Elementary Concepts in Numerical Analysis

Differentiation

5

a) Prove that the function $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = x$ is continuous and differentiable.

Solution:
As differentiability implies continuity, we only need to show that f is differentiable. For a function to be differentiable, the following equation

$$\forall x \in \mathbb{R} \quad \exists \quad \lim_{h \to 0} \frac{f(x_0 + \varepsilon) - f(x_0)}{\varepsilon}$$

needs to be satisfied. We thus obtain

$$\lim_{h \to 0} \frac{f(x_0 + \varepsilon) - f(x_0)}{\varepsilon} = \lim_{h \to 0} \frac{x_0 + \varepsilon - x_0}{\varepsilon} = \lim_{h \to 0} \frac{\varepsilon}{\varepsilon} = \lim_{h \to 0} 1 = 1.$$

3

b) Show mathematically why $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = |x|$ is not differentiable at $x_0 = 0$.

Solution:
We separately calculate the derivative from left and right at $x_0 = 0$. Since these have different values, we conclude that no unique limit exists and the function $|\cdot|$ is not differentiable.

Derivative from left:

$$\lim_{\varepsilon \mapsto 0, \varepsilon > 0} \frac{f(x_0 + \varepsilon) - f(x_0)}{\varepsilon} = \lim_{\varepsilon \mapsto 0, \varepsilon > 0} \frac{|0 + \varepsilon| - |0|}{\varepsilon} = \lim_{\varepsilon \mapsto 0, \varepsilon > 0} \frac{|\varepsilon|}{\varepsilon} = 1$$

Derivative from right:

$$\lim_{\varepsilon \mapsto 0, \varepsilon < 0} \frac{f(x_0 + \varepsilon) - f(x_0)}{\varepsilon} = \lim_{\varepsilon \mapsto 0, \varepsilon < 0} \frac{|0 + \varepsilon| - |0|}{\varepsilon} = \lim_{\varepsilon \mapsto 0, \varepsilon < 0} \frac{|\varepsilon|}{\varepsilon} = -1$$

1

c) Use the chain rule (amongst others) to calculate the derivative of $2x^3 + x^2 \log 4x$.

Solution:

$$
\begin{aligned}
&\frac{\mathrm{d}(2x^3 + x^2 \log 4x)}{\mathrm{d}x} \\
=&\frac{\mathrm{d}(2x^3)}{\mathrm{d}x} + \frac{\mathrm{d}(x^2 \log 4x)}{\mathrm{d}x} \\
=&6x^2 + 2x \log(4x) + x^2 \frac{\mathrm{d}(\log 4x)}{\mathrm{d}x} \\
=&6x^2 + 2x \log(4x) + 4x^2 \frac{1}{4x} \\
=&6x^2 + 2x \log(4x) + x
\end{aligned}
$$

3

d) Use R to derive the following expressions symbolically:

- $\frac{\mathrm{d}}{\mathrm{d}x}\left(x^4 + 2x^2 + \exp x\right)$
- $\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(x^4 + 2x^2 + \exp x\right)$

Calculate the value of the first derivative at point $x = 3$

Solution:

```
fx <- expression(x^4+2*x^2+exp(x))

D(fx, "x")

## 4 * x^3 + 2 * (2 * x) + exp(x)

D(D(fx, "x"), "x")

## 4 * (3 * x^2) + 2 * 2 + exp(x)

eval(D(fx, "x"), list(x=3))

## [1] 140.0855
```

3

e) Use R to derive the following expressions symbolically:

- $\frac{\mathrm{d}}{\mathrm{d}x}\left[x^4 + 2x^2 + \exp x\right]$
- $\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left[x^4 + 2x^2 + \exp x\right]$

Calculate the value of the first derivative at point $x = 3$.

Solution:

```
fx <- expression(x^4+2*x^2+exp(x))

D(fx, "x")

## 4 * x^3 + 2 * (2 * x) + exp(x)

D(D(fx, "x"), "x")

## 4 * (3 * x^2) + 2 * 2 + exp(x)

eval(D(fx, "x"), list(x=3))

## [1] 140.0855
```

3

f) Approximate the first derivative at point $x = 0.1$ via forward, backward and centered differences for

$$f(x) = 3x^4 + 2x \qquad \text{and} \qquad g(x) = \sin\frac{1}{x}.$$

Subsequently, compare the approximated derivatives to their true values. Use $h = 10^{-6}$.

Solution:

```
h <- 10^(-6)
x <- 0.1

f <- function(x) 3*x^4 + 2*x
# forward differences
(f(x+h) - f(x)) / h

## [1] 2.012

# backward differences
(f(x) - f(x-h)) / h

## [1] 2.012

# centered differences
(f(x+h) - f(x-h)) / (2*h)
```

```
## [1] 2.012

# true derivative
eval(D(expression(3*x^4 + 2*x), "x"), list(x=0.1))

## [1] 2.012

g <- function(x) sin(1/x)
# forward differences
(g(x+h) - g(x)) / h

## [1] 83.90903

# backward differences
(g(x) - g(x-h)) / h

## [1] 83.90527

# centered differences
(g(x+h) - g(x-h)) / (2*h)

## [1] 83.90715

# true derivative
eval(D(expression(sin(1/x)), "x"), list(x=0.1))

## [1] 83.90715
```

While the results are fairly equal for the first digits in case of f, the results differ for function g. Here, we see numerical errors of minor magnitude since this function is numerically more unstable. Usually, central differences provide the best approximation; this is also the case above.

1

g) Calculate the 2nd order central differences at point $x = 0.01$ for

$$g(x) = \sin\frac{1}{x}.$$

and compare to the true value from a symbolic differentiation. Use $h = 10^{-6}$.

Solution:

```
h <- 10^(-6)
x = 0.01

g <- function(x) sin(1/x)

# 2nd order central differences
(g(x+h) - 2*g(x) + g(x-h)) / h^2

## [1] 52360695

# true derivative
eval(D(D(expression(sin(1/x)), "x"), "x"), list(x=0.01))

## [1] 52361202
```

We see that the numerical approximation for g is here numerically very unstable; resulting in larger errors.

2

h) Derive a formulate to approximate a Hessian matrix of a function $f(x, y)$ using forward differences. Let h be an arbitrary step size.

Solution:
We approximate the Hessian Matrix of the function $f(x, y)$ at a given point (x_0, y_0). Then, we obtain

$$\nabla^2 f(x, y) = \begin{bmatrix} \frac{f(x+2h,y)-2f(x+h,y)+f(x,y)}{h^2} & \frac{f(x+h,y+h)-f(x+h,y)-f(x,y+h)+f(x,y)}{h^2} \\ \frac{f(x+h,y+h)-f(x+h,y)-f(x,y+h)+f(x,y)}{h^2} & \frac{f(x,y+2h)-2f(x,y+h)+f(x,y)}{h^2} \end{bmatrix}.$$

1

i) Write a user-defined function that can approximate a Hessian matrix of $f(x, y)$ using forward differences. Test your function with $f(x, y) = 2x + 3xy^2 + y^3 + 1$ at a point $(x_1, x_2) = (4, 5)$ with a given step size $h = 0.0001$.
Note: You can also pass function names as arguments; see the following example.

```
f <- function(x) x^2

g <- function(x, func) func(x)+2
g(3, f)

## [1] 11
```

Solution:

```
hessian <- function(f, x, y) {
  h <- 10^(-6)
  result <- matrix(rep(0, 4), nrow=2)

  result[1,1] <- (f(x+2*h, y) - 2*f(x+h, y) + f(x, y)) / h^2
  result[2,1] <- (f(x+h, y+h) - f(x+h, y) - f(x, y+h) + f(x, y)) / h^2
  result[1,2] <- (f(x+h, y+h) - f(x+h, y) - f(x, y+h) + f(x, y)) / h^2
  result[2,2] <- (f(x, y+2*h) - 2*f(x, y+h) + f(x, y)) / h^2

  return (result)
}

f <- function(x, y) (2*x + 3*x*y^2 - y^3 + 1)
hessian(f, 4, 5)

##              [,1]      [,2]
## [1,]  0.05684342 30.013325
## [2,] 30.01332516 -5.996981
```

1

j) Consider the following function $f(x, y) = 2x + 3xy^2 + y^3 + 1$. Use the function `optimHess` in R to approximate the Hessian matrix, i. e. with forward differences at a point $(x_1, x_2) = (4, 5)$ with a given step size $h = 0.0001$.

Solution:

```
f <- function(x) (2*x[1] + 3*x[1]*x[2]^2 - x[2]^3 + 1)
optimHess(c(4, 5), f, control=(ndeps=0.0001))

##               [,1] [,2]
## [1,] -1.421085e-08   30
## [2,]  3.000000e+01   -6
```

Taylor Approximation

1

k) Define a Taylor series mathematically.

Solution:
Taylor series is a representation of a function f as an infinite sum of terms that are calculated from the values of the function's derivatives at a single point via

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n.$$

1

l) What is a Maclaurin series?

Solution:
If the Taylor series is centered around $x_0 = 0$, then the series

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}(x)^n$$

is also called a **Maclaurin series**.

1

m) What is the Taylor series for e^x with $x_0 = 0$?

Solution:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad \text{for all } x$$

1

n) What is the Taylor series for $\ln 1 - x$ with $x_0 = 0$?

Solution:

$$\ln(1 - x) = -\sum_{n=1}^{\infty} \frac{x^n}{n} \quad \text{for } |x| < 1$$

1

o) What is the Taylor series for $\ln 1 + x$ with $x_0 = 0$?

Solution:

$$\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} \quad \text{for } |x| < 1$$

1

p) What is the Taylor series for $\frac{1}{1-x}$ with $x_0 = 0$?

Solution:

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots = \sum_{n=0}^{\infty} x^n$$

4

q) What is the 2-nd order Taylor approximation of a function $f(x, y) = \ln 1 + x{+}\ln 1 - y$ around $(x_0, y_0) = (0, 0)$?

Hint: a Taylor Series in 2 variables can be written as

$$\begin{aligned} f(x,y) = f(x_0,y_0) &+ \left[(x-x_0)\frac{\partial}{\partial x} + (y-y_0)\frac{\partial}{\partial y}\right] f(x_0,y_0) \\ &+ \frac{1}{2!}\left[(x-x_0)\frac{\partial}{\partial x} + (y-y_0)\frac{\partial}{\partial y}\right]^2 f(x_0,y_0) + \ldots \end{aligned}$$

Solution:

$$\begin{aligned} f(0,0) &= 0 \\ f_x(0,0) &= \left.\frac{1}{1+x}\right|_{(0,0)} = 1 \\ f_y(0,0) &= -\left.\frac{1}{1-y}\right|_{(0,0)} = -1 \\ f_{xx}(0,0) &= -\left.\frac{1}{(1+x)^2}\right|_{(0,0)} = -1 \\ f_{yy}(0,0) &= -\left.\frac{1}{(1-y)^2}\right|_{(0,0)} = -1 \\ f_{xy}(0,0) &= 0 \\ f(x,y) &= 0 + 1x + (-1)y + \frac{1}{2!}(-1)x^2 + \frac{1}{2!}(-1)y^2 + 0xy \\ &= x - y - \frac{1}{2}x^2 - \frac{1}{2}y^2 \end{aligned}$$

4

r) What is the 2-nd order Taylor approximation of a function

$$f(x, y) = \sqrt[5]{x^3 + e^y}$$

around $(x_0, y_0) = (0, 0)$?

Solution:

$$
\begin{aligned}
f(0,0) &= 1 \\
f_x(0,0) &= \frac{1}{5}\left(x^3+e^y\right)^{-\frac{4}{5}} 3x^2\Big|_{(0,0)} = 0 \\
f_y(0,0) &= \frac{1}{5}\left(x^3+e^y\right)^{-\frac{4}{5}} e^y\Big|_{(0,0)} = \frac{1}{5} \\
f_{xx}(0,0) &= \frac{6x}{5}\left(x^3+e^y\right)^{-\frac{4}{5}} e^y - \frac{3x^2}{5}\left(-\frac{4}{5}\right)\left(x^3+e^y\right)^{-\frac{9}{5}} 3x^2\Big|_{(0,0)} = 0 \\
f_{yy}(0,0) &= \frac{1}{5}\left(x^3+e^y\right)^{-\frac{4}{5}} e^y - \frac{4e^{2y}}{25}\left(x^3+e^y\right)^{-\frac{9}{5}}\Big|_{(0,0)} = \frac{1}{5} - \frac{4}{25} = \frac{1}{25} \\
f_{xy}(0,0) &= -\frac{4}{25}\left(x^3+e^y\right)^{-\frac{9}{5}} 3x^2 e^y\Big|_{(0,0)} = 0 \\
f(x,y) &= 1 + 0x + \frac{1}{5}y + \frac{1}{2!}0x^2 + \frac{1}{2!}\frac{1}{25}y^2 + 0xy \\
&= 1 + \frac{1}{5}y + \frac{1}{50}y^2
\end{aligned}
$$

2

s) Calculate the Taylor approximation of $f(x) = \sin x$ up to degree 4 around $x_0 = 0$. Then evaluate and compare it for $x = 0.1$.

Solution:

```
library(pracma)

##
## Attaching package:  'pracma'
## The following object is masked _by_ '.GlobalEnv':
##
##     hessian
```

```
f <- function(x) sin(x)
taylor.poly <- taylor(f, x0=0, n=4)
taylor.poly

## [1] -0.1666666  0.0000000  1.0000000  0.0000000

polyval(taylor.poly, 0.1) # x = 0.1

## [1] 0.09983333

sin(0.1) # for comparison

## [1] 0.09983342

polyval(taylor.poly, 0.5) - sin(0.5)

## [1] -0.0002588691
```

1

t) Visualize the function $f(x) = \log x + 1$ and its Taylor approximation for $x_0 = 0$.

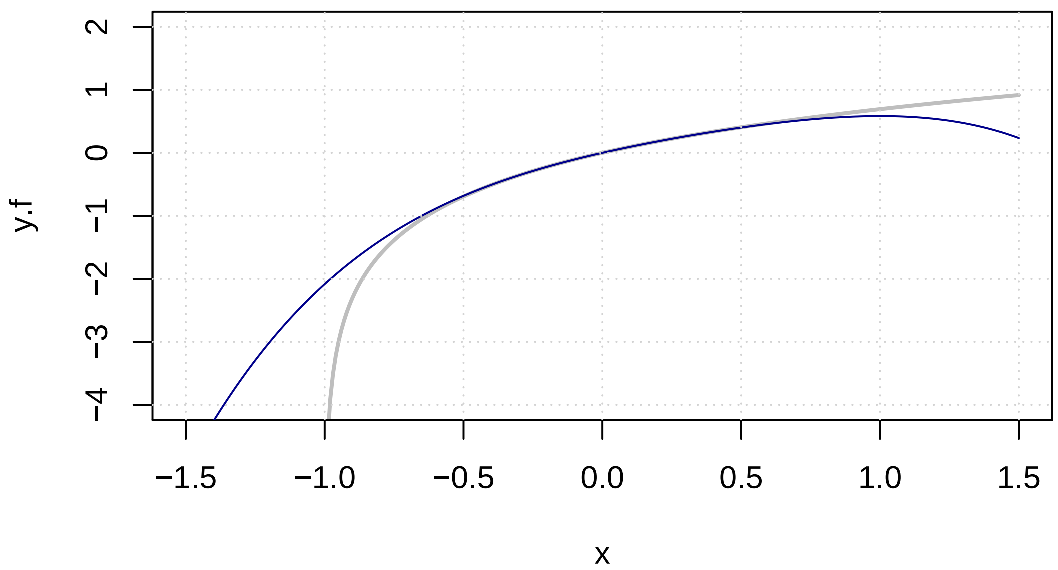
Solution:

```
f <- function(x) log(x+1)

taylor.poly <- taylor(f, x0=0, n=4)
x <- seq(-1.5, 1.5, by=0.01)
y.f <- f(x)

## Warning in log(x + 1):  NaNs produced

y.taylor <- polyval(taylor.poly, x)
plot(x, y.f, type="l", col="gray", lwd=2, ylim=c(-4, +2))
lines(x, y.taylor, col="darkblue")
grid()
```



Optimality Conditions

9

u) Find the stationary points of the function $f(x, y) = x^3 + 3y - y^3 - 3x$ and analyze their nature using Sylvester's Rule.

Solution:

1. We first utilize the first order optimality condition to find stationary points, i. e. those points where $\nabla f = 0$ is satisfied. Accordingly, we obtain

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)^T \stackrel{!}{=} 0$$

and

$$\begin{aligned} \frac{\partial f}{\partial x} &= 3x^2 - 3 = 0 \qquad &\Rightarrow \qquad & x = \pm 1, \\ \frac{\partial f}{\partial y} &= 3 - 3y^2 = 0 \qquad &\Rightarrow \qquad & y = \pm 1. \end{aligned}$$

Consequently, all four stationary points are given by

$$\begin{aligned}
\boldsymbol{p}_1 &= (1,1)^T, \\
\boldsymbol{p}_2 &= (-1,1)^T, \\
\boldsymbol{p}_3 &= (1,-1)^T, \\
\boldsymbol{p}_4 &= (-1,-1)^T.
\end{aligned}$$

2. Next, we calculate the Hessian matrix

$$\nabla^2 f = H = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 6x & 0 \\ 0 & -6y \end{pmatrix},$$

since we need this in order to understand the definiteness.

3. Now, we proceed to analyze the nature of point $\boldsymbol{p}_1 = (1,1)^T$. The Hessian and its determinant resolve to

$$\begin{aligned}
H = \nabla^2 f(\boldsymbol{p}_1) &= \begin{pmatrix} 6 & 0 \\ 0 & -6 \end{pmatrix}, \\
\det H_1 &= 6 > 0, \\
\det H_2 &= \det H = -36 < 0.
\end{aligned}$$

Hence, the Hessian matrix H is indefinite and $\boldsymbol{p}_1$ is a saddle point with value $f(\boldsymbol{p}_1) = f(1,1) = 0$.

4. The nature of point $\boldsymbol{p}_2 = (-1,1)^T$ is based on its Hessian matrix, as well as the determinant, given by

$$\begin{aligned}
H = \nabla^2 f(\boldsymbol{p}_2) &= \begin{pmatrix} -6 & 0 \\ 0 & -6 \end{pmatrix}, \\
\det H_1 &= -6 < 0, \\
\det H_2 &= \det H = 36 > 0.
\end{aligned}$$

The Hessian matrix H is thus negative definite; the point $\boldsymbol{p}_2$ represents a local maximum with value $f(\boldsymbol{p}_2) = f(-1,1) = 4$.

5. The nature of point $\boldsymbol{p}_3 = (1, -1)^T$ is given by

$$
\begin{aligned}
H = \nabla^2 f(\boldsymbol{p}_3) = \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix}, \\
\det H_1 = 6 > 0, \\
\det H_2 = \det H = 36 > 0,
\end{aligned}
$$

   where the Hessian matrix H is positive definite. Accordingly, we obtain a local minimum at $\boldsymbol{p}_3$ with value $f(\boldsymbol{p}_3) = f(1, -1) = -4$.

6. Similarly, the nature of point $\boldsymbol{p}_4 = (-1, -1)^T$ is

$$
\begin{aligned}
H = \nabla f(\boldsymbol{p}_4) = \begin{pmatrix} -6 & 0 \\ 0 & 6 \end{pmatrix}, \\
\det H_1 = -6 < 0, \\
\det H_2 = \det H = -36 < 0.
\end{aligned}
$$

   The Hessian matrix H is thus indefinite and $\boldsymbol{p}_4$ is a saddle point with value $f(\boldsymbol{p}_4) = f(-1, -1) = 0$. As an alternative approach, we could have also studied the eigenvalues of the Hessian matrix at point $\boldsymbol{p}_4$ via

$$
\begin{aligned}
& \det(H - \lambda I) = 0 \\
\Leftrightarrow & \begin{vmatrix} -6 - \lambda & 0 \\ 0 & 6 - \lambda \end{vmatrix} = 0 \\
\Leftrightarrow & (-6 - \lambda)(6 - \lambda) && = 0 \\
\Leftrightarrow & \lambda_1 = -6 \quad \text{and} \quad \lambda_2 = 6.
\end{aligned}
$$

   Given the above derivation, all eigenvalues satisfy $\lambda_i \neq 0$ and the Hessian matrix H is thus indefinite.

9

v) Consider the function $f(x, y) = \sin x \cdot \cos x$. First of all, plot the function nicely to get an impression of its curvature. Then, consider the points

$$
\boldsymbol{p}_1 = \begin{bmatrix} \frac{\pi}{2} \\ 0 \end{bmatrix} \text{ and } \boldsymbol{p}_2 = \begin{bmatrix} 0 \\ \frac{\pi}{2} \end{bmatrix}
$$

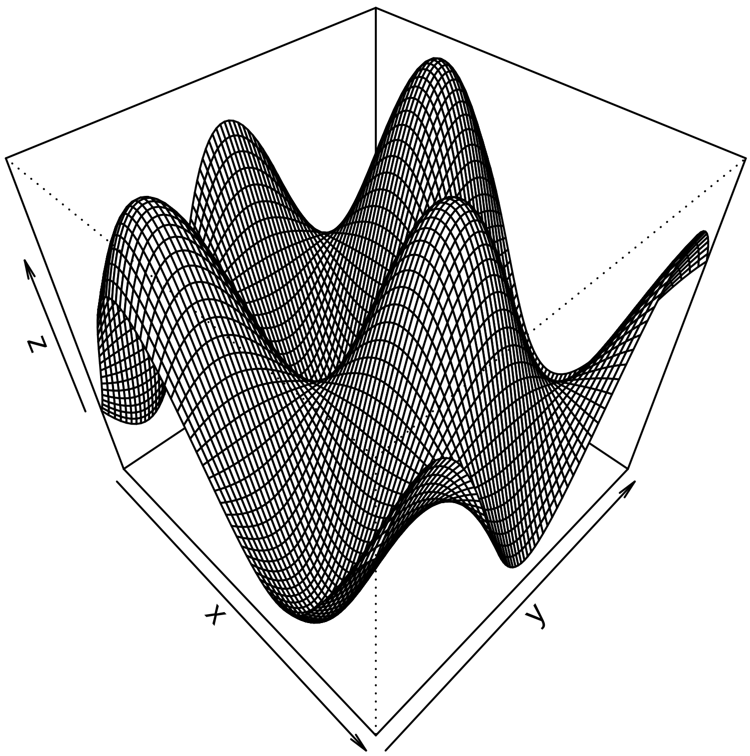
and check their first and second order optimality conditions using R. What type of stationary points do they belong to?

Solution:

Plot function in 3D

```
f <- function(x, y) sin(x)*cos(y)
x <- seq(-4, 4, 0.1)
y <- seq(-4, 4, 0.1)
z <- outer(x, y, f)

persp(x, y, z, theta=45, phi=45)
```




```
library(matrixcalc)
f <- expression(sin(x)*cos(y))
```

Check point $\boldsymbol{p}_1$ as follows:

- First order conditions are fulfilled

```r
eval(D(f, "x"), list(x=pi/2, y=0))

## [1] 6.123032e-17
```

- Second order conditions

```r
H <- optimHess(c(pi/2, 0),
               function(x) sin(x[1])*cos(x[2]),
               control=(ndeps=0.0001))
is.negative.definite(H)

## [1] TRUE
```

- Point $\boldsymbol{p}_1$ is a local maximum

Check point $\boldsymbol{p}_2$ as follows:

- First order conditions are fulfilled

```r
eval(D(f, "x"), list(x=0, y=pi/2))

## [1] 6.123032e-17
```

- Second order conditions

```r
H <- optimHess(c(0, pi/2),
               function(x) sin(x[1])*cos(x[2]),
               control=(ndeps=0.0001))
is.indefinite(H)

## [1] TRUE
```

- Point $\boldsymbol{p}_2$ is a saddle point

w) Prove that the function $f : \mathbb{R} \rightarrow \mathbb{R}$ with $f(x) = x^2$ is convex. 2

Solution:

A function is convex if

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2) \qquad \forall x_1, x_2 \in \mathbb{R} \quad \forall \alpha \in [0,1].$$

This means that every point on a line between x_1 and x_2 with $x_1 \leq x_2$ lies on or above the function. To prove this, we insert $f(x) = x^2$ and rearrange the above equation as follows

$$
\begin{aligned}
&\Leftrightarrow (\alpha x_1 + (1-\alpha)x_2)^2 \leq \alpha x_1^2 + (1-\alpha)x_2^2 \\
&\Leftrightarrow \alpha^2 x_1^2 + 2\alpha(1-\alpha)x_1 x_2 + (1-\alpha)^2 x_2^2 \leq \alpha x_1^2 + (1-\alpha)x_2^2 \\
&\Leftrightarrow \alpha^2 x_1^2 + 2\alpha(1-\alpha)x_1 x_2 + x_2^2 - 2\alpha x_2^2 + \alpha^2 x_2^2 \leq \alpha x_1^2 + x_2^2 - \alpha x_2^2 \\
&\Leftrightarrow 0 \geq (\alpha^2 - \alpha)x_1^2 + 2\alpha(1-\alpha)x_1 x_2 + (\alpha^2 - \alpha)x_2^2 \\
&\Leftrightarrow 0 \geq (\alpha^2 - \alpha)x_1^2 - 2(\alpha^2 - \alpha)x_1 x_2 + (\alpha^2 - \alpha)x_2^2 \\
&\Leftrightarrow 0 \geq (\alpha^2 - \alpha)(x_1 - x_2)^2.
\end{aligned}
$$

The final inequality is true for all $\alpha \in [0, 1]$ if $x_1 = x_2$, as well as $x_1 \neq x_2$.