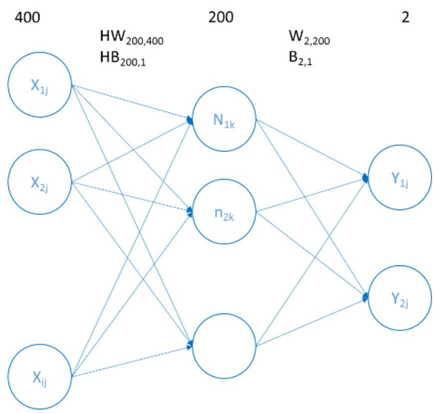


$$Y = WN + B = W(f(HW \cdot X + HB)) + B$$

Nj=HW*Xj+HB

$$\alpha\nabla_{HW}L = \frac{\alpha}{N}\frac{\partial}{\partial HW}L = -\frac{\alpha}{N}\frac{\partial}{\partial HW}\sum_{i} D(S(WX_j + b), L_{ij}) = -\frac{\alpha}{N}\frac{\partial}{\partial HW}\sum_{i}\sum_{j} L_{ij}\ln(S_{ij})$$

$$= -\frac{\alpha}{N}\frac{\partial}{\partial HW}\sum_{i=2}\sum_{j=50} L_{ij}\ln\left(\frac{e^{y_{ij}}}{e^{y_{1j}}+\cdots+e^{y_{ij}}}\right) = -\frac{\alpha}{N}\frac{\partial}{\partial HW}\sum_{i=2}\sum_{j=50} L_{ij}y_{ij} - L_{ij}\ln(e^{y_{1j}}+\cdots+e^{y_{ij}})$$

$$= -\frac{\alpha}{N}\frac{\partial N}{\partial HW}\frac{\partial}{\partial N}\sum_{i=2}\sum_{j=50} L_{ij}(W_iN_j + B_i) - L_{ij}\ln(e^{y_{1j}}+\cdots+e^{y_{ij}})$$

$$= -\frac{\alpha}{N}\frac{\partial N}{\partial HW}\frac{\partial}{\partial N}\sum_{j=50} L_{1j}\left(W_1N_j + B_1\right)+\cdots+L_{ij}\left(W_iN_j + B_i\right) - (L_{1j}+\cdots+L_{ij})\ln(e^{y_{1j}}+\cdots+e^{y_{ij}})$$

$$\begin{aligned}= -\frac{\alpha}{N}\frac{\partial N}{\partial HW}\frac{\partial}{\partial N}&[L_{11}(W_1N_1 + B_1)+\cdots+L_{i1}(W_iN_1 + B_i)+\cdots+L_{1j}\left(W_1N_j + B_1\right)+\cdots+L_{ij}\left(W_iN_j + B_i\right) \\ &- (L_{11}+\cdots+L_{i1})\ln(e^{y_{1j}}+\cdots+e^{y_{ij}}) - \cdots - \left(L_{1j}+\cdots+L_{ij}\right)\ln(e^{y_{1j}}+\cdots+e^{y_{ij}})]\end{aligned}$$

$$\begin{aligned}= -\frac{\alpha}{N}\frac{\partial N}{\partial HW}\frac{\partial}{\partial N}&[L_{11}(W_1N_1)+\cdots+L_{i1}(W_iN_1)+\cdots+L_{1j}\left(W_1N_j\right)+\cdots+L_{ij}\left(W_iN_j\right) \\ &- (L_{11}+\cdots+L_{i1})\ln(e^{y_{11}}+\cdots+e^{y_{i1}}) - \cdots - \left(L_{1j}+\cdots+L_{ij}\right)\ln(e^{y_{1j}}+\cdots+e^{y_{ij}})]\end{aligned}$$

$$= -\frac{\alpha}{N}\frac{\partial N}{\partial HW}\left\{\begin{bmatrix} L_{11}W_1 + \ldots + L_{i1}W_i \\ \vdots \\ L_{1j}W_1 + \cdots + L_{ij}W_i \end{bmatrix} - \begin{bmatrix} (L_{11}+\cdots+L_{i1})\dfrac{\frac{\partial}{\partial N1}(e^{y_{11}}+\cdots+e^{y_{i1}})}{e^{y_{11}}+\cdots+e^{y_{i1}}} \\ \vdots \\ \left(L_{1j}+\cdots+L_{ij}\right)\dfrac{\frac{\partial}{\partial N2}(e^{y_{11}}+\cdots+e^{y_{i1}})}{e^{y_{1j}}+\cdots+e^{y_{ij}}} \end{bmatrix}\right\}$$

$$= -\frac{\alpha}{N}\frac{\partial N}{\partial HW}\left\{L^T \cdot W - \begin{bmatrix} \dfrac{1}{e^{y_{11}}+\cdots+e^{y_{i1}}}[W_1e^{y11}+\cdots+W_ie^{yi1}] \\ \vdots \\ \dfrac{1}{e^{y_{1j}}+\cdots+e^{y_{ij}}}[W_1e^{y1j}+\cdots+W_ie^{yij}] \end{bmatrix}\right\}$$

$$= -\frac{\alpha}{N}\frac{\partial(R(HWX + HB))}{\partial HW}\left\{L^T \cdot W - \begin{bmatrix} W_1S(y_{11})+\cdots+W_iS(y_{i1}) \\ \vdots \\ W_1S(y_{1j})+\cdots+W_iS(y_{ij}) \end{bmatrix}\right\}$$

$$= -\frac{\alpha}{N}\left(\frac{\partial}{\partial HW}\begin{bmatrix} R(HW_1X_1 + HB_1) & \cdots & R\left(HW_1X_j + HB_1\right) \\ \vdots & \ddots & \vdots \\ R(HW_kX_1 + HB_k) & \cdots & R\left(HW_kX_j + HB_k\right) \end{bmatrix}\right)[L^T \cdot W - S^T \cdot W]$$

$$= -\frac{\alpha}{N}\begin{bmatrix} R'(HW_1X_1 + HB_1)X_1 & \cdots & R'\left(HW_1X_j + HB_1\right)X_j \\ \vdots & \ddots & \vdots \\ R'(HW_kX_1 + HB_k)X_1 & \cdots & R\left(HW_kX_j + HB_k\right)X_j \end{bmatrix}[L^T \cdot W - S^T \cdot W] = -\frac{\alpha}{N}[R'(HWX + HB) \cdot X][(L^T - S^T) \cdot W]$$

200*50   400*50   2*50   2*200

i=50, j=2,k=200,m=400

$$\alpha\nabla_{HW}L = \alpha\frac{\partial}{\partial hw_{mk}}\sum_i D(S(WX_i + b), L_i) = -\frac{\alpha}{N}\frac{\partial}{\partial hw_{mk}}\sum_i [\sum_j L_{ij}\ln(\frac{e^{y_{ij}}}{\sum_j e^{y_j}})] = -\frac{\alpha}{N}\frac{\partial}{\partial hw_{mk}}\sum_i [\sum_j L_{ij}\, y_{ij} - L_{ij}\ln(\sum_j e^{y_j})]$$

$$= -\frac{\alpha}{N}\frac{\partial}{\partial hw_{mk}}\sum_i [\sum_j L_{ij}\ (\sum_k (w_{jk}n_{ki}) + b_j) - L_{ij}\ln(e^{y_{1i}} + \cdots + e^{y_{ji}})]$$

$$= -\frac{\alpha}{N}\frac{\partial}{\partial hw_{mk}}\sum_i [\sum_j L_{ij}\ (\sum_k \left(w_{jk} \cdot R(\sum_m (hw_{km} \cdot x_{mi}) + hb_k)\right) + b_j) - L_{ij}\ln(e^{y_{1i}} + \cdots + e^{y_{ji}})]$$

$$= -\frac{\alpha}{N}\sum_i [\sum_j L_{ij}\, w_{jk} \cdot R'(\sum_m (hw_{km} \cdot x_{mi}) + hb_k)\, x_{mi} - L_{ij}\frac{(w_{1k} \cdot R'(\sum_m (hw_{km} \cdot x_{mi}) + hb_k)\, x_{mi}) \cdot e^{y_{1i}} + \cdots + \cdots}{e^{y_{1i}} + \cdots + e^{y_{ji}}}]$$

$$= -\frac{\alpha}{N}\sum_i [\sum_j L_{ij}(\, w_{jk} \cdot R'\left(\sum_m (hw_{km} \cdot x_{mi}) + hb_k\right) x_{mi}) - L_{ij}(\left(w_{1k} \cdot R'\left(\sum_m (hw_{km} \cdot x_{mi}) + hb_k\right) x_{mi}\right) S(y_{1i}) + \cdots$$

$$+ \left(w_{jk} \cdot R'\left(\sum_m (hw_{km} \cdot x_{mi}) + hb_k\right) x_{mi}\right) S(y_{ji}))]$$

$$= -\frac{\alpha}{N}\sum_i [\sum_j (L_{ij} - S_{ij})(\, w_{jk} \cdot R'\left(\sum_m (hw_{km} \cdot x_{mi}) + hb_k\right) \cdot x_{mi})]$$

[(L11 –S11) w1k + (L12 – S12) w2k] xm1 Rk1 + [(L21 –S21)w1k +(L22- S22) w2k] xm2 Rk2 +…+ [(L50,1 –S50,1)w1k +(L50,2-S50,2) w2k] xm50 Rk50

$$[w1k \quad w2k][\begin{matrix}(\mathrm{L11} - \mathrm{S11})\mathrm{xm1\ Rk1} + (\mathrm{L21} - \mathrm{S21})\mathrm{xm2\ Rk2} + \cdots \\ (\mathrm{L12} - \mathrm{S12})\mathrm{xm1\ Rk1} + (\mathrm{L22} - \mathrm{S22})\mathrm{xm2Rk2} + \cdots\end{matrix}]$$

$$= [w1k \quad w2k]\begin{bmatrix}(\mathrm{L11} - \mathrm{S11})\ (\mathrm{L21} - \mathrm{S21})\ \ldots \\ (\mathrm{L12} - \mathrm{S12})\ (\mathrm{L22} - \mathrm{S22})\ \ldots\end{bmatrix}[\mathrm{Rk1xm1\ Rk2xm2} \ldots \mathrm{Rk50xm50}]\mathrm{T}$$

$$= [w1k \quad w2k]\ (L - S)\ [\mathrm{Rk1xm1\ Rk2xm2} \ldots \mathrm{Rk50xm50}]\mathrm{T}$$

$$= [w1k \quad w2k](\ (L - S)([\mathrm{xm1\ xm2} \ldots \mathrm{xm50}\ ]T * [\mathrm{R'k1\ R'k2} \ldots \mathrm{R'k50}]\mathrm{T}))$$

$$\alpha\nabla_{HB}L = \alpha\frac{\partial}{\partial hb_k}\sum_i D(S(WX_i + b), L_i) = -\frac{\alpha}{N}\frac{\partial}{\partial hb_k}\sum_i [\sum_j L_{ij}\ln(\frac{e^{y_{ij}}}{\sum_j e^{y_j}})] = -\frac{\alpha}{N}\frac{\partial}{\partial hb_k}\sum_i [\sum_j L_{ij}\, y_{ij} - L_{ij}\ln(\sum_j e^{y_j})]$$

$$= -\frac{\alpha}{N}\frac{\partial}{\partial hb_k}\sum_i [\sum_j L_{ij}\ (\sum_k (w_{jk}n_{ki}) + b_j) - L_{ij}\ln(e^{\sum_k (w_{1k}n_{ki}) + b_1} + \cdots + e^{\sum_k (w_{jk}n_{ki}) + b_j})]$$

$$= -\frac{\alpha}{N}\frac{\partial}{\partial hb_k}\sum_i [\sum_j L_{ij}\ (\sum_k \left(w_{jk} \cdot R(\sum_m (hw_{km} \cdot x_{mi}) + hb_k)\right) + b_j) - L_{ij}\ln(e^{y_{1i}} + \cdots + e^{y_{ji}})]$$

$$= -\frac{\alpha}{N}\sum_i [\sum_j L_{ij}\ (w_{jk} \cdot R'\left(\sum_m (hw_{km} \cdot x_{mi}) + hb_k\right)) - L_{ij}\frac{(w_{jk} \cdot R'(\sum_m (hw_{km} \cdot x_{mi}) + hb_k)) \cdot e^{y_{1i}} + \cdots + \cdots}{e^{y_{1i}} + \cdots + e^{y_{ji}}}]$$

$$= -\frac{\alpha}{N}\sum_i [\sum_j L_{ij}\ (w_{jk} \cdot R'\left(\sum_m (hw_{km} \cdot x_{mi}) + hb_k\right)) - L_{ij}\frac{(w_{jk} \cdot R'(\sum_m (hw_{km} \cdot x_{mi}) + hb_k)) \cdot e^{y_{1i}} + \cdots + \cdots}{e^{y_{1i}} + \cdots + e^{y_{ji}}}]$$

$$= -\frac{\alpha}{N}\sum_i [\sum_j L_{ij}\ (w_{jk} \cdot R'\left(\sum_m (hw_{km} \cdot x_{mi}) + hb_k\right)) - L_{ij}\frac{(w_{1k} \cdot R'(\sum_m (hw_{km} \cdot x_{mi}) + hb_k)) \cdot e^{y_{1i}} + \cdots + \cdots}{e^{y_{1i}} + \cdots + e^{y_{ji}}}]$$

$$= -\frac{\alpha}{N}\sum_i[\sum_j L_{ij}\,(w_{jk}\cdot R'\left(\sum_m(hw_{km}\cdot x_{mi})+hb_k\right)) - L_{ij}(\left(w_{1k}\cdot R'\left(\sum_m(hw_{km}\cdot x_{mi})+hb_k\right)\right)S(y_{1i})+\cdots$$

$$+\left(w_{jk}\cdot R'\left(\sum_m(hw_{km}\cdot x_{mi})+hb_k\right)\right)S(y_{ji}))]$$

$$= -\frac{\alpha}{N}\sum_i[\sum_j(L_{ij}-S_{ij})\,(w_{jk}\cdot R'\left(\sum_m(hw_{km}\cdot x_{mi})+hb_k\right))]$$

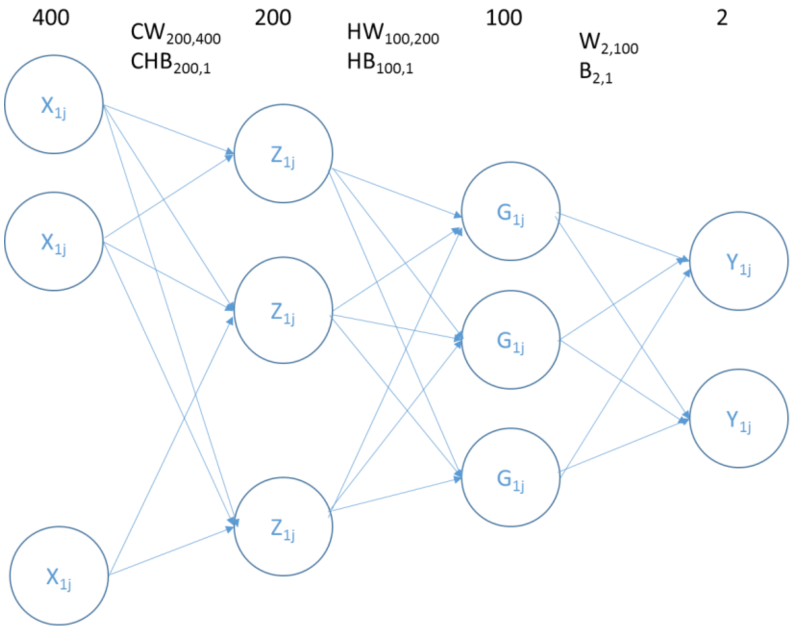
= w1k((L11-S11) Rk1+(L21- S21) Rk2+…(L50,1- S50,1) Rk50)+ w2k((L12-S12) Rk1+(L22-S22) Rk2+…(L50,2-S50,2) Rk50)

=w1k[(L11-S11) (L21- S21)…(L50,1- S50,1)][Rk1 Rk2…Rk50]^T+ w2k[(L12-S12) (L22-S22) …(L50,2-S50,2)] [Rk1 Rk2…Rk50]^T

$$[wk1\quad wk2]\begin{bmatrix}[(\mathrm{L}11-\mathrm{S}11)(\mathrm{L}21-\ \mathrm{S}21)\ldots(\mathrm{L}50{,}1-\ \mathrm{S}50{,}1)][\mathrm{Rk}1\ \mathrm{Rk}2\ldots\mathrm{Rk}50]\mathrm{T}\\ [(\ \mathrm{L}12-\mathrm{S}12)(\ \mathrm{L}22-\mathrm{S}22)\ldots(\ \mathrm{L}50{,}2-\mathrm{S}50{,}2)][\mathrm{Rk}1\ \mathrm{Rk}2\ldots\mathrm{Rk}50]\mathrm{T}\end{bmatrix}$$

$$[wk1\quad wk2](L-S)[\mathrm{Rk}1\ \mathrm{Rk}2\ldots\mathrm{Rk}50]\mathrm{T}$$

2*200 200*50 50*2



i=50, j=2,k=200,m=400, q=100

$$\alpha\nabla_{CW}L = -\frac{\alpha}{N}\frac{\partial}{\partial cw_{mk}}\sum_i[\sum_j L_{ij}\ln(\frac{e^{y_{ij}}}{\sum_j e^{y_j}})] = -\frac{\alpha}{N}\frac{\partial}{\partial cw_{mk}}\sum_i[\sum_j L_{ij}\,y_{ij} - L_{ij}\ln(\sum_j e^{y_j})]$$

$$= -\frac{\alpha}{N}\frac{\partial}{\partial cw_{mk}}\sum_i[\sum_j L_{ij}\,(\sum_q(w_{jq}g_{qi})+b_j) - L_{ij}\ln(e^{y_{1i}}+\cdots+e^{y_{ji}})]$$

$$= -\frac{\alpha}{N}\frac{\partial}{\partial cw_{mk}}\sum_i[\sum_j L_{ij}\,(\sum_q\left(w_{jq}\cdot R(\sum_k(hw_{qk}z_{ki})+hb_q)\right)+b_j) - L_{ij}\ln(e^{y_{1i}}+\cdots+e^{y_{ji}})]$$

$$= -\frac{\alpha}{N}\frac{\partial}{\partial cw_{mk}}\sum_i[\sum_j L_{ij}\,(\sum_q\left(w_{jq}\cdot R(\sum_k\left(hw_{qk}\cdot R(\sum_m(cw_{km}x_{mi})+cb_k)\right)+hb_q)\right)+b_j) - L_{ij}\ln(e^{y_{1i}}+\cdots+e^{y_{ji}})]$$

$$= -\frac{\alpha}{N}\sum_i[\sum_j L_{ij}\,(\sum_q\left(w_{jq}\cdot R'(\sum_k\left(hw_{qk}\cdot R(\sum_m(cw_{km}x_{mi})+cb_k)\right)+hb_q)\cdot hw_{qk}\cdot R'(\sum_m(cw_{km}x_{mi})+cb_k)\cdot x_{mi}\right))$$

$$-L_{ij}\frac{(\quad)e^{y_{1i}}+\cdots+(\quad)e^{y_{ji}}}{(e^{y_{1i}}+\cdots+e^{y_{ji}})}]$$

$$= -\frac{\alpha}{N}\sum_i[\sum_j L_{ij}\ (\sum_q\left(w_{jq}\cdot R'(\sum_k\left(hw_{qk}\cdot R(\sum_m(cw_{km}x_{mi})+cb_k)\right)+hb_q)\cdot hw_{qk}\cdot R'(\sum_m(cw_{km}x_{mi})+cb_k)\cdot x_{mi}\right)) - L_{ij}((\quad)S(y_{1i}) + \cdots (\quad)S(y_{ji})]$$

$$= -\frac{\alpha}{N}\sum_i[\sum_j L_{ij}\ (\sum_q\left(w_{jq}\cdot R'(\sum_k(hw_{qk}\cdot z_{ki})+hb_q)\cdot hw_{qk}\cdot R'(\sum_m(cw_{km}x_{mi})+cb_k)\cdot x_{mi}\right)) - L_{ij}((\quad)S(y_{1i})+\cdots(\quad)S(y_{ji})]$$

$$= -\frac{\alpha}{N}\sum_i[\sum_j(L_{ij}-S_{ij})\ (\sum_q\left(w_{jq}\cdot R'(\sum_k(hw_{qk}\cdot z_{ki})+hb_q)\cdot hw_{qk}\cdot R'(\sum_m(cw_{km}x_{mi})+cb_k)\cdot x_{mi}\right))]$$

$$\alpha\nabla_{CB}L = \alpha\frac{\partial}{\partial cb_k}\sum_i D(S(WX_i+b),L_i) = -\frac{\alpha}{N}\frac{\partial}{\partial cb_k}\sum_i[\sum_j L_{ij}\ \ln(\frac{e^{y_{ij}}}{\sum_j e^{y_j}})] = -\frac{\alpha}{N}\frac{\partial}{\partial cb_k}\sum_i[\sum_j L_{ij}\ y_{ij} - L_{ij}\ln(\sum_j e^{y_j})]$$

$$= -\frac{\alpha}{N}\frac{\partial}{\partial cb_k}\sum_i[\sum_j L_{ij}\ (\sum_q(w_{jq}g_{qi})+b_j) - L_{ij}\ln(e^{y_{1i}}+\cdots+e^{y_{ji}})]$$

$$= -\frac{\alpha}{N}\frac{\partial}{\partial cb_k}\sum_i[\sum_j L_{ij}\ (\sum_q\left(w_{jq}\cdot R(\sum_k(hw_{qk}z_{ki})+hb_q)\right)+b_j) - L_{ij}\ln(e^{y_{1i}}+\cdots+e^{y_{ji}})]$$

$$= -\frac{\alpha}{N}\frac{\partial}{\partial cb_k}\sum_i[\sum_j L_{ij}\ (\sum_q\left(w_{jq}\cdot R(\sum_k\left(hw_{qk}\cdot R(\sum_m(cw_{km}x_{mi})+cb_k)\right)+hb_q)\right)+b_j) - L_{ij}\ln(e^{y_{1i}}+\cdots+e^{y_{ji}})]$$

$$= -\frac{\alpha}{N}\sum_i[\sum_j L_{ij}\ (\sum_q\left(w_{jq}\cdot R'(\sum_k\left(hw_{qk}\cdot R\left(\sum_m(cw_{km}x_{mi})+cb_k\right)\right)+hb_q)\cdot hw_{qk}\cdot R'(\sum_m(cw_{km}x_{mi})+cb_k)\right)) - L_{ij}\frac{(\quad)e^{y_{1i}}+\cdots+(\quad)e^{y_{ji}}}{(e^{y_{1i}}+\cdots+e^{y_{ji}})}]$$

$$= -\frac{\alpha}{N}\sum_i[\sum_j L_{ij}\ (\sum_q\left(w_{jq}\cdot R'\left(\sum_k(hw_{qk}\cdot z_{ki})+hb_q\right)\cdot hw_{qk}\cdot R'\left(\sum_m(cw_{km}x_{mi})+cb_k\right)\right)) - L_{ij}(\sum_q\left(w_{1q}\cdot R'\left(\sum_k(hw_{qk}\cdot z_{ki})+hb_q\right)\cdot hw_{qk}\cdot R'\left(\sum_m(cw_{km}x_{mi})+cb_k\right)\right)S(y_{1i})+\cdots+(\ldots)S(y_{ji}))]$$

$$= -\frac{\alpha}{N}\sum_i[\sum_j(L_{ij}-S_{ij})(\sum_q\left(w_{jq}\cdot R'\left(\sum_k(hw_{qk}\cdot z_{ki})+hb_q\right)\cdot hw_{qk}\cdot R'\left(\sum_m(cw_{km}x_{mi})+cb_k\right)\right))]$$

(L11-S11) [w11 Rh’11 hwk,1 +…+ w1,100 Rh’1,100 hwk,100] RC’k1

+ (L12-S12)[w21 Rh’11 hwk1 +…+ w2,100 Rh’1,100 hwk,100] RC’k1

+…

+(L50,1-S50,1) [w11 Rh’1,50 hwk,1+…+ w1,100 Rh’50,100 hwk,100] RC’k,50

+(L50,2-S50,2) [w21 Rh’1,50 hwk,1+…+ w2,100 Rh’50,100 hwk,100] RC’k,50

=

$$\left([\mathrm{hwk},1 \quad \ldots \quad \mathrm{hwk},100]\begin{bmatrix}\mathrm{w1,1} & \mathrm{w1,100}\\ \vdots & \vdots\\ \mathrm{w2,1} & \mathrm{w2,100}\end{bmatrix}\cdot\begin{bmatrix}\mathrm{L11}-\mathrm{S11} & \ldots & \mathrm{L50,1}-\mathrm{S50,1}\\ \mathrm{L12}-\mathrm{S12} & \ldots & \mathrm{L50,2}-\mathrm{S50,2}\end{bmatrix} * \left(\begin{bmatrix}\mathrm{Rh'1,1} & \cdots & \mathrm{Rh'50,1}\\ \vdots & \ddots & \vdots\\ \mathrm{Rh'1,100} & \cdots & \mathrm{Rh'50,100}\end{bmatrix}\right)\right)\begin{bmatrix}\mathrm{RC'k1}\\ \vdots\\ \mathrm{RC'k,50}\end{bmatrix}$$