Summary of the process for full OnPLS

Jin Kweon

1/16/2019

1.Extract globally joint variations:

Apply SVD for all pairs of matrices X_i for i = 1, . . . , n, $(V_{i,j}\Sigma_{i,j}\hat{\hat{W}}_{i,j} = X_j^T X_i)$ for i $\neq$ j.

(Here, we do not really have to worry getting the same number of rows for every matrices, since we have a lot of data available nowadays!)

And, then, we need to augment the weighted matrices for another SVD. (below)

$\hat{W}_i\Sigma_i V_i^T = [\hat{\hat{W}}_{i,1}|$ ... $|\hat{\hat{W}}_{i,n}]$, where $\hat{\hat{W}}_{i,j}$ has the dimensions of $N_i \times A_{ij}$, where A_{ij} is the number of joint compoenents between X_i (M $\times$ N_i) and X_j (M $\times$ N_j).

$\hat{W}_i$ represents the part of X_i for globally jointed variations with all other matrices. (all the variation that X_i shares with all other matrices) And, then, we could simply repeat this for all i = 1, . . . , n, to get $\hat{W}_i$ for all i.

Here, it is important to realize that we started from double approximation of weighted matrix ($\hat{\hat{W}}_{i,j}$) to get real weight matrix W_i.

We need $\hat{W}_i$ to separate locally joint and unique variation. (Remember that the locally joint and unique variation of X_i and $\hat{W}_i$ are orthogonal.)

2. Extract locally joint and unique variations:

Our goal is really to get the score matrix T_i such that $X_{G,i} = T_i X_i^T$. As, $\hat{X}_{G,i} = \hat{T}_i\hat{W}_i^T = \hat{X}_i\hat{W}_i\hat{W}_i^T$, so, $\hat{T}_i$ $= X_i\hat{W}_i = (X_{G,i} + X_{LU,i} + E_i)\hat{W}_i$.

And, to make the second term $(X_{LU,i}\hat{W}_i)$ zero since this is not globally joint, we want to find the maximal overlap between uncorrected score T and $t_{LU,i}$, so, maximize the norm of the covariance between these two; so maximize $||\hat{T}_i^t t_{LU,i}|| = w_{LU,i}^T\hat{X}_{LU,i}^T\hat{T}_i\hat{T}_i^T\hat{X}_{LU,i}w_{LU,i}$ repsect to w_{LU,i}.

So, here, since we have the constraint $w_{LU,i}^T w_{LU,i} = 1$, we can use Langrange multiplier λ and take the derivative. Thus, maximum can be achieved if $w_{LU,i}$ is the eigenvector of $\hat{X}_{LU,i}^T\hat{T}_i\hat{T}_i^T\hat{X}_{LU,i}w_{LU,i} = \lambda_{max}w_{LU,i}$.

So, from there, we could get $t_{LU,i} = X_i w_{LU,i}$, with the loading vector $P_{LU,i} = \frac{X_i^T t_{LU,i}}{t_{LU,i}^T t_{LU,i}}$, and thus, the locall joint and unique varition is X_i - $t_{LU,i}P_{LU,i}^T$.

After, we want to maximize the sum of covariances between all connected scores such that $\sum_{i=1}^n\sum_{j=1}^n c_{i,j}t_i^T t_j$, where $i \neq j$ and $||w_i|| = 1$, and $c_{i,j} = 1$ if X_i and X_j are connected and 0 otherwise.

3. Distinguish locally joint and unique variations:

The idea is that globally joint variation is the variation for all n matrices, locally joint is the variation for 2 ~ n-1 matrices, and unique is the variation for only one matrix.