Homework 8

Submit answers for problems 1–3.

1. Exercise A7.5
2. Exercise A7.26.
3. Exercise A8.1.
4. *Maximum likelihood estimation from quantized measurements.* Consider the problem of estimating a vector $x \in \mathbf{R}^n$ from noisy linear measurements

$$y_i = a_i^T x + v_i, \quad i = 1, \ldots, m.$$

The random measurement errors v_i are independent and identically distributed with probability density function

$$f(t) = \frac{e^t}{(1+e^t)^2}.$$

In this problem we assume the measurements y_i are not directly observable. Instead they are quantized using three levels, by comparing them with thresholds b_i and c_i, where $b_i < c_i$. The observed, quantized measurements are $s_i \in \{0, 1, 2\}$, $i = 1, \ldots, m$, where

$$\begin{array}{ll} s_i = 0 & \text{if } y_i \leq b_i \\ s_i = 1 & \text{if } b_i < y_i \leq c_i \\ s_i = 2 & \text{if } y_i > c_i. \end{array}$$

Therefore s_i is a discrete random variable with probability distribution

$$\begin{array}{l} \mathbf{prob}(s_i = 0) = F(b_i - a_i^T x) \\ \mathbf{prob}(s_i = 1) = F(c_i - a_i^T x) - F(b_i - a_i^T x) \\ \mathbf{prob}(s_i = 2) = 1 - F(c_i - a_i^T x), \end{array}$$

where F is the cumulative distribution function

$$F(t) = \int_{-\infty}^{t} f(\tau) d\tau = \frac{e^t}{1+e^t}.$$

The log-likelihood function for the data $s_1, \ldots, s_m$ is

$$\begin{aligned} l(x) \;=\; & \sum_{i:s_i=0} \log F(b_i - a_i^T x) + \sum_{i:s_i=1} \log \Big(F(c_i - a_i^T x) - F(b_i - a_i^T x) \Big) \\ & + \sum_{i:s_i=2} \log(1 - F(c_i - a_i^T x)). \end{aligned}$$

(a) Show that the log-likelihood function l is a concave function of x.

(b) Explain why this holds for any log-concave probability density function f.

Solution.

(a) Each term in l is a sum of an affine function and the concave functions $-\log\exp(1+\exp(b_i - a_i^T x))$ or $-\log\exp(1+\exp(c_i - a_i^T x))$:

$$\begin{aligned}
\log F(b_i - a_i^T x) &= \log\left(\frac{\exp(b_i - a_i^T x)}{1+\exp(b_i - a_i^t x)}\right) \\
&= b_i - a_i^T x - \log(1+\exp(b_i - a_i^T x)) \\
\log\Big(F(c_i - a_i^T x) - F(b_i - a_i^T x)\Big) &= \log\left(\frac{\exp(c_i - a_i^T x)}{1+\exp(c_i - a_i^T x)} - \frac{\exp(b_i - a_i^T x)}{1+\exp(b_i - a_i^T x)}\right) \\
&= \log(\exp(c_i - a_i^T x) - \exp(b_i - a_i^T x)) \\
&\quad - \log(1+\exp(c_i - a_i^T x)) - \log(1+\exp(b_i - a_i^T x)) \\
&= \log(e^{c_i} - e^{b_i}) - a_i^T x - \log(1+\exp(c_i - a_i^T x)) \\
&\quad - \log(1+\exp(b_i - a_i^T x)) \\
\log(1 - F(c_i - a_i^T x) &= \log\frac{1}{1+\exp(c_i - a_i^T x)} \\
&= -\log(1+\exp(c_i - a_i^T x)).
\end{aligned}$$

The functions $\log\exp(1+\exp(b_i - a_i^T x))$ and $\log\exp(1+\exp(c_i - a_i^T x))$ are convex because they are compositions of the log-sum-exp function $\log(1+\exp(y))$ with affine functions.

(b) The terms for $s_i = 0$ in l are log-concave, because the cumulative distribution F of a log-concave density f is log-concave (exercise T3.55).
For the terms in the second sum, we note that

$$F(c_i - t) - F(b_i - t) = \int_{b_i - t}^{c_i - t} f(\tau)d\tau = \int_{b_i}^{c_i} f(\tau - t)du = \int_{-\infty}^{\infty} f(\tau - t)g(u)du$$

where $g(\tau) = 1$ for $u \in [b_i, c_i]$ and zero otherwise. The result is log-concave in t by the property on page 3.29 of the slides.
The last terms are log-concave because

$$1 - F(t) = \int_t^{\infty} f(\tau)d\tau = \int_{-\infty}^{-t} f(-\tau)d\tau$$

and $f(-\tau)$ is log-concave, so this again follows from exercise T3.55.