# Advanced Statistical Computing

~~Computational Statistics~~

Exercises for week 6: Numerical Integration, Convex Optimization

**Exercise 1** The `R` function `integrate` computes an approximation to the integral of a function by a variant of Gaussian quadrature. Two other methods to find an approximation are stochastic simulation (possibly by importance sampling, as in week 3) and the *trapezium rule*. In this exercise we are interested in computing

$$\frac{1}{\sqrt{2\pi}} \int_1^\infty x^2 e^{-x^2} \, dx.$$

a Write a function `trapeziumrule` that takes arguments `f, lower, upper, h`, equal to a function, lower and upper limits of the integral and a step size, that returns the integral of the function over the interval $[\texttt{lower}, \texttt{upper}]$.

b Use this function to find an approximation to the given integral. Try the 4 step sizes $h = 0.01, 0.001, 0.0001, 0.0001$ and use an appropriate value for `upper`.

c Compute an approximation by using the function `integrate`. Read the help page of the function to see how to handle the upper infinite limit!

**Exercise 2** In *portfolio management* according to Markowitz the problem is to decide on the number of assets of various types to invest in so that the variance of the value of the portfolio ("risk") is small but the expected capital gain is big. Consider $m$ assets $1, \ldots, m$ and let $(v_1, \ldots, v_m)$ be the relative weights of the assets in the portfolio, so that

$$\sum_{i=1}^{m} v_i = 1, \qquad v_i \geq 0, \quad i = 1, \ldots m. \tag{1}$$

The returns (daily changes) of the assets are random variables with known covariance matrix $\Sigma$, giving the change in value of the portfolio $v = (v_1, \ldots, v_m)$ a variance equal to $v^T \Sigma v$. The expected gain on asset $i$ is $\mu_i$ giving the portfolio an expected gain $\sum_{i=1}^m v_i \mu_i = v^T \mu$. Markowitz's approach is to maximize $v^T \mu$, but minimize $v^T \Sigma v$, under the constraint (1). These aims are contradictory in general and must be weighted. The relative importance of maximizing profit versus minimizing risk is a *risk taking profile*, and can be implemented by deciding that overall we minimize under constraint (1), for given $r > 0$,

$$v^T \Sigma v - r \, v^T \mu.$$

a The file `assets.txt` contains the returns of $m = 50$ assets over 300 trading days. Read the data with `x= read.table(``file=assets.txt'')` and estimate $\mu$ and $\Sigma$ by `m=apply(x,1,mean)` and `Sigma=var(t(x))`.

b Determine an optimal portfolio for risk profile $r = 1$ using the function `solve.QP` from library `quadprog`. Read the help page to see how to include an equality constraint.

c Repeat this for $r = 10$.

d Does $r = 10$ give the profile of someone taking more or less risk than $r = 1$?

**Exercise 3** Consider observations $(x_1, y_1), \dots (x_n, y_n)$ in the univariate regression model (without intercept) $Y = \beta x + e$, were $x \in \mathbb{R}$, $\beta \in \mathbb{R}$ and $e$ are univariate errors which are known to have $\tau$-quantile equal to 0, i.e. $\mathrm{P}(e \leq 0) = \tau$, for given $\tau \in (0, 1)$. A reasonable estimator for $\beta$ in this case is the minimizer $\hat{\beta}$ of

$$\sum_{i=1}^{n} \Big[\tau(y_i - x_i\beta)^{+} + (1-\tau)(y_i - x_i\beta)^{-}\Big].$$

Here we use the notation $u^+$ and $u^-$ for the "positive and negative parts of a number" :

$$u^+ = \begin{cases} u & \text{if } u > 0, \\ 0 & \text{if } u \leq 0 \end{cases}, \qquad u^- = \begin{cases} 0 & \text{if } u > 0, \\ -u & \text{if } u \leq 0 \end{cases}.$$

The choice $\tau = 1/2$ gives median absolute deviation regression, but throughout the following, fix $\tau = 3/4$.

a Write a program that computes $\hat{\beta}$, using linear programming with one of the `R` functions `solveLP` or `lp` in the library `linprog`.

b Apply your program to the data set in the file `qrdata.txt`

c Check the validity of your outcome against the output of the standard function `rq` for quantile regression in the package `quantreg`.