DATE
PAGE

$$x_i \to u_{0i} \triangleq w_0^T x_i + \delta_{0i}$$
$$x_i \to u_{1i} \triangleq w_1^T x_i + \delta_{1i}$$

$$y_i = \mathbb{I}(u_{1i} > u_{0i})$$

define $v_i = u_{1i} - u_{0i} + \cancel{\epsilon_i}$ $(\delta_{1i} - \delta_{0i})$

$$z_i \triangleq w^T x_i + \epsilon_i, \quad \epsilon_i \sim N(0,1)$$
$$y_i = 1 = \mathbb{I}(z_i \geq 0)$$

$$P(y_i = 1 \mid x_i, w) = \int \mathbb{I}(z_i \geq 0)\, N(z_i \mid w^T x_i, I)\, dz_i$$
$$= P(w^T x_i + \epsilon \geq 0) = P(\epsilon \geq -w^T x_i)$$
$$= 1 - \phi(-w^T x_i) = \underline{\phi(w^T x_i)}$$

this model can be trained using EM. $P(y_i = 1 \mid z_i) = \mathbb{I}(z_i \geq 0)$
$z_i \sim N(w^T x_i, 1)$ is latent var.

complete log likelihood $\ell(z, w \mid V_0) = \ln(P(y \mid z)) + \ln(z \mid x w, I) + \ln(w \mid 0, V_0)$

$$= \sum_i \ln(P(y_i \mid z_i)) - \frac{1}{2}(z - Xw)^T (z - Xw) - \frac{1}{2} w^T V_0^{-1} w + \text{const.}$$

① the posterior of E-step is a trunc. gaussian

$$P(z_i \mid y_i, x_i, w) = \begin{cases} N(z_i \mid w^T x_i, 1)\, \mathbb{I}(z_i > 0) & \text{if } y_i = 1 \\ N(z_i \mid w^T x_i, 1)\, \mathbb{I}(z_i < 0) & \text{if } y_i = 0 \end{cases}$$

note $\ell(z, w|V_0)$ is linearly dependent on z

so
$$E[z_i|y_i, w, x_i] = \begin{cases} \mu_i + \dfrac{\phi(\mu_i)}{1-\Phi(-\mu_i)} = \mu_i + \dfrac{\phi(\mu_i)}{\Phi(\mu_i)} & y_i=1 \\ \mu_i - \dfrac{\phi(\mu_i)}{\Phi(-\mu_i)} = \mu_i - \dfrac{\phi(\mu_i)}{1-\Phi(\mu_i)} & y_i=0 \end{cases}$$

In the M-step estimate w using ridge regression, where $\mu = E[Z]$ is the output we wish to predict.

$$\hat{w} = (V_0^{-1} + X^T X)^{-1} X^T \mu$$