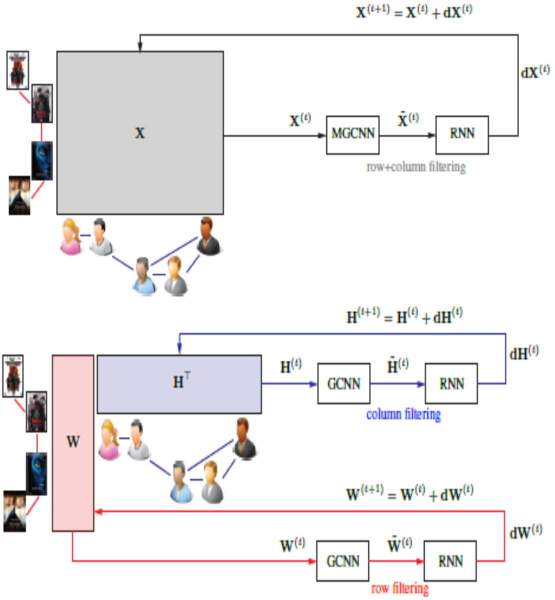
Geometric Matrix Completion with Recurrent Multi-Graph Neural Networks

Jongjin Lee

Presented by Jongjin Lee

2017.11.22

- Recommender systems, estimating rating matrix X
- Geometric structure on the rows and columns(row graph, column graph)
- An extension of CNNs to graphs
- Learning parameters by using LSTM
- Estimating rating Matirx / latent parameters (RMGCNN, sRMGCNN)



$X^{(t+1)} = X^{(t)} + dX^{(t)}$
$dX^{(t)}$
X
$X^{(t)}$
MGCNN
$\tilde{X}^{(t)}$
RNN
row+column filtering
$H^{(t+1)} = H^{(t)} + dH^{(t)}$
$dH^{(t)}$
H^T
$H^{(t)}$
GCNN
$\tilde{H}^{(t)}$
RNN
column filtering
W
$W^{(t+1)} = W^{(t)} + dW^{(t)}$
$dW^{(t)}$
$W^{(t)}$
GCNN
$\tilde{W}^{(t)}$
RNN
row filtering

- Matrix completion problem

$$min \ \|X\|_\star + \frac{\mu}{2}\|\Omega \circ (X - Y)\|_F^2$$

- Geometric structure on the rows and columns
  - Column, row graph $\mathcal{G}_r = (1, ..., m, \mathcal{E}_r, W_r)$, $\mathcal{G}_c = (1, ..., n, \mathcal{E}_c, W_c)$
  - Graph Laplacian $\Delta = \mathrm{I} - D^{-1/2}WD^{-1/2}$, $D = diag(\sum_{i \neq j} w_i j)$
  - Dirichlet norm $\|X\|_{\mathcal{G}_c}^2 = trace(X^T \Delta_c X)$, $\|X\|_{\mathcal{G}_r}^2 = trace(X^T \Delta_r X)$
- Geometric matrix completion(RMGCNN)

$$min \ \|X\|_{\mathcal{G}_r}^2 + \|X\|_{\mathcal{G}_c}^2 + \frac{\mu}{2}\|\Omega \circ (X - Y)\|_F^2$$

- Factorized models(sRMGCNN)

$$min \ \|W\|_{\mathcal{G}_r}^2 + \|H\|_{\mathcal{G}_c}^2 + \frac{\mu}{2}\|\Omega \circ (WH^T - Y)\|_F^2$$

- CNNs formulated in the spectral domain
  - $\Delta = \Phi \Lambda \Phi^T$
- Spectral Convolution
  - $X \star Y = \Phi_r(\hat{X} \circ \hat{Y})\Phi_c^T, \ \hat{X} = \Phi_r X \Phi_c^T$
  - $X \star Y = \tau(\tilde{\lambda}_c, \tilde{\lambda}_r)X,$
    $\rightarrow \hat{Y}_{k,k'} = \tau(\lambda_{c,k}, \lambda_{r,k'})$, using Chevychev polynomial filters of degree p
    $\rightarrow \tau_\Theta(\lambda_c, \lambda_r) = \sum_{j,j'}^{p} \theta_{jj'} T_j(\tilde{\lambda_c}) T_{j'}(\tilde{\lambda_r})$
  - $\tilde{\Delta}_c = 2\lambda_{c,n}^{-1}\Delta_c - \mathrm{I}, \ \tilde{\Delta}_r = 2\lambda_{r,m}^{-1}\Delta_r - \mathrm{I}$
  - $\tilde{\Lambda}_c = 2\lambda_{c,n}^{-1}\Lambda_c - \mathrm{I}, \ \tilde{\Lambda}_r = 2\lambda_{r,m}^{-1}\Lambda_r - \mathrm{I}$
- A multi-graph convolutional layer(q' input, q output)

$$\rightarrow \tilde{X}_l = \xi\left(\sum_{l'=1}^{q'} X_{l'} \star Y_{ll'}\right)$$

$$= \xi\left(\sum_{l'=1}^{q'} \sum_{j,j'=0}^{p} \theta_{jj',ll'} T_j(\tilde{\Delta}_c) X_{l'} T_{j'}(\tilde{\Delta}_r)\right), l = 1, \ldots, q$$

$\rightarrow$ input $(m \times n \times q')$, output $(m \times n \times q)$, ξ is non-linearity

- Separable convolution (factorized version $X = WT^T$)
  - $\tilde{w}_l = \sum_{j=0}^{p} \theta_j^r T_j(\tilde{\Delta_c}) w_l$, $\tilde{h}_l = \sum_{j=0}^{p} \theta_j^c T_j(\tilde{\Delta_c}) h_l$, l=1,...,r
  - $\tilde{w}_l = \xi \left(\sum_{j=0}^{p} \theta_j^r T_j(\tilde{\Delta_r}) w_l \right)$, $\tilde{h}_l = \xi \left(\sum_{j=0}^{p} \theta_j^c T_j(\tilde{\Delta_c}) h_l \right)$
- Loss function (LSTM)

$$l(\theta, \sigma) = min \ \|X_{\theta,\sigma}^{(T)}\|_{\mathcal{G}_r}^2 + \|X_{\theta,\sigma}^{(T)}\|_{\mathcal{G}_c}^2 + \frac{\mu}{2} \|\Omega \circ (X_{\theta,\sigma}^{(T)} - Y)\|_F^2$$

$$min \ \|W_{\theta,\sigma}^{(T)}\|_{\mathcal{G}_r}^2 + \|H_{\theta,\sigma}^{(T)}\|_{\mathcal{G}_c}^2 + \frac{\mu}{2} \|\Omega \circ (W_{\theta,\sigma}^{(T)} (H_{\theta,\sigma}^{(T)})^T - Y)\|_F^2$$