{cases} 16x_1^3 - 8x_2x_1 + 6x_1 = 0 \\ -4x_1 + 2x_2 - 1 = 0 \end{cases}$$

$\Rightarrow \quad x_2 = \dfrac{1 + 4x_1^2}{2}$

$16x_1^3 - 4(1 + 4x_1^2)x_1 + 6x_1$

$= -x_1 + 6x_1 = 5x_1 = 0 \quad \Rightarrow \quad x_2 = \frac{1}{2}$

only <u>one</u> stationary point!

f coercive $\Rightarrow \exists$ opt $\Rightarrow$ stat. is opt $\rightarrow$ YES!

Ex 5 (Newton's method with Armijo's step length rule)

$$\min f(x) = \frac{1}{2}(x_1 - 2x_2)^2 + x_1^4$$

a) Begin at $x^0 = (2\ 1)^T$ use $\mu = 0.3$

Do one iter. of Newton + Armijo's s.l.-rule

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

$$\nabla^2 f(x) = \begin{pmatrix} 12x_1^2 + 1 & -2 \\ -2 & 4 \end{pmatrix}$$

$$\nabla^2 f(x^0) = \begin{pmatrix} 44 & -2 \\ -2 & 4 \end{pmatrix}$$

pos. def. MATLAB

$$\nabla f(x^0) = \begin{pmatrix} 32 \\ 0 \end{pmatrix}$$

solve $\nabla^2 f(x^0) p = -\nabla f(x^0)$ solved by $p_1 = -\frac{2}{3}$ $p_2 = -\frac{1}{3}$

choose $\quad x^0 + \alpha p = \begin{pmatrix} 2 - \frac{2}{3}\alpha \\ 1 - \frac{1}{3}\alpha \end{pmatrix}$

$$\varphi(\alpha) = f(x^0 + \alpha p) = \frac{1}{2}\left(2 - \frac{2}{3}\alpha - 2\left(1 - \frac{1}{3}\alpha\right)\right)^2 + \left(2 - \frac{2}{3}\alpha\right)^4$$

$$= \left(2 - \frac{2}{3}\alpha\right)^4$$

$$\varphi'(\alpha) = 4\left(2 - \frac{2}{3}\alpha\right)^3\left(-\frac{2}{3}\right) = -\frac{8}{3}\left(2 - \frac{2}{3}\alpha\right)^3 \qquad \varphi'(0) = -\frac{8}{3}\cdot 2^3$$

$$\varphi(0) = 2^4$$

choose α s.t. $\quad \underbrace{\varphi(\alpha) - \varphi(0)}_{-21.33} \leq \underbrace{\mu\alpha\varphi'(0)}_{-6.4} \qquad \mu = 0.3$

$\alpha = 1$ ok

$$x^1 = x^0 + p = \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} -\frac{2}{3} \\ -\frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{4}{3} \\ \frac{2}{3} \end{pmatrix}$$

b) For which $\mu \in (0,1)$ does the rule accept $\alpha = 1$?

$$\left(2 - \frac{2}{3}\right)^4 - 2^4 \leq \mu - \frac{8}{3}\cdot 2^3$$

$$\Rightarrow \mu \leq \frac{\left(2 - \frac{2}{3}\right)^4 - 2^4}{-\frac{8}{3}\cdot 2^3} \approx 0.6019$$

Ex 6 (Levenberg – Marquardt's method)

$$\min \quad f(x) = \frac{1}{2} x^T Q x + q^T x$$

where $Q \in \mathbb{R}^{n \times n}$ symetric and pos. semi definite, but not pos. def.
At iter. step k, using Lev-Mar method, we find dir. p_k.

$$x_{k+1} = x_k + p_k$$

Show that x_{k+1} ~~$x_k + p_k$~~ is an opt. sol. to

$$\min_{y \in \mathbb{R}^n} g(y) = f(y) + \frac{\gamma}{2} \|y - x_k\|^2$$

where $\gamma > 0$ is the "shift".

$$p = \underbrace{(\nabla^2 f(x_k) + \gamma I)^{-1}}_{A} \nabla f(x_k)$$

sym. matrix

$$\nabla f(x) = Q x + q$$

$$\nabla^2 f(x) = Q$$

$$A = (Q + \gamma I)^{-1}$$

$$x_{k+1} = x_k + (Q + \gamma I)^{-1} (-Q x_k - q) \qquad (1)$$

solve $\min_{y \in \mathbb{R}^2} g(y) = f(y) + \frac{\gamma}{2} \|y - x_k\|^2$

$\|y - x_k\|^2 = (y - x_k)^T (y - x_k) = y^T y - 2 y^T x_k - x_k^T x_k$

$\nabla \|y - x_k\|^2 = 2y - 2x_k$

$\nabla^2 \|y - x_k\|^2 = 2 I$

$\Rightarrow \nabla g(y) = Qy + q + \gamma y - \gamma x_k$

$\nabla^2 g(y) = Q + \gamma I$ pos. def. since $\gamma > 0$ and Q pos. semi def.

$\Rightarrow g$ strictly convex $\Rightarrow \nabla g = 0$ suff + necessary opt. cond.

$\nabla g = 0 \Leftrightarrow Qy + \gamma y = -q + \gamma x_k$

$\Leftrightarrow y = (Q + \gamma I)^{-1}(-q) + \gamma (Q + \gamma I)^{-1} x_k$

compare (1):

$x_{k+1} = (Q + \gamma I)^{-1}(-q) + (Q + \gamma I)^{-1}(-Q x_k) + x_k$

$= (Q + \gamma I)^{-1}(-q) + (Q + \gamma I)^{-1}(\underbrace{-Q x_k + (Q + \gamma I) x_k}_{\gamma x_k})$

$\therefore x_{k+1} = y$ $\boxtimes$

if $\gamma \to 0$ $\qquad \min\limits_{y \in \mathbb{R}^n} g(y) \longrightarrow \min\limits_{y \in \mathbb{R}^n} f(y)$

<u>But</u> for small γ, the problem becomes numerically unstable !!!