Exercise: Multi-output linear regression

(Source: Jaakkola.)

Consider a linear regression model with a 2 dimensional response vector $\mathbf{y}_i \in \mathbb{R}^2$. Suppose we have some binary input data, $x_i \in \{0, 1\}$. The training data is as follows:

$$\begin{array}{c|cc} \mathbf{x} & \mathbf{y} \\ \hline \mathbf{0} & (-1,-1)^T \\ \mathbf{0} & (-1,-2)^T \\ \mathbf{0} & (-2,-1)^T \\ \mathbf{1} & (1,1)^T \\ \mathbf{1} & (1,2)^T \\ \mathbf{1} & (2,1)^T \\ \end{array}$$

Let us embed each x_i into 2d using the following basis function:

$$\phi(0) = (1,0)^T, \quad \phi(1) = (0,1)^T$$
 (1)

The model becomes

$$\hat{\mathbf{y}} = \mathbf{W}^T \boldsymbol{\phi}(x) \tag{2}$$

where ${\bf W}$ is a 2×2 matrix. Compute the MLE for ${\bf W}$ from the above data.