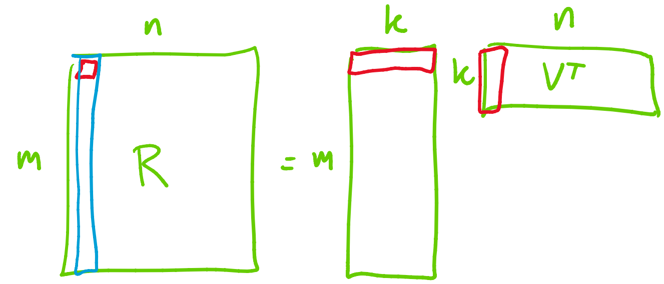
Section B

Wednesday, November 30, 2022 12:49 PM

Low-Rank Factorization

Case #1 Given: $m \times n$ matrix R of rank $k \ll \min(m,n)$

Claim: $\exists$ $m \times k$ matrix U
$n \times k$ matrix V
s.t. $R = UV^T$



Note: ① columns of U are the k basis vectors of the cols of R
② rows of V^T are the " " " " " rows of R

Case #2 Given: $m \times n$ matrix R of rank $\ell > k$
where k = # of cols of U and V

Claim: can approximate R with rank-k matrices U and V

$$R \approx UV^T$$

Approx Error

Def: residual matrix $E = R - UV^T$

Frobenius Norm

MSE: $(\|E\|_2)^2 = (\|R - UV^T\|_2)^2$ where $(\|E\|_2)^2 = \sum_i \sum_j (E_{ij})^2$

Unconstrained Matrix Factorization

Opt. Problem #1: (fully observed R)

$$\hat{U}, \hat{V} = \underset{U,V}{\operatorname{argmin}}\, J(U,V) \quad \text{where} \quad J(U,V) = \frac{1}{2}\|R - UV^T\|_2^2$$

Opt Problem #2: (partially observed R)

Let $r_{ij} \triangleq R_{ij}$ ⟵ rating of item i by user j

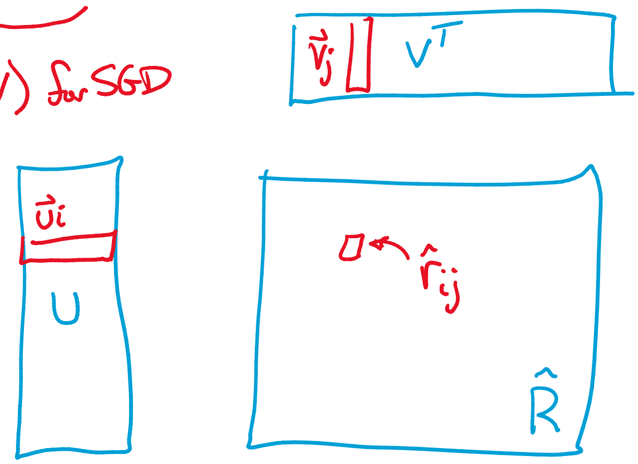
$\vec{u}_i \triangleq U_{i,\cdot}$ ⟵ user factor (learned feature vector)

$\vec{u}_i \triangleq U_{i,\cdot}$ ⟵ user factor (learned feature vector)

$\vec{v}_j \triangleq V_{\cdot,j}$ ⟵ item factor (learned feature vector)

Let $Z = \{(i,j) : r_{ij} \text{ is observed}\}$

$$\star \quad J(U,V) = \frac{1}{2} \sum_{(i,j)\in Z} \underbrace{\left(r_{ij} - \vec{u}_i^T \vec{v}_j\right)^2}_{J_{ij}(U,V) \text{ for SGD}}$$



Model Predictions

$$\hat{r}_{ij} = \vec{u}_i^T \vec{v}_j$$

Gradient Descent:

while not converged:

$g_U \leftarrow \nabla_U J(U,V)$

$g_V \leftarrow \nabla_V J(U,V)$

$U \leftarrow U - \gamma g_U$

$V \leftarrow V - \gamma g_V$

SGD:

while not converged:

1. sample (i,j) from Z
2. step opposite the gradient of $\underbrace{J_{ij}(U,V)}$

ALS

<see slides>