**Professor Daniel Drucker's Calculation for the proof of Eckart-Young (page 75)**

**Gilbert Strang, Linear Algebra and Learning from Data**

The matrices $A, C$, and $R$ have dimensions $m\times n$, $m\times k$ and $k\times n$. Set $E = ||A-CR||_F^2 = \sum_{r=1}^{m}\sum_{s=1}^{n}\{(A-CR)_{rs}\}^2$.

$$\begin{aligned}
\frac{\partial E}{\partial C_{ij}} &= \frac{\partial}{\partial C_{ij}}\sum_{r=1}^{m}\sum_{s=1}^{n}\left(A_{rs}-\sum_{t=1}^{k}C_{rt}\,R_{ts}\right)^2 \\
&= 2\sum_{r=1}^{m}\sum_{s=1}^{n}\left(A_{rs}-\sum_{t=1}^{k}C_{rt}\,R_{ts}\right)\frac{\partial}{\partial C_{ij}}\left(A_{rs}-\sum_{u=1}^{k}C_{ru}\,R_{us}\right) \\
&= 2\sum_{r=1}^{m}\sum_{s=1}^{n}\left(A_{rs}-\sum_{t=1}^{k}C_{rt}\,R_{ts}\right)\left(-\sum_{u=1}^{k}\delta_{ri}\,\delta_{uj}\,R_{us}\right) \\
&= 2\sum_{r=1}^{m}\sum_{s=1}^{n}\left(\sum_{t=1}^{k}C_{rt}\,R_{ts}-A_{rs}\right)\delta_{ri}\,R_{js} = 2\sum_{s=1}^{n}\left(\sum_{t=1}^{k}C_{it}\,R_{ts}-A_{is}\right)R_{js} \\
&= 2\left\{\sum_{s=1}^{n}(CR)_{is}\,R_{js}-\sum_{s=1}^{n}A_{is}\,R_{js}\right\} = 2\left\{\sum_{s=1}^{n}(CR)_{is}\,(R^{\mathrm{T}})_{sj}-\sum_{s=1}^{n}A_{is}\,(R^{\mathrm{T}})_{sj}\right\} \\
&= 2\left\{(CRR^{\mathrm{T}})_{ij}-(AR^{\mathrm{T}})_{ij}\right\} = \mathbf{2}\left\{\boldsymbol{(CR-A)R^{\mathrm{T}}}\right\}_{\boldsymbol{ij}}
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial E}{\partial R_{ij}} &= \frac{\partial}{\partial R_{ij}}\sum_{r=1}^{m}\sum_{s=1}^{n}\left(A_{rs}-\sum_{t=1}^{k}C_{rt}\,R_{ts}\right)^2 \\
&= 2\sum_{r=1}^{m}\sum_{s=1}^{n}\left(A_{rs}-\sum_{t=1}^{k}C_{rt}\,R_{ts}\right)\frac{\partial}{\partial R_{ij}}\left(A_{rs}-\sum_{u=1}^{k}C_{ru}\,R_{us}\right) \\
&= 2\sum_{r=1}^{m}\sum_{s=1}^{n}\left(A_{rs}-\sum_{t=1}^{k}C_{rt}\,R_{ts}\right)\left(-\sum_{u=1}^{k}\delta_{ui}\,\delta_{sj}\,C_{ru}\right) \\
&= 2\sum_{r=1}^{m}\sum_{s=1}^{n}\left(\sum_{t=1}^{k}C_{rt}\,R_{ts}-A_{rs}\right)\delta_{sj}\,C_{ri} = 2\sum_{r=1}^{m}\left(\sum_{t=1}^{k}C_{rt}\,R_{tj}-A_{rj}\right)C_{ri} \\
&= 2\left\{\sum_{r=1}^{m}(CR)_{rj}\,C_{ri}-\sum_{r=1}^{m}A_{rj}\,C_{ri}\right\} = 2\left\{\sum_{r=1}^{m}(C^{\mathrm{T}})_{ir}\,(CR)_{rj}-\sum_{r=1}^{m}(C^{\mathrm{T}})_{ir}\,A_{rj}\right\} \\
&= 2\left\{C^{\mathrm{T}}(CR)_{ij}-(C^{\mathrm{T}}A)_{ij}\right\} = \mathbf{2}\left\{\boldsymbol{C^{\mathrm{T}}(CR-A)}\right\}_{\boldsymbol{ij}}, \quad \text{so}
\end{aligned}$$

$$\boxed{\frac{\partial E}{\partial C} = 2(CR-A)R^{\mathrm{T}} \quad \text{and} \quad \frac{\partial E}{\partial R} = 2C^{\mathrm{T}}(CR-A).}$$

Page 75 had the transpose of the correct $\partial E/\partial R$. Fortunately still okay since we set $\partial E/\partial R = 0$ and can safely transpose 0.