Nonlinear Programming 1

Exercise 1: Lagrange multipliers

Solve the constrained optimization problem

$$\begin{aligned} \min_{(x,y)\in\mathbb{R}^2} \quad & x + 2y \\ \text{s.t.} \quad & x^2 + 4y^2 = 4 \end{aligned}$$

by Lagrange multipliers.

Exercise 2: Nonlinear SVM

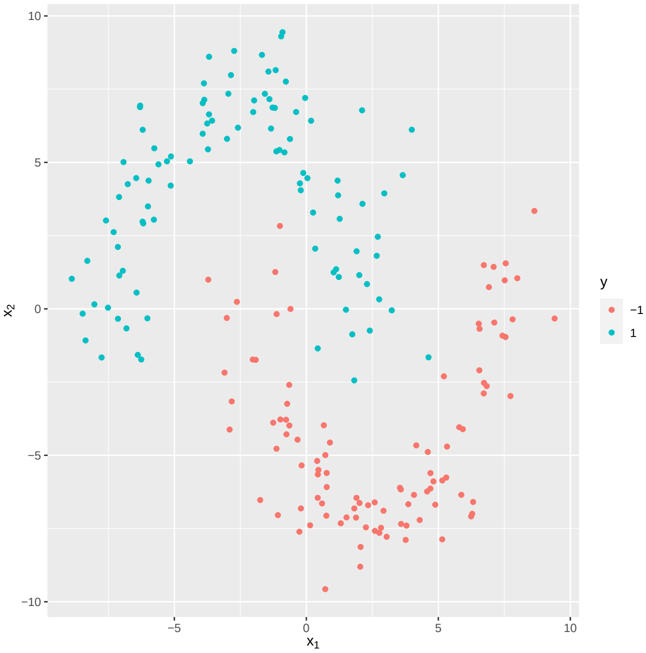
You are given the following data situation:

```
library(ggplot2)
library(mlr3)

# generate 200 nonlinear separable binary observations
set.seed(123)
moon_data = tgen("moons")$generate(200)$data()

moon_data$y = ifelse(moon_data$y == "A", 1, -1)
moon_data$y_dec = as.factor(moon_data$y)

ggplot(moon_data, aes(x=x1, y=x2)) +
  geom_point(aes(color=y_dec)) +
  xlab(expression(x[1])) +
  ylab(expression(x[2])) +
  labs(color=expression(y))
```



We can extend the linear SVM to a nonlinear SVM by transforming the features via a nonlinear transformation

$\phi : \mathbb{R}^d \to \mathbb{R}^l$. With this, the primal form with soft constraints becomes

$$\min_{\boldsymbol{\theta}, \theta_0, \boldsymbol{\zeta}} 0.5\|\boldsymbol{\theta}\|^2 + C \sum_{i=1}^{n} \zeta^{(i)}$$

s.t.

$$y^{(i)}(\langle \boldsymbol{\theta}, \phi(\mathbf{x}^{(i)}) \rangle + \theta_0) \geq 1 - \zeta^{(i)} \quad \forall i \in \{1, \ldots, n\}$$

and

$$\zeta^{(i)} \geq 0 \quad \forall i \in \{1, \ldots, n\}.$$

(a) Write down the general Lagrangian function of the nonlinear SVM

(b) Solve the primal nonlinear SVM for $C = 1$ and third order polynomial transformation (without intercept) $\phi(\mathbf{x}) = (x_1, x_2, x_1^2, x_2^2, x_1^3, x_2^3, x_1 x_2, x_1^2 x_2, x_1 x_2^2)^\top$ with `cvxr` in `R`.
Hint: Examples how quadratic problems can be solved with `cvxr` can be found here.

(c) State the KKT conditions of the general nonlinear SVM

(d) Derive the dual form of the nonlinear SVM. State an advantage of the dual form over the primal form.
Hint: Use the KKT conditions to transform the primal form

(e) Solve the dual form of b) with `cvxr` in `R`.
Hint 1: For a polynomial transformation ϕ of order l (without intercept) it holds that there exists a invertible diagonal matrix $\mathbf{D} \in \mathbb{R}^9$ such that $\langle \mathbf{D}\phi(\mathbf{x}), \mathbf{D}\phi(\mathbf{z}) \rangle = (\mathbf{x}^\top \mathbf{z} + 1)^l - 1$
Hint 2: Add $10^{-7} \cdot \mathbf{I}$ to the kernel matrix to ensure that the resulting matrix is invertible.