MATH60005/70005: Optimisation (Autumn 24-25)

Chapter 2: exercises, solutions, and additional notes

Dr Dante Kalise, Dr Estefanía Loayza-Romero and Sara Bicego
Department of Mathematics
Imperial College London, United Kingdom
{dkaliseb,kloayzar,s.bicego21}@imperial.ac.uk

1. Find the global minimum and maximum points of $f(x_1, x_2) = x_1 + x_2$ over the unit ball in $\mathbb{R}^2$, $S = B[0,1] = \left\{(x_1, x_2)^T : x_1^2 + x_2^2 \leq 1\right\}$. Repeat with $f(x_1, x_2) = 2x_1 - 3x_2$ over the set $S = \{(x_1, x_2) : 2x_1^2 + 5x_2^2 \leq 1\}$.

   Answer: as seen in lectures. Consider

   $$f(x_1, x_2) = (1 \;\; 1)^\top \mathbf{x} \leq \|(1 \;\; 1)\| \|\mathbf{x}\| \leq \sqrt{2}\,.$$

   The upper bound is attained with the maximizer $\mathbf{x} = (1 \;\; 1)^\top/\sqrt{2}$ and a similar lower bound is attained with the minimizer $\mathbf{x} = -(1 \;\; 1)^\top/\sqrt{2}$. For the second part, we consider the change of variables

   $$u = \sqrt{2}x_1\,, \qquad v = \sqrt{5}x_2\,,$$

   so that S becomes the unit ball in (u, v), where we optimize $f(u, v) = \sqrt{2}u - \frac{3}{\sqrt{5}}v$ as in the first part of the exercise. Don't forget to go back to the original variables (x_1, x_2) at the very end.

2. Classify the matrices

   $$\mathbf{A} = \begin{pmatrix} 4 & 1 \\ 1 & 3 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0.1 \end{pmatrix}.$$

   Answer: the matrix $\mathbf{A}$ is direct from $det(\mathbf{A}) = 10$ and $tr(\mathbf{A}) = 7$, hence positive definite. For $\mathbf{B}$, both diagonally dominant and principal minors criteria are inconclusive, and therefore we must go back to the definition of indefinite matrix. Note that

   $$\mathbf{e}_1^\top \mathbf{B} \mathbf{e}_1 = 1 > 0\,, \quad \text{and} \quad (\mathbf{e}_2 - \mathbf{e}_3)^\top \mathbf{B} (\mathbf{e}_2 - \mathbf{e}_3) = -0.9 < 0\,,$$

   and we conclude the matrix $\mathbf{B}$ is indefinite.

3. Use a computational tool of your preference to classify

$$\begin{pmatrix} 2 & 2 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 3 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} -5 & 1 & 1 \\ 1 & -7 & 1 \\ 1 & 1 & -5 \end{pmatrix}$$

Answer: positive definiteness can be easily check in a computer by looking at the eigenvalues of the matrix. For example, in matlab this is done by using the command `eig()`. For example

```
A=[2 2 0 0;2 2 0 0 ;0 0 3 1;0 0 1 3]
eig(A)
```

will give as an output the eigenvalues 0,2,4,4, from where we directly conclude that the matrix is positive semidefinite.

4. Classify the stationary points of

a) $f(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_1x_3 + x_2x_3 + \left(x_1^2 + x_2^2 + x_3^2\right)^2$.

Answer: In this case we find that $\mathbf{x} = 0$ is a stationary point and that the Hessian can be expressed as

$$\nabla^2 f = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} + 4\|\mathbf{x}\|^2 I_3 + 8\mathbf{x}\mathbf{x}^\top ,$$

where each term corresponds to a positive semi-definite matrix for all $\mathbf{x}$ in $\mathbb{R}^3$. The first matrix can be checked by the diagonally dominant criterion, the second matrix is a diagonal matrix with non-negative entries, and for the third matrix we observe that given a vector $\mathbf{x}$, the product

$$\mathbf{v}^\top \mathbf{x}\mathbf{x}^\top \mathbf{v} = (\mathbf{x}^\top \mathbf{v})^2 \geq 0 , \quad \forall \mathbf{v} \in \mathbb{R}^3 ,$$

hence the matrix $8\mathbf{x}^\top \mathbf{x}$ is positive semidefinite. The sum of positive semidefinite matrices is positive semidefinite. Since the above is valid for all $\mathbf{x} \in \mathbb{R}^3$, from global optimality conditions we conclude that $\mathbf{x} = 0$ is a global minimizer.

b) $f(x_1, x_2) = x_1^4 + 2x_1^2x_2 + x_2^2 - 4x_1^2 - 8x_1 - 8x_2$.

Answer: The gradient is given by

$$\nabla f = \begin{pmatrix} 4x_1^3 + 4x_1x_2 - 8x_1 - 8 \\ 2x_1^2 + 2x_2 - 8 \end{pmatrix}$$

from where $\nabla f = 0$ gives $x_1 = 1$ and $x_2 = 3$. The Hessian is given by

$$\nabla^2 f = \begin{pmatrix} 12x_1^2 + 4x_2 - 8 & 4x_1 \\ 4x_1 & 2 \end{pmatrix}, \quad \nabla^2 f(1,3) = \begin{pmatrix} 16 & 4 \\ 4 & 2 \end{pmatrix} \succ 0 ,$$

hence it is a strict local minimizer. We also observe that

$$f(x_1, x_2) = (x_1^2 + x_2 - 4)^2 + 4(x_1 - 1)^2 - 20 \geq -20\,,$$

and that $f(1, 3) = -20$, therefore it is a global minimizer.

c) $f(x_1, x_2) = x_1^2 + 4x_1x_2 + x_2^2 + x_1 - x_2$.

Answer: this case is solved by noting that f is quadratic function

$$f(\mathbf{x}) = \mathbf{x}^\top \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \mathbf{x} + \frac{1}{2}[2 \quad -2]^\top \mathbf{x}\,,$$

and noting that $det(\mathbf{A}) = -3 < 0$, hence the Hessian is indefinite and $\mathbf{x}^* = -\mathbf{A}^{-1}\mathbf{b}$ is a saddle point.

Additional notes

Page 4, example

to show that the stationary points are maximizers and minimizers, we need to use Cauchy-Schwarz inequality. We note that

$$f(\mathbf{x}) = \frac{x_1 + x_2}{x_1^2 + x_2^2 + 1} \leq \sqrt{2}\frac{\sqrt{x_1^2 + x_2^2}}{x_1^2 + x_2^2 + 1} \leq \sqrt{2} \max_{t \geq 0} \frac{t}{t^2 + 1} \leq \frac{\sqrt{2}}{2}\,,$$

which is the value attained at the stationary point (or with a minus for the minimizer).

Page 8, example

For the stationary point $(0.5, 0)$ the Hessian is positive semidefinite, and it is a strict local min (not global since the function is not bounded below, check $f(-1, x_2)$ and $x_2 \to \infty$). For $(-0.5, 0)$ the Hessian is indefinite, from we directly conclude it is a saddle point. For the point $(0, 0)$ it is more complicated, as the Hessian is negative semidefinite, so it can be either a local max or a saddle point. Here, we will show it is a saddle point by using trajectories. Note that

$$f(\alpha^4, \alpha) = \alpha^6(-2\alpha^2 + 1 + 4\alpha^{10})\,,$$

which is positive as $\alpha \to 0$. Instead, if we now take

$$f(-\alpha^4, \alpha) = \alpha^6(-2\alpha^2 - 1 + 4\alpha^{10})\,,$$

this is negative as $\alpha \to 0$. This means that in any ball surrounding $(0, 0)$, we will find larger and smaller values than $f(0, 0) = 0$, hence it is a saddle point.

Note that for this to happen, you can only play with he term $x_1x_2^2$, as the other terms won't change their sign no matter what you try .

Now, more generally, define a family of curves $x_1 = \alpha^\beta$, and $x_2 = \alpha^\gamma$, where $\beta, \gamma > 0$ so you don't have problems as $\alpha \longrightarrow 0$. If you repeat the calculations, trying to factor out a power such that the term ± 1 remains inside the parenthesis -this is the one that won't vanish and allows you to play with the sign-, you'll find that any positive β, γ that satisfy

$$\beta - 2\gamma > 0 \qquad \text{and} \qquad 3\beta - 2\gamma > 0$$

do the job (the second inequality is redundant, I'm posting it for completeness). In particular, taking $\beta = 4$ and $\gamma = 1$ as I did, but you can also try with $\beta = 3, \gamma = 1$, or $\beta = 5$ and $\gamma = 2$, and many others!