

The condition  $\sum_{i=1}^n y_i = 1$  is obviously necessary, as well as the conditions  $y_i > 0$ , for  $i = 1, \dots, n$ . Conversely, if  $\mathbf{1}^\top y = 1$  and  $y > 0$ , then  $x_j = \log y_i$  for  $i = 1, \dots, n$  is a solution. Since  $(*)$  implies that

$$x_i = \log y_i + \log \left( \sum_{i=1}^n e^{x_i} \right), \quad (**)$$

we get

$$\begin{aligned} y^\top x - f(x) &= \sum_{i=1}^n y_i x_i - \log \left( \sum_{i=1}^n e^{x_i} \right) \\ &= \sum_{i=1}^n y_i \log y_i + \sum_{i=1}^n y_i \log \left( \sum_{i=1}^n e^{x_i} \right) - \log \left( \sum_{i=1}^n e^{x_i} \right) \quad \text{by } (**) \\ &= \sum_{i=1}^n y_i \log y_i + \left( \sum_{i=1}^n y_i - 1 \right) \log \left( \sum_{i=1}^n e^{x_i} \right) \\ &= \sum_{i=1}^n y_i \log y_i \quad \text{since } \sum_{i=1}^n y_i = 1. \end{aligned}$$

Consequently, if  $f^*(y)$  is defined, then  $f^*(y) = \sum_{i=1}^n y_i \log y_i$ . If we agree that  $0 \log 0 = 0$ , then it is an easy exercise (or, see Boyd and Vandenberghe [29], Section 3.3, Example 3.25) to show that

$$f^*(y) = \begin{cases} \sum_{i=1}^n y_i \log y_i & \text{if } \mathbf{1}^\top y = 1 \text{ and } y \geq 0 \\ \infty & \text{otherwise.} \end{cases}$$

Thus we obtain the negative entropy restricted to the domain  $\mathbf{1}^\top y = 1$  and  $y \geq 0$ .

If  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is convex and differentiable, then  $x^*$  maximizes  $x^\top y - f(x)$  iff  $x^*$  minimizes  $-x^\top y + f(x)$  iff

$$\nabla f_{x^*} = y,$$

and so

$$f^*(y) = (x^*)^\top \nabla f_{x^*} - f(x^*).$$

Consequently, if we can solve the equation

$$\nabla f_z = y$$

for  $z$  given  $y$ , then we obtain  $f^*(y)$ .

It can be shown that if  $f$  is twice differentiable, strictly convex, and surlinear, which means that

$$\lim_{\|y\| \rightarrow +\infty} \frac{f(y)}{\|y\|} = +\infty,$$