Zakladni rovnice kterou chceme naprogramovat.

$$\min_\theta -\mathbb{E}_{p_{\text{data}}(x,y)} \left[\alpha \log q_\theta\left(y|x\right) + (1-\alpha)\log q_\theta\left(x|y\right)\right] \tag{1}$$

$$\approx \min_\theta -\mathbb{E}_{p_{\text{data}}(x,y)} \left[\alpha \log \frac{\exp\left(f_\theta\left(x\right)\left[y\right]\right)}{\sum_{y'}\exp\left(f_\theta\left(x\right)\left[y'\right]\right)} + (1-\alpha)\log \frac{\exp\left(f_\theta\left(x\right)\left[y\right]\right)}{\sum_{i=1}^{K}\exp\left(f_\theta\left(x_i\right)\left[y\right]\right)}\right] \tag{2}$$

Dle znaceni v clanku, **Joint Energy models**

$$p(x,y) = \frac{\exp\left(f_\theta(x)[y]\right)}{Z(\theta)} \tag{3}$$

kde

$$f_\theta(x)[y] = -E_\theta(x,y) \tag{4}$$

a pro normalizacni konstantu plati

$$Z(\theta) = \sum_x\sum_y \exp\left(-E_\theta(x,y)\right) \tag{5}$$

Jak to premostit na jednorozmernou linearni regresi?

$$p(x,y) = p(y,x) = p(y|x)\cdot p(x) \tag{6}$$

$$p(y|x) = \mathcal{N}\left(\theta_0 + \theta_1 x, \sigma^2\right) \tag{7}$$

a $p(x)$ muzeme zvolit libovolne? Napriklad pro jednoduchost

$$p(x) = \mathcal{N}\left(0, \sigma^2\tau^2\right), \tag{8}$$

pricemz uvazujeme σ za zname a τ volime dle potreby. Z cehoz plyne ze

$$p(x,y) = \mathcal{N}\left(0,\sigma^2\tau^2\right)\cdot\mathcal{N}\left(\theta_0+\theta_1 x,\sigma^2\right) = \frac{1}{2\pi\sigma\tau}\exp\left(-\frac{(y-\theta_0-\theta_1 x)^2}{2\sigma^2} - \frac{x^2}{2\sigma^2\tau^2}\right) \tag{9}$$

a tudiz

$$f_\theta(x)[y] = -\frac{(y-\theta_0-\theta_1 x)^2}{2\sigma^2} - \frac{x^2}{2\sigma^2\tau^2}. \tag{10}$$

Nyni dosadme (10) do (2) a pouzijeme pravidla pro pocitani s logaritmy.

$$\min_\theta -\mathbb{E}_{p_{\text{data}}(x,y)} \left[\alpha \log q_\theta\left(y|x\right) + (1-\alpha)\log q_\theta\left(x|y\right)\right] \tag{11}$$

$$\min_\theta -\mathbb{E}_{p_{\text{data}}(x,y)} \left[\alpha \log \frac{\exp\left(f_\theta\left(x\right)\left[y\right]\right)}{\sum_{y'}\exp\left(f_\theta\left(x\right)\left[y'\right]\right)} + (1-\alpha)\log \frac{\exp\left(f_\theta\left(x\right)\left[y\right]\right)}{\sum_{i=1}^{K}\exp\left(f_\theta\left(x_i\right)\left[y\right]\right)}\right] = \tag{12}$$

$$\min_\theta -\mathbb{E}_{p_{\text{data}}(x,y)} \left[\alpha \log \frac{\exp\left(-\frac{(y-\theta_0-\theta_1 x)^2}{2\sigma^2} - \frac{x^2}{2\sigma^2\tau^2}\right)}{\sum_y \exp\left(-\frac{(y-\theta_0-\theta_1 x)^2}{2\sigma^2} - \frac{x^2}{2\sigma^2\tau^2}\right)} + (1-\alpha)\log \frac{\exp\left(-\frac{(y-\theta_0-\theta_1 x)^2}{2\sigma^2} - \frac{x^2}{2\sigma^2\tau^2}\right)}{\sum_x \exp\left(-\frac{(y-\theta_0-\theta_1 x)^2}{2\sigma^2} - \frac{x^2}{2\sigma^2\tau^2}\right)}\right] \tag{13}$$

potom pro prvni q faktor dostaneme

$$\alpha\log q_\theta\left(y|x\right) = \alpha\left(\frac{(y-\theta_0-\theta_1 x)^2}{2\sigma^2} + \frac{x^2}{2\sigma^2\tau^2}\right) - \alpha\log\sum_y \exp\left(-\frac{(y-\theta_0-\theta_1 x)^2}{2\sigma^2} - \frac{x^2}{2\sigma^2\tau^2}\right) \tag{14}$$

a pro druhy

$$(1-\alpha)\log q_\theta\left(x|y\right) = (1-\alpha)\left(\frac{(y-\theta_0-\theta_1 x)^2}{2\sigma^2} + \frac{x^2}{2\sigma^2\tau^2}\right) - (1-\alpha)\log\sum_x \exp\left(-\frac{(y-\theta_0-\theta_1 x)^2}{2\sigma^2} - \frac{x^2}{2\sigma^2\tau^2}\right) \tag{15}$$