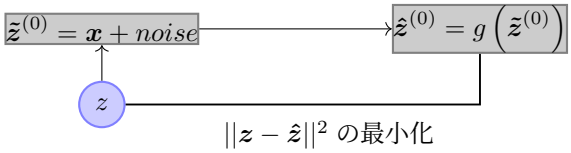
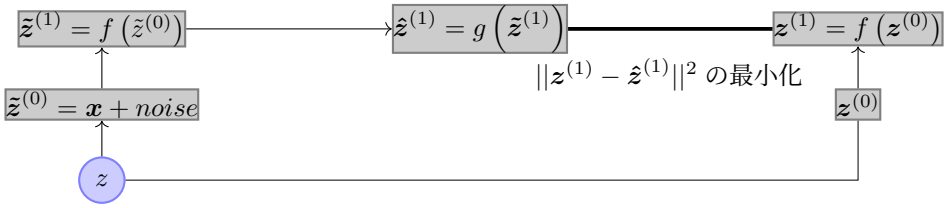
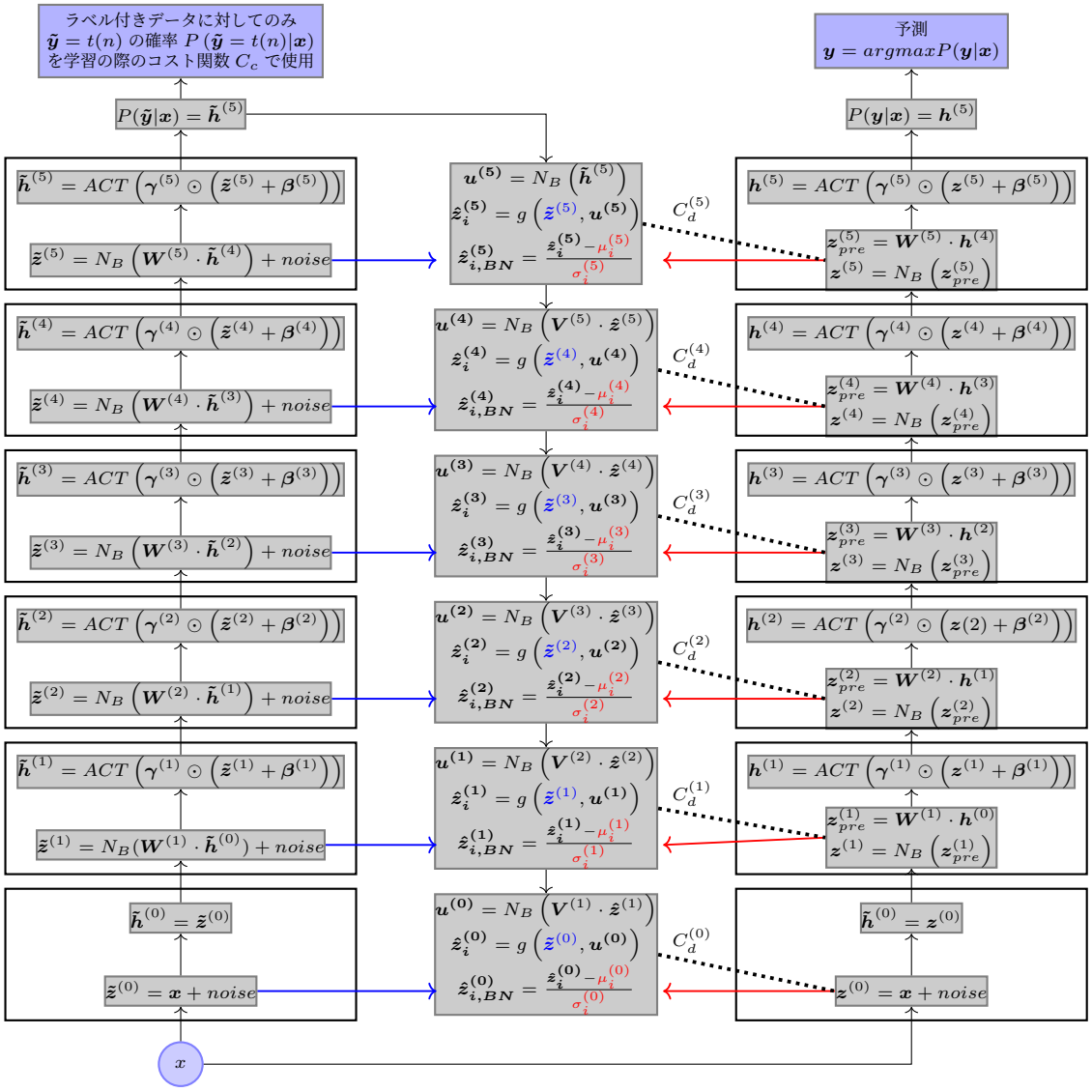

$\tilde{\boldsymbol{z}}^{(0)} = \boldsymbol{x} + noise$
$\hat{\boldsymbol{z}}^{(0)} = g\left(\tilde{\boldsymbol{z}}^{(0)}\right)$
z
$||\boldsymbol{z} - \hat{\boldsymbol{z}}||^2$ の最小化


$\tilde{\boldsymbol{z}}^{(1)} = f\left(\tilde{z}^{(0)}\right)$
$\hat{\boldsymbol{z}}^{(1)} = g\left(\tilde{\boldsymbol{z}}^{(1)}\right)$
$\boldsymbol{z}^{(1)} = f\left(\boldsymbol{z}^{(0)}\right)$
$||\boldsymbol{z}^{(1)} - \hat{\boldsymbol{z}}^{(1)}||^2$ の最小化
$\tilde{\boldsymbol{z}}^{(0)} = \boldsymbol{x} + noise$
$\boldsymbol{z}^{(0)}$
z



$$\boldsymbol{\mu}^{(l)} = batchmean\left(\boldsymbol{z}_{pre}^{(l)}\right)$$
$$\boldsymbol{\sigma}^{(l)} = batchstd\left(\boldsymbol{z}_{pre}^{(l)}\right)$$
$$\hat{z}_i^{(l)} = g_i\left(\tilde{z}_i^{(l)}, u_i^{(l)}\right) = \left(\tilde{z}_i^{(l)} - \mu_i\left(u_i^{(l)}\right)\right) \cdot v_i\left(u_i^{(l)}\right) + \mu_i\left(u_i^{(l)}\right)$$
$$\mu_i\left(u_i^{(l)}\right) = a_{1,i}^{(l)} \cdot sigmoid\left(a_{2,i}^{(l)} \cdot u_i^{(l)} + a_{3,i}^{(l)}\right) + a_{4,i}^{(l)} \cdot u_i^{(l)} + a_{5,i}^{(l)}$$
$$v_i\left(u_i^{(l)}\right) = a_{6,i}^{(l)} \cdot sigmoid\left(a_{7,i}^{(l)} \cdot u_i^{(l)} + a_{8,i}^{(l)}\right) + a_{9,i}^{(l)} \cdot u_i^{(l)} + a_{10,i}^{(l)}$$
$$C_d = \sum_{l=0}^{L} \lambda_l C_d^{(l)} = \sum_{l=0}^{L} \lambda_l ||\boldsymbol{z}^{(l)} - \hat{\boldsymbol{z}}_{BN}^{(l)}||^2$$