Math 408 - Homework 9

$$\frac{\partial f}{\partial x_1} = 2x_1 + x_2 - 1$$

4.1.3 $f(x_1, x_2) = x_1^2 + x_2^2 + x_1 x_2 - x_1$ $\qquad \frac{\partial f}{\partial x_2} = 2x_2 + x_1$

$x^{(0)} = (0,0)$ $\quad t_k$ - line minimization

$$\nabla f = \begin{pmatrix} 2x_1 + x_2 - 1 \\ 2x_2 + x_1 \end{pmatrix} \qquad Hf = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \succ 0 \qquad \phi_0(t) = f\left(\begin{pmatrix}0\\0\end{pmatrix} + t\,d^{(k)}\right) = f(t d^{(k)})$$

ⓐ $d^{(0)} = -\nabla f(x^{(0)}) = -\begin{pmatrix}-1\\0\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}$

$\phi_0(t) = f\left(\begin{pmatrix}t\\0\end{pmatrix}\right) = t^2 - t \qquad \phi_0'(t) = 2t - 1 = 0 \Rightarrow t = 1/2$

$\phi_0''(t) = 2 > 0 \Rightarrow t = 1/2$ is the minimizer of $\phi_0(t)$

$x^{(1)} = x^{(0)} + t\,d^{(0)} = \begin{pmatrix}0\\0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1/2\\0\end{pmatrix}$ //

ⓑ $P^{(0)} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \qquad P^{(0)^{-1}} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}$

$d^{(0)} = -\begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}\begin{pmatrix}-1\\0\end{pmatrix} = \begin{pmatrix}+1/2\\0\end{pmatrix} \qquad x^{(1)} = \begin{pmatrix}0\\0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}1/2\\0\end{pmatrix} = \begin{pmatrix}1/4\\0\end{pmatrix}$ //

ⓒ eigenvalues of $Hf = -3, -1 \qquad \gamma_0 = \max\{0, 1+3\} = 4$

$P^{(0)} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} + 4\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 1 \\ 1 & 6 \end{bmatrix}$

$d^{(0)} = -\frac{1}{35}\begin{bmatrix} 6 & -1 \\ -1 & 6 \end{bmatrix}\begin{pmatrix}-1\\0\end{pmatrix} = \frac{1}{35}\begin{pmatrix}6\\-1\end{pmatrix} = \begin{pmatrix}6/35\\-1/35\end{pmatrix}$

$x^{(1)} = \begin{pmatrix}0\\0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}6/35\\-1/35\end{pmatrix} = \begin{pmatrix}3/35\\-1/70\end{pmatrix}$ //

4.1.4 If $d^{(k)} = -\nabla f(x^{(k)})$ then $-\nabla f(x^{(k)})^T d^{(k)} = -\|\nabla f(x^{(k)})\|^2$

$\Rightarrow$ we can take $\beta = 1$ in Thm 4.1.1 (a)

$d^{(k)} = -\nabla f(x^{(k)}) = -\begin{pmatrix} 2x_1^{(k)} + x_2^{(k)} - 1 \\ 2x_2^{(k)} + x_1^{(k)} \end{pmatrix}$ $\qquad$ Suppose $x^{(k)} = \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \end{pmatrix}$ converges to x^*

$\|d^{(k)}\| = \sqrt{(2x_1^{(k)} + x_2^{(k)} - 1)^2 + (2x_2^{(k)} + x_1^{(k)})^2}$

is bounded for every converging subsequence of $x^{(k)}$

$\therefore$ By Thm 4.1.1, $f(x^{(k+1)}) \leq f(x^{(k)})$ $\forall$ $k \in \{0,1,2,\dots\}$ with $\nabla f(x^{(k)}) \neq 0$

Now we have to see which of (i) – (iii) is possible:

- The function f is coercive since $Hf \succ 0$ (Ex 3.3.5 on p 35)

$\therefore$ $f \to \infty$ as $\|x\| \to \infty$.

- The function f is also convex since $Hf \succ 0$.

- If $\nabla f = 0$ $\Rightarrow$ $\begin{pmatrix} 2x_1 + x_2 - 1 \\ 2x_2 + x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ $\Rightarrow$ $x_1 = -2x_2$

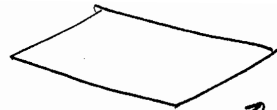
$-4x_2 + x_2 - 1 = 0 \Rightarrow -3x_2 = 1$

$\Rightarrow (x_1, x_2) = \begin{pmatrix} 2/3 \\ -1/3 \end{pmatrix}$ $\therefore$ f has only one critical pt with $Hf \succ 0$

$\Rightarrow \begin{pmatrix} 2/3 \\ -1/3 \end{pmatrix}$ is the global minimizer of f with $f\begin{pmatrix} 2/3 \\ -1/3 \end{pmatrix} = -1/3$

$\therefore$ (ii) is not possible ($f(x^{(k)}) \not\to -\infty$ as $k \to \infty$)

$\exists$ only one critical point $\therefore$ (i) or (iii) is possible.

③  graph is a plane in $\mathbb{R}^3$ — no critical points.

4.1.6 $f(x_1,x_2) = x_1 + x_2$ $\quad x^{(0)} = \begin{pmatrix}0\\0\end{pmatrix} \quad \nabla f = \begin{pmatrix}1\\1\end{pmatrix} \quad d^{(k)} = \begin{pmatrix}-1\\-1\end{pmatrix}$

In Thm 4.1.1 ⓐ is true by choosing $\beta = 1$ and ⓑ also true since $d^{(k)}$ is a constant vector

Choose t_k by Armijo's rule and get $x^{(k+1)} = x^{(k)} + t_k d^{(k)}$

$$= \begin{pmatrix}x_1^{(k)}\\x_2^{(k)}\end{pmatrix} + t_k\begin{pmatrix}-1\\-1\end{pmatrix} = \begin{pmatrix}x_1^{(k)} - t_k\\x_2^{(k)} - t_k\end{pmatrix}$$

$$\therefore f\left(x^{(k+1)}\right) = x_1^{(k)} + x_2^{(k)} - 2t_k < x_1^{(k)} + x_2^{(k)} = f\left(x^{(k)}\right)$$

(i) is never true in Thm 4.1.1 since $\nabla f(x^{(k)}) = \begin{pmatrix}1\\1\end{pmatrix}$ always.

To look at (ii) & (iii) we need to understand t_k

Armijo's rule

$$\phi_k(t) = f\left(x^{(k)} + t d^{(k)}\right) = f\left(\begin{pmatrix}x_1^{(k)}\\x_2^{(k)}\end{pmatrix} + t\begin{pmatrix}-1\\-1\end{pmatrix}\right) = f\left(\begin{pmatrix}x_1^{(k)} - t\\x_2^{(k)} - t\end{pmatrix}\right)$$

$$= x_1^{(k)} + x_2^{(k)} - 2t \qquad \phi_k(0) = x_1^{(k)} + x_2^{(k)}$$

$\phi_k'(t) = -2 < 0 \quad \forall\, t$ $\qquad$ In particular, $\phi_k'(0) = -2$

How to choose t_k? $\quad t_k =$ largest $t \in \{1, 1/2, 1/4, 1/8, \ldots\}$

s.t $\dfrac{\phi_k(t) - \phi_k(0)}{t} \le .1\, \phi_k'(0) = .1(-2) = -.2$

$$\Rightarrow \frac{x_1^{(k)} + x_2^{(k)} - 2t - \left(x_1^{(k)} + x_2^{(k)}\right)}{t} = -2 < -.2 \quad \text{always.}$$

$\therefore t_k = 1$ always.

$\therefore$ If you look at the sequence $\{x^{(k)}\}$

$$x^{(k+1)} = x^{(k)} + 1\begin{pmatrix}-1\\-1\end{pmatrix} = x^{(k)} - \begin{pmatrix}1\\1\end{pmatrix}$$

$\therefore \{x^{(k)}\}$ does no converge & no subsequence of it converges too. $\quad \therefore$ (iii) does not need to be considered

As $k \to \infty$ $\quad f(x^{(k)}) \to -\infty \quad \therefore$ (ii) is true.

The sequence $\{x^{(k)}\}$ moves ~~fro~~ like

$$\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}-1\\-1\end{pmatrix}, \begin{pmatrix}-2\\-2\end{pmatrix}, \begin{pmatrix}-3\\-3\end{pmatrix}, \ldots$$

4.1.7 $f(\underline{x}) = \frac{1}{2}x^T A x + b^T x$ A $n\times n$ symm matrix, $b \in \mathbb{R}^n$

$Hf = A$

(a) It is true that f is convex $\iff A \succeq 0$.

(b) Suppose $A \succ 0 \Rightarrow f$ convex. Then by Thm 3.5.2 if a $x^* \in \mathbb{R}^n$ satisfies $\nabla f(x^*) = 0$ then x^* is a global opt of f.

(remember A is symm.

$$f(\underline{x}) = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}x_i x_j + \sum b_i x_i$$

$$\therefore \frac{\partial f}{\partial x_i} = \frac{1}{2}\sum_{j=1}^{n} 2a_{ij}x_j + b_i \qquad \nabla f(x) = \begin{pmatrix} \cancel{\frac{2}{2}}\sum a_{1j}x_j + b_1 \\ \cancel{\frac{2}{2}}\sum a_{2j}x_j + b_2 \\ \vdots \\ \cancel{\frac{2}{2}}\sum a_{nj}x_j + b_n \end{pmatrix}$$

$$\therefore \nabla f = 0 \Rightarrow \sum_{j=1}^{n} a_{ij}x_j = -b_i \Rightarrow \sum a_{ij}x_j = -b_i \ \forall i$$

$$\Rightarrow Ax = -b \Rightarrow x = A^{-1}(-b) = -A^{-1}b$$

(A^{-1} exists since $A \succ 0$) & x unique

$\therefore \exists$ only one global min and it is $x = -A^{-1}b$ //

(c) $A \succ 0$, $x^{(0)} = 0_n$ $d^{(k)} = -Hf(x^{(k)})^{-1}\nabla f(x^{(k)}) = -A^{-1}\nabla f(x^{(k)})$

$$\nabla f(x^{(k)}) = \frac{1}{2}Ax^{(k)} + b$$

$$\therefore d^{(k)} = -A^{-1}\left(\frac{1}{2}Ax^{(k)} + b\right) = -\frac{1}{2}x^{(k)} - A^{-1}(b) = -\frac{1}{2}x^{(k)} - A^{-1}b$$

$$\phi_k(t) = f(x^{(k)} + td^{(k)}) \Rightarrow \phi_0(t) = f(td^{(0)})$$

$$= f(t(-A^{-1}b)) = f(-tA^{-1}b) = \frac{1}{2}(-tA^{-1}b)^T A(-tA^{-1}b) + b(-tA^{-1}b)$$

$$= \frac{t^2}{2}\big(b^T \underbrace{(A^{-1})^T A}_{I} A^{-1}b\big) - tb^TA^{-1}b \qquad \left(\begin{matrix}\text{recall } A \succ 0 \Rightarrow A^{-1} \succ 0 \\ \Rightarrow A^{-1} \text{ symm}\end{matrix}\right)$$

$$= \frac{t^2}{2}(b^TA^{-1}b) - tb^TA^{-1}b = \underbrace{b^TA^{-1}b}_{>0}\left(\frac{t^2}{2} - t\right)$$

$$\phi_k'(t) = \frac{2t}{2} - 1 = 0 \Rightarrow t - 1 = 0 \Rightarrow t = 1. //$$

$$\therefore x^{(1)} = \begin{pmatrix}0\\0\end{pmatrix} + 1\left(-\frac{1}{2}\begin{pmatrix}0\\0\end{pmatrix} - A^{-1}b\right) = -A^{-1}b \ //$$