Section B

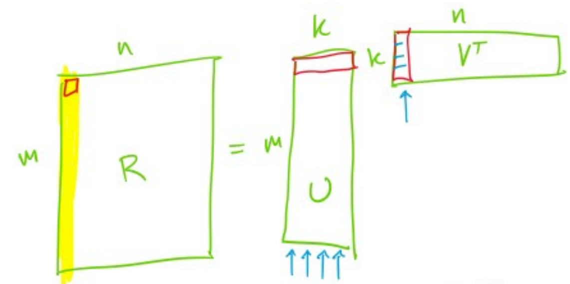
Wednesday, April 19, 2023   11:45 AM

Low-Rank Factorization

Case #1

Given: $m \times n$ matrix R of $\underbrace{\text{rank } k}_{\text{has } k \text{ linearly independent columns}} \ll \min(m,n)$

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 " " " " " rows of R

every row of R is a linear comb. of these basis vectors

Case #2

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

Claim: can approximate R with rank-k matrices U and V
$R \approx UV^T$

Approx Error

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

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

Unconstrained Matrix Factorization

Opt. Problem #1: (fully observed R)

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

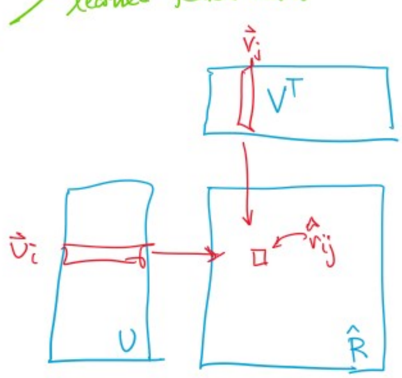
Opt. Problem #2: (partially observed R)

Let $r_{ij} \triangleq R_{ij}$ ← rating of item i by user j
$\vec{u}_i \triangleq U_{i,\cdot}$ ← user factor
$\vec{v}_j \triangleq V_{\cdot,j}$ ← item factor

user factor and item factor: learned feature vectors

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

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



Model Predictions

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

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

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:

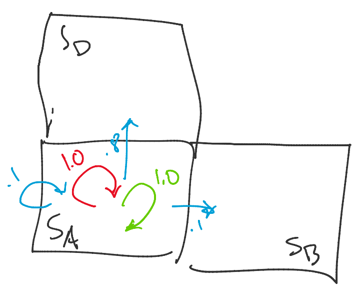
while not converged:

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

Block Coord. Descent:

ALS - see slides

$\boxed{\text{OH}}$



left
right
→ up