数値最適化提出課題解答例

第一部の課題は資料中に十分な解説があるので解答例を省略します。
以下、第二部・第三部の提出課題解答例を説明します。

演習問題1

次の3つの関数を目的関数とした制約なし最小化問題を考える。

(A) $x_1{}^2 - 2x_1x_2$,　(B) $x_1{}^2 - x_2{}^2$,　(C) $2x_1{}^2 + x_2{}^2$

問1：各目的関数の最急降下ベクトルを示せ。

最急降下ベクトルは目的関数fにより$-\nabla f$で表されるので、目的関数毎にそれぞれ以下の通り、

(A) $-[2x_1 - 2x_2,\ -2x_1]^{\mathrm{T}}$,　(B) $-[2x_1,\ -2x_2]^{\mathrm{T}}$,
(C) $-[4x_1,\ 2x_2]^{\mathrm{T}}$

問2：初期点$x_1=1$, $x_2=1$から最急降下法を用いて（最小化問題を）解くことを考える。最初の1ステップを解き、更新後の変数値を示せ。直線探索では厳密な解が解けたものとして良い。

（ヒント）初期点を$\boldsymbol{x}_0$とすれば最急降下法の次のステップの近似解は$\boldsymbol{x}_1=\boldsymbol{x}_0-t\nabla f(\boldsymbol{x}_0)$と表すことができます。直線探索問題は $f(\boldsymbol{x}_1)=f(\boldsymbol{x}_0-t\nabla f(\boldsymbol{x}_0))$の最小化問題です。

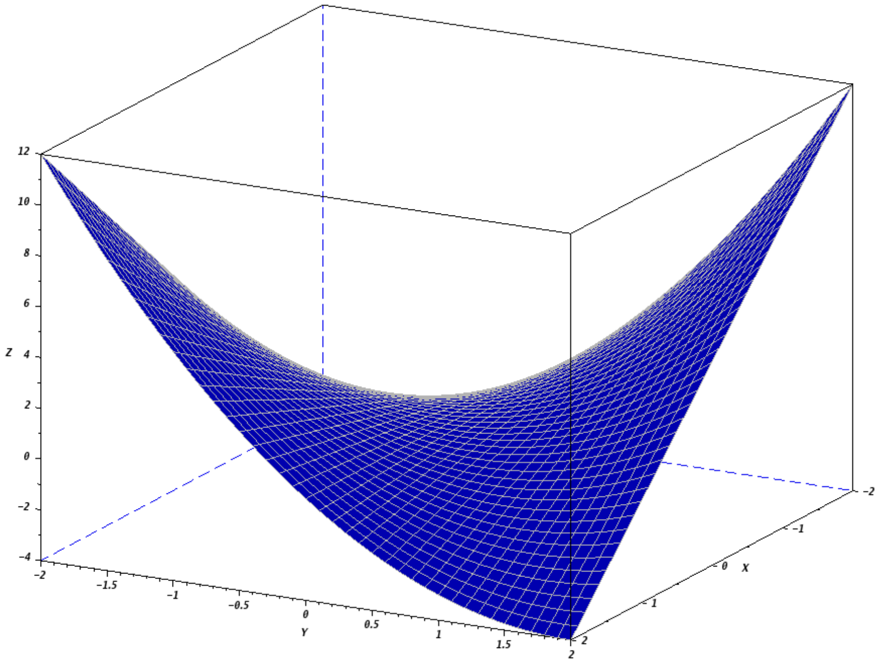
この問題で示された(A)〜(C)であれば、1変数tの停留点を求めることは易しいので近似計算をする必要はありません。

ヒントに沿って解きます。
(A)の$\boldsymbol{x}_0$における直線探索問題は次の最小化問題です。

$$\min.\ f([1,1]^{\mathrm{T}}-t[0,-2]^{\mathrm{T}}) \quad \Leftrightarrow \quad \min.\ -4t-1$$

これは1次関数の最小化問題なので変数への制約が無い場合最適解は存在しません。

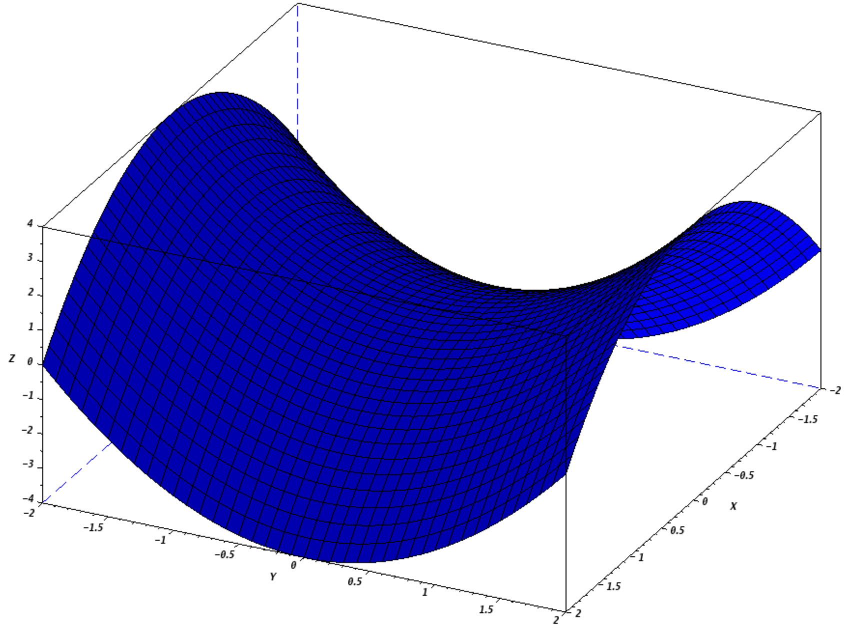
(A)を図示した下図を見ればこの目的関数の停留点が極大値・極小値のいずれもとらないことが判ります。



(B)の$\boldsymbol{x}_0$における直線探索問題は次の最小化問題です。

$$\min.\ f([1,1]^{\mathrm{T}}-t[2,-2]^{\mathrm{T}}) \quad \Leftrightarrow \quad \min.\ -8t$$

これも1次関数なので、制約なしでは極値を持ちません。下図のように(A)の目的関数を回転させたようなものになっています。



(C)の$\boldsymbol{x}_0$における直線探索問題は次の最小化問題です。

$$\min.\ f([1,1]^{\mathrm{T}}-t[4,2]^{\mathrm{T}}) \quad \Leftrightarrow \quad \min.\ 36t^2-20t+3$$

これはすぐに　$36t^2-20t+3=[6t-(5/3)]^2+(2/9)$　より停留解$t=5/18$,停留値2/9が極小値であることが判ります。

$$\therefore \quad \boldsymbol{x}_1=[1,1]^{\mathrm{T}}-(5/18)[4,2]^{\mathrm{T}}=[-1/9,4/9]^{\mathrm{T}}$$

演習問題2

次の3つの関数を目的関数とした制約なし最小化問題を考える。

(A) $x_1^2 - 2x_1x_2$,　(B) $x_1^2 - x_2^2$,　(C) $2x_1^2 + x_2^2$

問1：各目的関数のニュートン方向ベクトルを示せ。

ニュートン方向ベクトルは目的関数のヘッセ行列と勾配ベクトルを用いて$-[\mathrm{H}f(\boldsymbol{x})]^{-1}[\nabla f(\boldsymbol{x})]$で表されるので、目的関数毎に

$$(\mathrm{A})\ -\begin{bmatrix} 2 & -2 \\ -2 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 2x_1 - 2x_2 \\ -2x_1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2x_1 - 2x_2 \\ -2x_1 \end{bmatrix},$$

$$\text{(B)}\ -\begin{bmatrix}2 & 0\\ 0 & -2\end{bmatrix}^{-1}\begin{bmatrix}2x_1\\ -2x_2\end{bmatrix}=\frac{1}{2}\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}2x_1\\ -2x_2\end{bmatrix},$$

$$\text{(C)}\ -\begin{bmatrix}4 & 0\\ 0 & 2\end{bmatrix}^{-1}\begin{bmatrix}4x_1\\ 2x_2\end{bmatrix}=-\frac{1}{4}\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix}\begin{bmatrix}4x_1\\ 2x_2\end{bmatrix}$$

問2:初期点$x_1=1$, $x_2=1$から~~最急降下~~ニュートン法を用いて(最小化問題を)解くことを考える。最初の1ステップを解き、更新後の変数値を示せ。(問題では「最急降下法」と誤っていました。)

$\boldsymbol{x}_1=\boldsymbol{x}_0-[\mathrm{H}f(\boldsymbol{x}_0)]^{-1}[\nabla f(\boldsymbol{x}_0)]$を計算すれば良い。

(A)〜(C)で全て$\boldsymbol{x}_1=[0,0]^{\mathrm{T}}$となることを確認してください。

演習問題3

次の最小化問題にLagrangeの未定係数法を適用し計算過程とともに最小解を示せ。

$$\min.\quad f(x_1,x_2)=ax_1+bx_2+c,\ \text{s.t.}\ \ g(x_1,x_2)=x_1{}^2+x_2{}^2=1.$$

ヒントと注意:

例題の制約式と目的関数を入れ替えた問題です。例題の図から最適解は明らかでしょう。

停留点が複数あることに注意してください。

配付資料の解説部分に表現を合せ$g(x_1,x_2)=x_1{}^2+x_2{}^2-1=0$とすべきでした。申し訳ありません。

$$\frac{\partial}{\partial x_1} f(\boldsymbol{x}) - \lambda \frac{\partial}{x_1} g(\boldsymbol{x}) = a - \lambda x_1 = 0,$$
$$\frac{\partial}{\partial x_2} f(\boldsymbol{x}) - \lambda \frac{\partial}{x_2} g(\boldsymbol{x}) = b - \lambda x_2 = 0,$$

x_1, x_2が求まる（与式から消去できる）のでこれを制約式に持ち込みλの条件を得ます。

$$\left(\frac{a}{2\lambda}\right)^2 + \left(\frac{b}{2\lambda}\right)^2 = 1$$

λは2通り求まりますが、これは停留点が2つ存在することが理由ですから、目的関数の最小値を与えるものが最適解です。

$$\lambda = \frac{1}{2}\sqrt{a^2+b^2}, x_1 = -a/\sqrt{a^2+b^2}, x_2 = -b/\sqrt{a^2+b^2}$$

演習問題4

次の最小化問題のKKT条件を示し、最適解を求めよ。

$$\text{min.}\quad f(x_1,x_2)=x_1{}^2+x_2{}^2,\ \text{s.t.}\ h(x_1,x_2)\cancel{\geqq}=ax_1+bx_2+c.$$

「≧」は「＝」の誤りでした。これで制約関数hを定義し、不等式制約「$\mathrm{h}(x_1,x_2)\geqq 0$」を課すことを意図した問題です。そもそも訂正前の問題は問題として成立しませんので、解答例はこの意図に沿ったものです。

提出された回答では多くの方が意図した問題として解釈してくださっていました。採点はそれ以外のそれぞれの解釈に沿ったもの、あるいは問題が成立しないことを指摘していただいたものも正解として行っています。

配付資料で解説したKKT条件は目的関数と有効制約から定められます。ここでは、制約条件はただ一つの不等式で与えられているので、これが有効制約である場合とそうでない場合を考えれば十分です。

まず$h(x)≥0$が有効制約である場合を考えます。KKT条件は

$$\frac{\partial}{\partial x_1} f(\boldsymbol{x}) - \mu \frac{\partial}{x_1} h(\boldsymbol{x}) = 0, \frac{\partial}{\partial x_2} f(\boldsymbol{x}) - \mu \frac{\partial}{x_2} h(\boldsymbol{x}) = 0, \mu \geq 0,$$

を満たす$\boldsymbol{x}=(x_1,x_2)$,μによって成立します。すなわち、この問題では

$$2x_1 - \mu a = 0,\ 2x_2 - \mu b = 0,\ \mu \geq 0,$$

をKKT条件として示すことになります。一般的には有効制約の組み合わせとともにKKT条件を満たす変数・パラメタを探索しますが、この問題ではさらに条件を絞ることができます。

$h(\boldsymbol{x})\geq 0$が有効制約であるためには不等式が等号で成立する必要があります。単一の不等式であれば、それは単に等式制約を考えることと一致します。

$$h(\boldsymbol{x})=ax_1+bx_2+c=0$$

KKT条件を併せればパラメタμの条件が得られます。

$$(\mu/2)(a^2+b^2)+c=0,\ \mu\geq 0,$$

この条件は$c\leq 0$のときのみ$\mu=-2c/(a^2+b^2)\geq 0$のもとで成立します。（$a^2+b^2=0$の場合を考える必要はないでしょう。）
これは、$c\leq 0$のとき以下で最適性条件を満たすということです。

$$ax_1+bx_2+c=0,\ \mu=-2c/(a^2+b^2),\ x_1=\mu a/2,\ x_2=\mu b/2$$

さらにμを消去して条件を整理します。

$h(\boldsymbol{x})\geq 0$が有効制約である場合（$h(\boldsymbol{x})=ax_1+bx_2+c=0$）
最適解は以下を満たす

$$[a,b][x_1,x_2]^{\mathrm{T}}\geq 0,\ x_1=-ac/(a^2+b^2),\ x_2=-bc/(a^2+b^2)$$

この点$[x_1,x_2]$が原点から見て点$[a,b]$と同じ側にあれば
⇔境界$h(\boldsymbol{x})=0$が点$[a,b]$と同じ側にあれば
⇔領域$h(\boldsymbol{x})\geq 0$に原点が含まれていなければ
$[x_1,x_2]$は停留点である。

$h(\boldsymbol{x})\geq 0$が有効制約でない場合（$h(\boldsymbol{x})=ax_1+bx_2+c>0$）

最適解は制約なし問題の最適性条件を満たす。

この問題では制約なし問題の停留点は原点のみで、目的関数は凸なので、

「領域$h(\boldsymbol{x})\geq 0$に原点が含まれていなければ

$$[x_1, x_2] = -c/(a^2+b^2) \times [a, b]$$

は最小解である。

領域$h(\boldsymbol{x})\geq 0$に原点が含まれていれば

$$[x_1, x_2] = [0, 0]$$

すなわち原点は最小解である。」

ということが言えます。