ECE/MATH 520, Spring 2008

Homework Problems 1

Solutions (version: February 5, 2008, 11:7)

1.5 Suppose you are shown four cards, laid out in a row. Each card has a letter on one side and a number on the other. On the visible side of the cards are printed the symbols:

$$S \qquad 8 \qquad 3 \qquad A$$

Determine which cards you should turn over to decide if the following rule is true or false: "If there is a vowel on one side of the card, then there is an even number on the other side."

Ans.: The cards that you should turn over are 3 and A. The remaining cards are irrelevant to ascertaining the truth or falsity of the rule. The card with S is irrelevant because S is not a vowel. The card with 8 is not relevant because the rule does not say that if a card has an even number on one side, then it has a vowel on the other side.

Turning over the A card directly verifies the rule, while turning over the 3 card verifies the contraposition.

2.6 Show that for any two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$.
Hint: Write $\boldsymbol{x} = (\boldsymbol{x} - \boldsymbol{y}) + \boldsymbol{y}$, and use the Triangle inequality. Do the same for $\boldsymbol{y}$.

Ans.: We have $\|\boldsymbol{x}\| = \|(\boldsymbol{x} - \boldsymbol{y}) + \boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{y}\| + \|\boldsymbol{y}\|$ by the Triangle Inequality. Hence, $\|\boldsymbol{x}\| - \|\boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$. On the other hand, from the above we have $\|\boldsymbol{y}\| - \|\boldsymbol{x}\| \leq \|\boldsymbol{y} - \boldsymbol{x}\| = \|\boldsymbol{x} - \boldsymbol{y}\|$. Combining the two inequalities, we obtain $|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$.

3.2 Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of the matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$. Show that the eigenvalues of the matrix $\boldsymbol{I}_n - \boldsymbol{A}$ are $1 - \lambda_1, \ldots, 1 - \lambda_n$.

Ans.: Suppose $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are eigenvectors of $\boldsymbol{A}$ corresponding to $\lambda_1, \ldots, \lambda_n$, respectively. Then, for each $i = 1, \ldots, n$, we have

$$(\boldsymbol{I}_n - \boldsymbol{A})\boldsymbol{v}_i = \boldsymbol{v}_i - \boldsymbol{A}\boldsymbol{v}_i = \boldsymbol{v}_i - \lambda_i \boldsymbol{v}_i = (1 - \lambda_i)\boldsymbol{v}_i$$

which shows that $1 - \lambda_1, \ldots, 1 - \lambda_n$ are the eigenvalues of $\boldsymbol{I}_n - \boldsymbol{A}$.

Alternatively, we may write the characteristic polynomial of $\boldsymbol{I}_n - \boldsymbol{A}$ as

$$\pi_{\boldsymbol{I}_n - \boldsymbol{A}}(1 - \lambda) = \det((1 - \lambda)\boldsymbol{I}_n - (\boldsymbol{I}_n - \boldsymbol{A})) = \det(-[\lambda \boldsymbol{I}_n - \boldsymbol{A}]) = (-1)^n \pi_{\boldsymbol{A}}(\lambda),$$

which shows the desired result.

3.12a Consider the matrix

$$\boldsymbol{Q} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

a. Is this matrix positive definite, negative definite, or indefinite?

Ans.: a. The matrix $\boldsymbol{Q}$ is indefinite, since $\Delta_2 = -1$ and $\Delta_3 = 2$.

5.5 Let $\boldsymbol{x}(t) = [e^t+t^3, t^2, t+1]^T$, $t \in \mathbb{R}$, and $f(\boldsymbol{x}) = x_1^3 x_2 x_3^2 + x_1 x_2 + x_3$, $\boldsymbol{x} = [x_1, x_2, x_3]^T \in \mathbb{R}^3$. Find $(d/dt) f(\boldsymbol{x}(t))$ in terms of t.

Ans.: We have

$$Df(\boldsymbol{x}) = [3x_1^2 x_2 x_3^2 + x_2,\ x_1^3 x_3^2 + x_1,\ 2x_1^3 x_2 x_3 + 1]$$

and

$$\frac{d\boldsymbol{x}}{dt}(t) = \begin{bmatrix} e^t + 3t^2 \\ 2t \\ 1 \end{bmatrix}.$$

By the chain rule,

$$\begin{aligned}
&\frac{d}{dt} f(\boldsymbol{x}(t)) \\
&= Df(\boldsymbol{x}(t)) \frac{d\boldsymbol{x}}{dt}(t) \\
&= [3x_1(t)^2 x_2(t) x_3(t)^2 + x_2(t),\ x_1(t)^3 x_3(t)^2 + x_1(t),\ 2x_1(t)^3 x_2(t) x_3(t) + 1] \begin{bmatrix} e^t + 3t^2 \\ 2t \\ 1 \end{bmatrix} \\
&= (3x_1(t)^2 x_2(t) x_3(t)^2 + x_2(t))(e^t + 3t^2) + (x_1(t)^3 x_3(t)^2 + x_1(t))(2t) \\
&\quad + 2x_1(t)^3 x_2(t) x_3(t) + 1 \\
&= 12t(e^t + 3t^2)^3 + 2te^t + 6t^2 + 2t + 1.
\end{aligned}$$

(Actually, you're not expected or required to expand out and simplify to the last formula.)

5.8 Let

$$\begin{aligned}
f_1(x_1, x_2) &= x_1^2 - x_2^2; \\
f_2(x_1, x_2) &= 2x_1 x_2.
\end{aligned}$$

Sketch the level sets associated with $f_1(x_1, x_2) = 12$ and $f_2(x_1, x_2) = 16$ on the same diagram. Indicate on the diagram the values of $\boldsymbol{x} = [x_1, x_2]^T$ for which $\boldsymbol{f}(\boldsymbol{x}) = [f_1(x_1, x_2), f_2(x_1, x_2)]^T = [12, 16]^T$.

Ans.: We have that

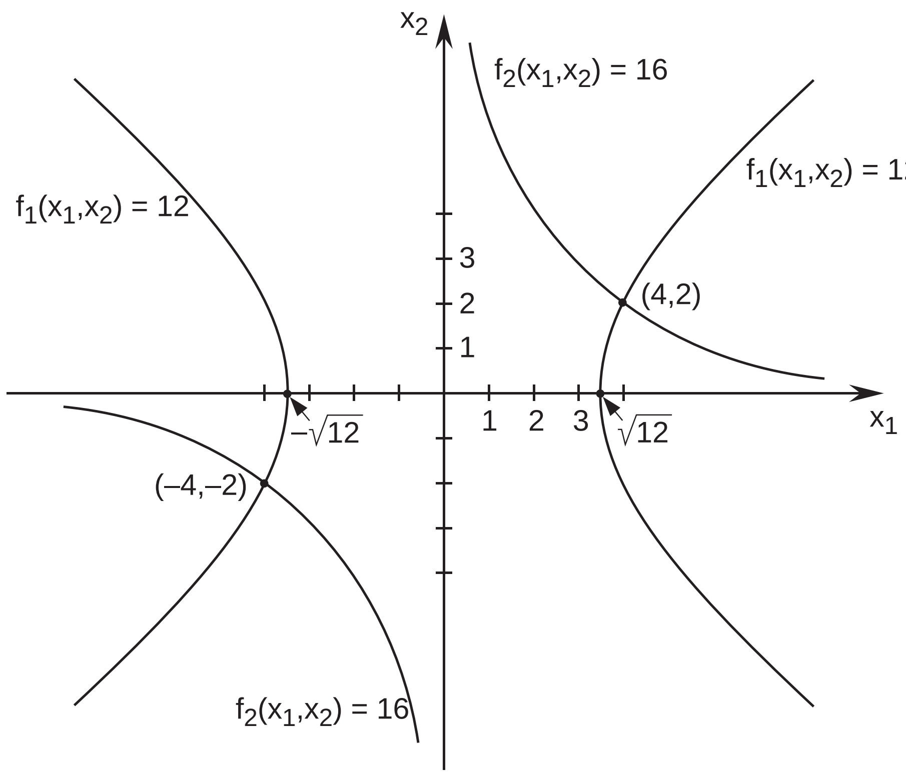
$$\{\boldsymbol{x} : f_1(\boldsymbol{x}) = 12\} = \{\boldsymbol{x} : x_1^2 - x_2^2 = 12\},$$

and

$$\{\boldsymbol{x} : f_2(\boldsymbol{x}) = 16\} = \{\boldsymbol{x} : x_2 = 8/x_1\}.$$

To find the intersection points, we substitute $x_2 = 8/x_1$ into $x_1^2 - x_2^2 = 12$ to get $x_1^4 - 12x_1^2 - 64 = 0$. Solving gives $x_1^2 = 16, -4$. Clearly, the only two possibilities for x_1 are $x_1 = +4, -4$, from which we obtain $x_2 = +2, -2$. Hence, the intersection points are located at $[4, 2]^T$ and $[-4, -2]^T$.

The level sets associated with $f_1(x_1, x_2) = 12$ and $f_2(x_1, x_2) = 16$ are shown as follows.



5.9 Write down the Taylor series expansion of the following functions about the given points $\boldsymbol{x}_0$. Neglect terms of order three or higher.

a. $f(\boldsymbol{x}) = x_1 e^{-x_2} + x_2 + 1$, $\boldsymbol{x}_0 = [1, 0]^T$

b. $f(\boldsymbol{x}) = x_1^4 + 2x_1^2 x_2^2 + x_2^4$, $\boldsymbol{x}_0 = [1, 1]^T$

c. $f(\boldsymbol{x}) = e^{x_1 - x_2} + e^{x_1 + x_2} + x_1 + x_2 + 1$, $\boldsymbol{x}_0 = [1, 0]^T$.

Ans.: a. We have

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_o) + Df(\boldsymbol{x}_o)(\boldsymbol{x} - \boldsymbol{x}_o) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_o)^T D^2 f(\boldsymbol{x}_o)(\boldsymbol{x} - \boldsymbol{x}_o) + \cdots.$$

We compute

$$\begin{aligned} Df(\boldsymbol{x}) &= [e^{-x_2}, -x_1e^{-x_2}+1], \\ D^2f(\boldsymbol{x}) &= \begin{bmatrix} 0 & -e^{-x_2} \\ -e^{-x_2} & x_1e^{-x_2} \end{bmatrix}. \end{aligned}$$

Hence,

$$\begin{aligned} f(\boldsymbol{x}) &= 2+[1,0]\begin{bmatrix} x_1-1 \\ x_2 \end{bmatrix} + \frac{1}{2}[x_1-1, x_2]\begin{bmatrix} 0 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} x_1-1 \\ x_2 \end{bmatrix} + \cdots \\ &= 1+x_1+x_2-x_1x_2+\frac{1}{2}x_2^2+\cdots. \end{aligned}$$

b. We compute

$$\begin{aligned} Df(\boldsymbol{x}) &= [4x_1^3+4x_1x_2^2, 4x_1^2x_2+4x_2^3], \\ D^2f(\boldsymbol{x}) &= \begin{bmatrix} 12x_1^2+4x_2^2 & 8x_1x_2 \\ 8x_1x_2 & 4x_1^2+12x_2^2 \end{bmatrix}. \end{aligned}$$

Expanding f about the point $\boldsymbol{x}_o$ yields

$$\begin{aligned} f(\boldsymbol{x}) &= 4+[8,8]\begin{bmatrix} x_1-1 \\ x_2-1 \end{bmatrix} + \frac{1}{2}[x_1-1, x_2-1]\begin{bmatrix} 16 & 8 \\ 8 & 16 \end{bmatrix}\begin{bmatrix} x_1-1 \\ x_2-1 \end{bmatrix} + \cdots \\ &= 8x_1^2+8x_2^2-16x_1-16x_2+8x_1x_2+12+\cdots. \end{aligned}$$

c. We compute

$$\begin{aligned} Df(\boldsymbol{x}) &= [e^{x_1-x_2}+e^{x_1+x_2}+1, -e^{x_1-x_2}+e^{x_1+x_2}+1], \\ D^2f(\boldsymbol{x}) &= \begin{bmatrix} e^{x_1-x_2}+e^{x_1+x_2} & -e^{x_1-x_2}+e^{x_1+x_2} \\ -e^{x_1-x_2}+e^{x_1+x_2} & e^{x_1-x_2}+e^{x_1+x_2} \end{bmatrix}. \end{aligned}$$

Expanding f about the point $\boldsymbol{x}_o$ yields

$$\begin{aligned} f(\boldsymbol{x}) &= 2+2e+[2e+1,1]\begin{bmatrix} x_1-1 \\ x_2 \end{bmatrix} + \frac{1}{2}[x_1-1, x_2]\begin{bmatrix} 2e & 0 \\ 0 & 2e \end{bmatrix}\begin{bmatrix} x_1-1 \\ x_2 \end{bmatrix} + \cdots \\ &= 1+x_1+x_2+e(1+x_1^2+x_2^2)+\cdots. \end{aligned}$$

6.2 Show that, if $\boldsymbol{x}^*$ is a global minimizer of f over Ω, and $\boldsymbol{x}^* \in \Omega' \subset \Omega$, then $\boldsymbol{x}^*$ is a global minimizer of f over Ω'.

Ans.: Suppose $\boldsymbol{x}^*$ is a global minimizer of f over Ω, and $\boldsymbol{x}^* \in \Omega' \subset \Omega$. Let $\boldsymbol{x} \in \Omega'$. Then, $\boldsymbol{x} \in \Omega$ and therefore $f(\boldsymbol{x}^*) \le f(\boldsymbol{x})$. Hence, $\boldsymbol{x}^*$ is a global minimizer of f over Ω'.

6.5 Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given below:

$$f(\boldsymbol{x}) = \boldsymbol{x}^T \begin{bmatrix} 1 & 2 \\ 4 & 7 \end{bmatrix} \boldsymbol{x} + \boldsymbol{x}^T \begin{bmatrix} 3 \\ 5 \end{bmatrix} + 6$$

a. Find the gradient and Hessian of f at the point $[1,1]^T$.

b. Find the directional derivative of f at $[1,1]^T$ in the direction of maximal rate of increase.

c. Find a point that satisfies the FONC (interior case) for f. Does this point satisfy the SONC (for a minimizer)?

Ans.: a. The gradient and Hessian of f are

$$\begin{aligned}\nabla f(\boldsymbol{x}) &= 2\begin{bmatrix}1 & 3\\ 3 & 7\end{bmatrix}\boldsymbol{x} + \begin{bmatrix}3\\5\end{bmatrix}\\ \boldsymbol{F}(\boldsymbol{x}) &= 2\begin{bmatrix}1 & 3\\ 3 & 7\end{bmatrix}.\end{aligned}$$

Hence, $\nabla f([1,1]^T) = [11,25]^T$, and $\boldsymbol{F}([1,1]^T)$ is as shown above.

b. The direction of maximal rate of increase is the direction of the gradient. Hence, the directional derivative with respect to a unit vector in this direction is

$$\left(\frac{\nabla f(\boldsymbol{x})}{\|\nabla f(\boldsymbol{x})\|}\right)^T \nabla f(\boldsymbol{x}) = \frac{(\nabla f(\boldsymbol{x}))^T \nabla f(\boldsymbol{x})}{\|\nabla f(\boldsymbol{x})\|} = \|\nabla f(\boldsymbol{x})\|.$$

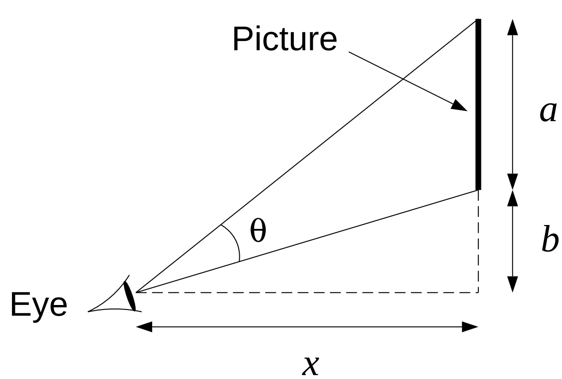
At $\boldsymbol{x} = [1,1]^T$, we have $\|\nabla f([1,1]^T)\| = \sqrt{11^2 + 25^2} = 27.31$.

c. The FONC in this case is $\nabla f(\boldsymbol{x}) = \boldsymbol{0}$. Solving, we get

$$\boldsymbol{x} = \begin{bmatrix}3/2\\ -1\end{bmatrix}.$$

The point above does not satisfy the SONC because the Hessian is not positive semidefinite (its determinant is negative).

6.11 An art collector stands at distance of x feet from the wall where a piece of art (picture) of height a feet is hung, b feet above his eyes, as shown below.



Find the distance from the wall for which the angle θ subtended by the eye to the picture is maximized.
Hint: (1) Maximizing θ is equivalent to maximizing $\tan(\theta)$;
(2) If $\theta = \theta_2 - \theta_1$, then $\tan(\theta) = (\tan(\theta_2) - \tan(\theta_1))/(1 + \tan(\theta_2)\tan(\theta_1))$.

Ans.: Let θ_1 be the angle from the horizontal to the bottom of the picture, and θ_2 the angle from the horizontal to the top of the picture. Then, $\tan(\theta) = (\tan(\theta_2) - \tan(\theta_1))/(1 + \tan(\theta_2)\tan(\theta_1))$. Now, $\tan(\theta_1) = b/x$ and $\tan(\theta_2) = (a+b)/x$. Hence, the objective function that we wish to maximize is

$$f(x) = \frac{(a+b)/x - b/x}{1 + b(a+b)/x^2} = \frac{a}{x + b(a+b)/x}.$$

We have

$$f'(x) = -\frac{a^2}{(x + b(a+b)/x)^2}\left(1 - \frac{b(a+b)}{x^2}\right).$$

Let x^* be the optimal distance. Then, x^* must satisfy the FONC. Now, the point $x^* = 0$ does not satisfy the FONC (why?). Therefore, x^* must be an interior point of the constraint set $\Omega = \{x : x \geq 0\}$. Hence, we have $f'(x^*) = 0$, which gives

$$\begin{aligned} 1 - \frac{b(a+b)}{(x^*)^2} &= 0 \\ \Rightarrow \qquad x^* &= \sqrt{b(a+b)}. \end{aligned}$$

6.20 *Line Fitting*. Let $[x_1, y_1]^T, \ldots, [x_n, y_n]^T$, $n \geq 2$, be points on the $\mathbb{R}^2$ plane (each $x_i, y_i \in \mathbb{R}$). We wish to find the straight line of "best fit" through these points ("best" in the sense that the average squared error is minimized); that is, we wish to find $a, b \in \mathbb{R}$ to minimize

$$f(a,b) = \frac{1}{n}\sum_{i=1}^{n}(ax_i + b - y_i)^2$$

a. Let

$$\begin{aligned} \overline{X} &= \frac{1}{n}\sum_{i=1}^{n} x_i \\ \overline{Y} &= \frac{1}{n}\sum_{i=1}^{n} y_i \\ \overline{X^2} &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 \\ \overline{Y^2} &= \frac{1}{n}\sum_{i=1}^{n} y_i^2 \\ \overline{XY} &= \frac{1}{n}\sum_{i=1}^{n} x_i y_i. \end{aligned}$$

Show that $f(a,b)$ can be written in the form $\boldsymbol{z}^T\boldsymbol{Q}\boldsymbol{z} - 2\boldsymbol{c}^T\boldsymbol{z} + d$, where $\boldsymbol{z} = [a,b]^T$, $\boldsymbol{Q} = \boldsymbol{Q}^T \in \mathbb{R}^{2\times 2}$, $\boldsymbol{c} \in \mathbb{R}^2$, and $d \in \mathbb{R}$, and find expressions for $\boldsymbol{Q}$, $\boldsymbol{c}$, and d in terms of $\overline{X}$, $\overline{Y}$, $\overline{X^2}$, $\overline{Y^2}$, and $\overline{XY}$.

b. Assume that the x_i, $i = 1,\ldots,n$, are not all equal. Find the parameters a^* and b^* for the line of best fit in terms of $\overline{X}$, $\overline{Y}$, $\overline{X^2}$, $\overline{Y^2}$, and $\overline{XY}$. Show that the point $[a^*, b^*]^T$ is the only local minimizer of f.
Hint: $\overline{X^2} - (\overline{X})^2 = (1/n)\sum_{i=1}^n (x_i - \overline{X})^2$.

c. Show that if a^* and b^* are the parameters of the line of best fit, then $\overline{Y} = a^*\overline{X} + b^*$ (and hence once we have computed a^*, we can compute b^* using the formula $b^* = \overline{Y} - a^*\overline{X}$).

Ans.: a. We write

$$\begin{aligned}
f(a,b) &= \frac{1}{n}\sum_{i=1}^n a^2x_i^2 + b^2 + y_i^2 + 2x_iab - 2x_iy_ia - 2y_ib \\
&= a^2\left(\frac{1}{n}\sum_{i=1}^n x_i^2\right) + b^2 + 2\left(\frac{1}{n}\sum_{i=1}^n x_i\right)ab \\
&\quad - 2\left(\frac{1}{n}\sum_{i=1}^n x_iy_i\right)a - 2\left(\frac{1}{n}\sum_{i=1}^n y_i\right)b + \left(\frac{1}{n}\sum_{i=1}^n y_i^2\right) \\
&= [a\ b]\begin{bmatrix} \frac{1}{n}\sum_{i=1}^n x_i^2 & \frac{1}{n}\sum_{i=1}^n x_i \\ \frac{1}{n}\sum_{i=1}^n x_i & 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} \\
&\quad - 2\left[\frac{1}{n}\sum_{i=1}^n x_iy_i, \frac{1}{n}\sum_{i=1}^n y_i\right]\begin{bmatrix} a \\ b \end{bmatrix} + \frac{1}{n}\sum_{i=1}^n y_i^2 \\
&= \boldsymbol{z}^T\boldsymbol{Q}\boldsymbol{z} - 2\boldsymbol{c}^T\boldsymbol{z} + d,
\end{aligned}$$

where $\boldsymbol{z}$, $\boldsymbol{Q}$, $\boldsymbol{c}$ and d are defined in the obvious way.

b. If the point $\boldsymbol{z}^* = [a^*, b^*]^T$ is a solution, then by the FONC, we have $\nabla f(\boldsymbol{z}^*) = 2\boldsymbol{Q}\boldsymbol{z}^* - 2\boldsymbol{c} = \boldsymbol{0}$, which means $\boldsymbol{Q}\boldsymbol{z}^* = \boldsymbol{c}$. Now, since $\overline{X^2} - (\overline{X})^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \overline{X})^2$, and the x_i are not all equal, then $\det \boldsymbol{Q} = \overline{X^2} - (\overline{X})^2 \neq 0$. Hence, $\boldsymbol{Q}$ is nonsingular, and hence

$$\boldsymbol{z}^* = \boldsymbol{Q}^{-1}\boldsymbol{c} = \frac{1}{\overline{X^2} - (\overline{X})^2}\begin{bmatrix} 1 & -\overline{X} \\ -\overline{X} & \overline{X^2} \end{bmatrix}\begin{bmatrix} \overline{XY} \\ \overline{Y} \end{bmatrix} = \begin{bmatrix} \frac{\overline{XY} - (\overline{X})(\overline{Y})}{\overline{X^2} - (\overline{X})^2} \\ \frac{(\overline{X^2})(\overline{Y}) - (\overline{X})(\overline{XY})}{\overline{X^2} - (\overline{X})^2} \end{bmatrix}.$$

Since $\boldsymbol{Q} > 0$, then by the SOSC, the point $\boldsymbol{z}^*$ is a strict local minimizer. Since $\boldsymbol{z}^*$ is the only point satisfying the FONC, then $\boldsymbol{z}^*$ is the only local minimizer.

c. We have

$$a^*\overline{X} + b^* = \left(\frac{\overline{XY} - (\overline{X})(\overline{Y})}{\overline{X^2} - (\overline{X})^2}\right)\overline{X} + \frac{(\overline{X^2})(\overline{Y}) - (\overline{X})(\overline{XY})}{\overline{X^2} - (\overline{X})^2} = \overline{Y}.$$