Decomposed Principal Component Analysis for Obtaining Interpretable Solutions

Kohei Adachi, Osaka University, Japan
Nickolay T. Trendafilov, The Open University, UK

1. Formulation
For an $n$-observations $\times m$-variables data matrix $X$, principal component analysis (PCA) can be formulated as minimizing $f(F, A) = \|X - FA\|^2$ over $F$ ($n \times p$) and $A$ ($m \times p$) subject to $FF = AA'$ with the number of components $p \leq \min(n, m)$ and $\alpha$ a given positive constant. A drawback of PCA is that the resulting $A$ is not necessarily interpretable. For dealing with this problem, we consider constraining $A$ as

$$A = (a_{jk})$$

having a property easily interpretable

and propose to attain this constrained PCA by using the fact that the PCA loss function can be decomposed

$$f(F, A) = \|X - FC' + FC' - FA\|^2 = \|X - FC'\|^2 + \alpha \times \|C - A\|^2$$

with $C = (c_{jk}) = X'(FF')^{-1} = \alpha^{-1}XF$. It should be noticed that minimizing (2) over $A$ subject to (1) amounts to the minimization of a simple function

$$g(A) = \|C - A\|^2,$$

which is expected to be easily attained. We use the term decomposed PCA (DPCA) for referring to a family of the procedures in which decomposition (2) is used for obtaining the loading matrix $A$ constrained as (1).

2. Constraints
We present three examples of (1) and show the optimal $A = (a_{jk})$ which minimizes (2) for a given $F$.

**Sparseness.** Under the constraint that $A$ includes $q$ zero loadings, (2) is minimized by setting $a_{jk} = 0$ if $|c_{jk}|$ is less than the $(q+1)$-th smallest of the absolute values of the elements in $C$ and setting $a_{jk} = c_{jk}$ otherwise. This procedure is different from the existing sparse PCA with penalty functions (Trendafilov, 2013).

**Threshold.** Under the constraint that $|a_{jk}|$ must be 0 or not be less than a positive value $t$, (2) is minimized by setting $a_{jk} = c_{jk}$ if $|c_{jk}| > t$, $a_{jk} = 0$ if $|c_{jk}| < t/2$, and $a_{jk} = t$ otherwise.

**Cluster.** Under the constraint that each variable should load only one factor, (2) is minimized by setting $a_{jk} = c_{jk}$ if $|c_{jk}| = \max_{1 \leq i \leq p}|c_{ij}|$ and $a_{jk} = 0$ otherwise.

3. Algorithm
Given $S = n^{-1}XX'$, the solution of constrained $A$ can be obtained by the following algorithm:

1. Initialize $A$.
2. Perform the eigenvalue decomposition $A'SA = LA'L'$ and obtain $C = (c_{jk}) = SAL^{-1}L'$.
3. Update $A$ as illustrated in Section 2.
4. Finish if convergence is reached; otherwise, go back to [2].

This algorithm is derived from the fact that the minimization of (2) over $F$ for a given $A$ is attained for $F = \alpha^{-1/2}KL'$ under the constraint $FF = nI_m$, where $KL'$ are obtained by the singular value decomposition (SVD) $\alpha^{-1/2}XA = KAL'$ with $K$ and $L$ being column-orthonormal and $A$ being a diagonal matrix. This SVD implies $K = \alpha^{-1/2}XLAL^{-1}$, which can be substituted in $F = \alpha^{-1/2}KL'$ for rewriting it as

$$F = XAL^{-1}L'.$$

Using (4) and $FF = \alpha I_m$, (2) is rewritten as $f(A) = ntrS - 2ntrSAL^{-1}L'A' + \alpha trAA'$, which is as a function of a single unknown $A$ with $F$ vanishing. Further, the substitution of (4) into $C = n^{-1}XF$, it can be rewritten as $C = SAL^{-1}L'$ in Step 2, which can be obtained using the decomposition $A'SA = LA'L'$ following from the above SVD. This algorithm can be viewed as a PCA version of Adachi's (2012) factor analysis procedure.

References

Acknowledgement
The works were partially supported by Grant #4387 by The Great Britain Sasakawa Foundation.