

UVD via SGD:

$$R_{m \times n} \approx U_{m \times k} V_{k \times n}$$
$$\begin{bmatrix} R \end{bmatrix} \approx \begin{bmatrix} U \end{bmatrix} \begin{bmatrix} V \end{bmatrix}$$

Diagram illustrating the matrix approximation $R \approx UV$. Arrows point from the U and V matrices to the equation $r_{ij} \approx u_{i-} \cdot v_{-j}$.

Recall:

$$\operatorname{argmin}_{U, V} \frac{1}{2} \sum_{(i, j)} (r_{ij} - u_{i-} \cdot v_{-j})^2$$
$$= \operatorname{argmin}_{U, V} \frac{1}{2} \sum E_{ij}$$

Compute U and V:

1. Init U and V randomly.
2. Choose random indices (i, j) .
3. Estimate $\tilde{r}_{ij} = u_{i-} \cdot v_{-j}$
4. Update u_{i-} and v_{-j} as:
$$\Delta u_{i-} = -\gamma \frac{\partial E_{ij}}{\partial u_{i-}}$$
$$\Delta v_{-j} = -\gamma \frac{\partial E_{ij}}{\partial v_{-j}}$$
5. Repeat until adequate convergence.

What is $\frac{\partial E_{ij}}{\partial u_{i-}}$ and $\frac{\partial E_{ij}}{\partial v_{-j}}$? Let's derive $\frac{\partial E_{ij}}{\partial u_{i-}}$:

First let's look at a ~~single value~~ ^{particular element} within the u_{i-} vector,

$$\frac{\partial E_{ij}}{\partial u_{ik}} = -(r_{ij} - u_{i-} \cdot v_{-j})^2 v_{kj}$$

u_{ik}
↳ we'll calculate the partial derivative at u_{ik} .

Now we can see:

$$\frac{\partial E_{ij}}{\partial u_{i-}} = -(r_{ij} - u_{i-} \cdot v_{-j})^2 v_{-j}$$