$$\mathcal{L}(u,v) = \sum_{i,j} \left(R_{i,j} - u_{i\_} \cdot v_{\_j}\right)^2$$

we want $\dfrac{\nabla \mathcal{L}(u,v)}{u_{i\_}}$ the gradient w.r.t. $u_{i\_}$ (or $v_{\_j}$) of $\mathcal{L}$.

The gradient is the vector of partial derivatives of $u_{i\_}$ ($v_{\_j}$). This is a length k vector

$$\frac{\nabla \mathcal{L}(u,v)}{u_{i\_}} = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial u_{i1}} \\ \frac{\partial \mathcal{L}}{\partial u_{i2}} \\ \vdots \\ \frac{\partial \mathcal{L}}{\partial u_{ik}} \end{bmatrix}$$

For SGD, we choose a single i,j. So we drop $\sum_{i,j}$. And to compute ∇ we worry about one of these partials, since the rest are identical:

$$\frac{\partial \mathcal{L}(u,v)}{\partial u_{ik^*}}$$

$$= \frac{\partial}{\partial u_{ik^*}} \left(R_{i,j} - u_{i\_} v_{\_j} \right)^2$$

$$= \frac{\partial}{\partial u_{ik^*}} \left(R_{i,j} - \sum_{\ell=1}^{k} u_{i\ell} \cdot v_{\ell j} \right)^2$$

chain rule $\quad = 2\left(R_{i,j} - u_{i\_} v_{\_j}\right) \dfrac{\partial}{\partial u_{ik^*}} \left(R_{i,j} - \sum_{\ell=1}^{k} u_{i\ell} \cdot v_{\ell j}\right)$

$$= -2 v_{k^* j} \left(R_{i,j} - u_{i\_} v_{\_j}\right)$$