Mathematical Concepts 2

Solution 1:
Matrix Calculus

(a) $\frac{\partial \|\mathbf{x}-\mathbf{c}\|_2^2}{\partial \mathbf{x}} = \frac{\partial \|\mathbf{u}\|_2^2}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{x}} = \frac{\partial \mathbf{u}^\top \mathbf{u}}{\partial \mathbf{u}} \frac{\partial \mathbf{x}-\mathbf{c}}{\partial \mathbf{x}} = \frac{\partial \mathbf{u}^\top \mathbf{I} \mathbf{u}}{\partial \mathbf{u}} (\mathbf{I} - \mathbf{0}) = \mathbf{u}^\top (\mathbf{I} + \mathbf{I}^\top) = 2(\mathbf{x}-\mathbf{c})^\top$

(b) $\frac{\partial \|\mathbf{x}-\mathbf{c}\|_2}{\partial \mathbf{x}} = \frac{\partial \sqrt{\|\mathbf{x}-\mathbf{c}\|_2^2}}{\partial \mathbf{x}} = \frac{0.5}{\sqrt{\|\mathbf{x}-\mathbf{c}\|_2^2}} \frac{\partial \|\mathbf{x}-\mathbf{c}\|_2^2}{\partial \mathbf{x}} \overset{(a)}{=} \frac{(\mathbf{x}-\mathbf{c})^\top}{\|\mathbf{x}-\mathbf{c}\|_2}$

(c) $\frac{\partial \mathbf{u}^\top \mathbf{v}}{\partial \mathbf{x}} = \frac{\partial \mathbf{u}^\top \mathbf{I} \mathbf{v}}{\partial \mathbf{x}} = \mathbf{u}^\top \mathbf{I} \frac{\partial \mathbf{v}}{\partial \mathbf{x}} + \mathbf{v}^\top \mathbf{I}^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}} = \mathbf{u}^\top \frac{\partial \mathbf{v}}{\partial \mathbf{x}} + \mathbf{v}^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}}$

(d) $\frac{\partial \mathbf{Y}^\top \mathbf{u}}{\partial \mathbf{x}} = \frac{\partial \begin{pmatrix} \mathbf{y}_1^\top \mathbf{u} \\ \vdots \\ \mathbf{y}_d^\top \mathbf{u} \end{pmatrix}}{\partial \mathbf{x}} \overset{(c)}{=} \begin{pmatrix} \mathbf{y}_1^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \mathbf{u}^\top \frac{\partial \mathbf{y}_1}{\partial \mathbf{x}} \\ \vdots \\ \mathbf{y}_d^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \mathbf{u}^\top \frac{\partial \mathbf{y}_d}{\partial \mathbf{x}} \end{pmatrix}$

(e) Note for $\mathbf{y} : \mathbb{R}^d \to \mathbb{R}^d, \mathbf{x} \mapsto \mathbf{y}(\mathbf{x})$ the $i-$th column of $\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$ is $\frac{\partial \mathbf{y}}{\partial x_i}$. With this it follows that

$$
\begin{aligned}
\frac{\partial^2 \mathbf{u}^\top \mathbf{v}}{\partial \mathbf{x} \partial \mathbf{x}^\top} &= \frac{\partial}{\partial \mathbf{x}} \left(\frac{\partial \mathbf{u}^\top \mathbf{v}}{\partial \mathbf{x}^\top} \right) \\
&= \frac{\partial}{\partial \mathbf{x}} \left[\left(\frac{\partial \mathbf{u}^\top \mathbf{v}}{\partial \mathbf{x}} \right)^\top \right] \\
&\overset{(c)}{=} \frac{\partial (\mathbf{u}^\top \frac{\partial \mathbf{v}}{\partial \mathbf{x}} + \mathbf{v}^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}})^\top}{\partial \mathbf{x}} \\
&= \frac{\partial \left(\left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right)^\top \mathbf{u} + \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right)^\top \mathbf{v} \right)}{\partial \mathbf{x}} \\
&\overset{(d)}{=} \begin{pmatrix} \mathbf{u}^\top \frac{\partial^2 \mathbf{v}}{\partial x_1 \partial \mathbf{x}} + \frac{\partial \mathbf{v}}{\partial x_1}^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \\ \vdots \\ \mathbf{u}^\top \frac{\partial^2 \mathbf{v}}{\partial x_d \partial \mathbf{x}} + \frac{\partial \mathbf{v}}{\partial x_d}^\top \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \end{pmatrix} + \begin{pmatrix} \mathbf{v}^\top \frac{\partial^2 \mathbf{u}}{\partial x_1 \partial \mathbf{x}} + \frac{\partial \mathbf{u}}{\partial x_1}^\top \frac{\partial \mathbf{v}}{\partial \mathbf{x}} \\ \vdots \\ \mathbf{v}^\top \frac{\partial^2 \mathbf{u}}{\partial x_d \partial \mathbf{x}} + \frac{\partial \mathbf{u}}{\partial x_d}^\top \frac{\partial \mathbf{v}}{\partial \mathbf{x}} \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{u}^\top \frac{\partial^2 \mathbf{v}}{\partial x_1 \partial \mathbf{x}} \\ \vdots \\ \mathbf{u}^\top \frac{\partial^2 \mathbf{v}}{\partial x_d \partial \mathbf{x}} \end{pmatrix} + \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right)^\top + \frac{\partial \mathbf{v}}{\partial \mathbf{x}} \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)^\top + \begin{pmatrix} \mathbf{v}^\top \frac{\partial^2 \mathbf{u}}{\partial x_1 \partial \mathbf{x}} \\ \vdots \\ \mathbf{v}^\top \frac{\partial^2 \mathbf{u}}{\partial x_d \partial \mathbf{x}} \end{pmatrix}
\end{aligned}
$$

Solution 2:
Optimality in 1d

Let $f : [-1, 2] \to \mathbb{R}, x \mapsto \exp(x^3 - 2x^2)$

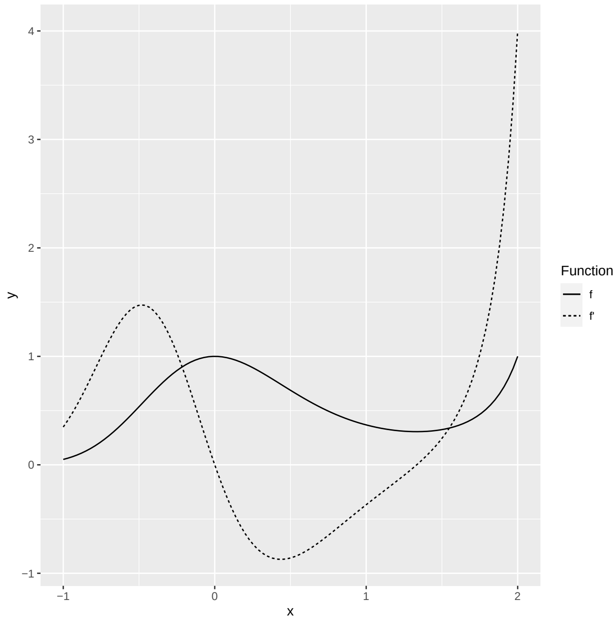
(a) $f'(x) = \exp(x^3 - 2x^2) \cdot (3x^2 - 4x)$

(b)
```
library(ggplot2)

f <- function(x) exp(x^3 - 2*x^2)
df <- function(x) f(x) * (3*x^2 - 4*x)

ggplot(data.frame(x = seq(-1, 2, by=0.005)), aes(x)) +
```

```
  geom_function(fun = f, aes(linetype = "f")) +
  geom_function(fun = df, aes(linetype = "f'"))    +
  scale_linetype_discrete(name = "Function")
```



(c) f is continuously differentiable $\Rightarrow$ candidates can only be stationary points and boundary points.
Find stationary points, i.e., points where
$f'(x) = 0 \iff \underbrace{\exp(x^3 - 2x^2)}_{>0} \cdot (3x^2 - 4x) = 0 \iff 3x^2 - 4x = 0 \iff x(3x-4) = 0.$
$\Rightarrow x_1 = 0, x_2 = 4/3$. The other candidates are boundary points, i.e., $x_3 = -1, x_4 = 2$.

(d) $f''(x) = \exp(x^3 - 2x^2) \cdot (3x^2 - 4x)^2 + \exp(x^3 - 2x^2) \cdot (6x - 4)$

(e) $f''(x_1) = \exp(0) \cdot (-4) < 0$
$\Rightarrow x_1$ is a local maximum

$f''(x_2) = \exp((4/3)^3 - 2(4/3)^2) \cdot (4) > 0$
$\Rightarrow x_2$ is a local minimum.

The boundary points x_3 and x_4 are not considered as *local* optima.

(f) $f(x_1) = \exp(0) = 1$
$f(x_2) = \exp((4/3)^3 - 2(4/3)^2) \approx 0.3057$
$f(x_3) = \exp(-3) \approx 0.05$
$f(x_4) = \exp(0) = 1$
$\Rightarrow x_1, x_4$ are global maxima. x_3 is global minimum.