Calculus rules $\frac{d}{dx}x^a = ax^{a-1}$ $\frac{d}{dx}kf(x) = kf'(x)$

$\frac{d}{dx}[f(x)+g(x)] = f'(x)+g'(x)$

$\frac{d}{dx}\big(f(g(x))\big) = g'(x)\cdot f'(g(x))$

$\frac{d}{dx}\log x = \frac{1}{x}$ $\frac{d}{dx}e^x = e^x$ $\frac{d}{dx}f(w) = 0$

Log rules $\log ab = \log a + \log b$

$\log a^n = n\log a$ $\log e^a = \log(\exp(a)) = a$

Update rules

Basic logistic update:

$$w_j \leftarrow w_j + \varepsilon x_j^i\left(y^i - g(\boldsymbol{w}^T\boldsymbol{x})\right)$$

SVM optimization with slack:

$\underset{w}{\operatorname{argmin}}\ \boldsymbol{w}^T\boldsymbol{w} + C\sum_i \xi^i$

such that

$\boldsymbol{w}^T\boldsymbol{x}^+ + b \geq +1 - \xi^i$

$\boldsymbol{w}^T\boldsymbol{x}^- + b \leq -1 + \xi^i$

$$\boldsymbol{w} = \sum_i \alpha^i y^i \boldsymbol{x}^i$$

For PCA: $z_q^i = (\boldsymbol{u}^q)^T\boldsymbol{x}^i$ (Linear algebra version: $\boldsymbol{z} = \boldsymbol{U}^T\boldsymbol{x}^i$)

For Neural Network:

$\delta_k^{top,i} = \left(1 - r_k^{top,i}\right)\left(y^i - r_k^{top,i}\right)r_k^{top,i}$

(Linear algebra version: $\boldsymbol{\delta^{top}} = \left(\boldsymbol{r^{top,i}}\right)^{\boldsymbol{T}}\left(\left(1-\boldsymbol{r^{top,i}}\right)\circ\left(\boldsymbol{y^i} - \boldsymbol{r_k^{top,i}}\right)\right)$)

HMM

$$\beta_t(i) = \sum_j A_{j,i}\phi_{o_{t+1},j}\beta_{t+1}(j)$$

$$S_t(i,j) = \frac{\alpha_t(j)A_{i,j}\phi_{o_{t+1},i}\beta_{t+1}(i)}{\sum_{f,g}\alpha_t(g)A_{f,g}\phi_{o_{t+1},f}\beta_{t+1}(f)}$$

Linear algebra for $\boldsymbol{\beta_t}$

$$\boldsymbol{\beta_t} = \boldsymbol{A}^T\left(\boldsymbol{\phi_{o_{t+1}}}\circ\boldsymbol{\beta_{t+1}}\right)$$

From M-step: $A_{i,j} = \frac{\sum_t S_t(i,j)}{\sum_k\sum_t S_t(k,j)}$