Supplementary Materials for EVD Dualdating based Online Subspace Learning

In many online applications, it is impossible to store the original data because of the limitation of the physical medium and the consideration about efficiency. Described in a mathematical form, this means that the original data matrix A is unobtainable and replaced by its best rank-k approximation which can be calculated by U_k and Λ_k. Theorem 2 proofs that under the *low-rank-plus-shift* structure, when A is replaced by $best_k(A)$, the information discarded will also be discarded after EVD dualdating. In other words, EVD dualdating is an optimal rank-k estimator in the sequential usage.

Theorem 2. *Given a matrix $A \in \mathbb{R}^{d\times n}$, with its best-$k$ approximation $\hat{A}$, the deleted data $D \in \mathbb{R}^{d\times p}$ from A, the added data $B \in \mathbb{R}^{d\times m}$, $d > n > p$. Let $C = \begin{bmatrix} A & \underline{D} \end{bmatrix}$ be the remained data from A, $F = \begin{bmatrix} C & D & B \end{bmatrix}$ be the full data, $E = \begin{bmatrix} C & B \end{bmatrix}$ be the final data, where the underline means deletion. Let $\hat{C} = \begin{bmatrix} best_k(A) & \underline{D} \end{bmatrix}$ be the remained matrix after deleting columns corresponds to D from A's best-k approximation, $\hat{E} = \begin{bmatrix} \hat{C} & B \end{bmatrix}$ be the final data from $\hat{A}$. Assume F satisfies the low-rank-plus-shift structure, i.e.*

$$F^TF = X_F + \sigma^2 I_n, \qquad \sigma > 0,$$

where X_F is symmetric and positive semidefinite with $rank(X_F) = k$, then

$$best_k(S_E) = best_k(\hat{S_E}).$$

Proof.

$$F^TF = \begin{bmatrix} C^TC & C^TD & C^TB \\ D^TC & D^TD & D^TB \\ B^TC & B^TD & B^TB \end{bmatrix} = X_F + \sigma^2 I_{n+m}.$$

Then,

$$X_F = F^TF - \sigma^2 I_{n+m} = \begin{bmatrix} C^TC - \sigma^2 I_{n-p} & C^TD & C^TB \\ D^TC & D^TD - \sigma^2 I_p & D^TB \\ B^TC & B^TD & B^TB - \sigma^2 I_m \end{bmatrix}.$$

Since $rank(X_F) = k$, it follows that:

$$\begin{aligned} rank(C^TC - \sigma^2 I_{n-p}) &= k_C \le k, \\ rank(D^TD - \sigma^2 I_p) &= k_D \le k, \\ rank(B^TB - \sigma^2 I_m) &= k_B \le k. \end{aligned}$$

Let the EVDs of $C^TC - \sigma^2 I_{n-p}$, $D^TD - \sigma^2 I_p$, and $B^TB - \sigma^2 I_m$ are:

$$\begin{aligned} C^TC - \sigma^2 I_{n-p} &= V_C \begin{bmatrix} \Lambda'_C & \\ & 0 \end{bmatrix} V_C^T, \\ D^TD - \sigma^2 I_p &= V_D \begin{bmatrix} \Lambda'_D & \\ & 0 \end{bmatrix} V_D^T, \\ B^TB - \sigma^2 I_m &= V_B \begin{bmatrix} \Lambda'_B & \\ & 0 \end{bmatrix} V_B^T, \end{aligned}$$

where $\Lambda'_C = \Sigma_C^2 - \sigma^2 I_{k_C}$, $\Lambda'_D = \Sigma_D^2 - \sigma^2 I_{k_D}$, $\Lambda'_B = \Sigma_B^2 - \sigma^2 I_{k_B}$ and $\Sigma_C \in \mathbb{R}^{k_C\times k_C}$, $\Sigma_D \in \mathbb{R}^{k_D\times k_D}$, $\Sigma_B \in \mathbb{R}^{k_B\times k_B}$ are from the singular values of C, D and B.

So the SVDs of C, D, and B can be written as:

$$
\begin{aligned}
C &= U_C \begin{bmatrix} \Sigma_C & \\ & \sigma I_{t_C} \end{bmatrix} V_C^T = \begin{bmatrix} U_{Ck_C} & U_{Cl} \end{bmatrix} \begin{bmatrix} \Sigma_C & \\ & \sigma I_{t_C} \end{bmatrix} \begin{bmatrix} V_{Ck_C} & V_{Cl} \end{bmatrix}^T, \\
D &= U_D \begin{bmatrix} \Sigma_D & \\ & \sigma I_{t_D} \end{bmatrix} V_D^T = \begin{bmatrix} U_{Dk_D} & U_{Dl} \end{bmatrix} \begin{bmatrix} \Sigma_D & \\ & \sigma I_{t_D} \end{bmatrix} \begin{bmatrix} V_{Dk_D} & V_{Dl} \end{bmatrix}^T, \\
B &= U_B \begin{bmatrix} \Sigma_B & \\ & \sigma I_{t_B} \end{bmatrix} V_B^T = \begin{bmatrix} U_{Bk_B} & U_{Bl} \end{bmatrix} \begin{bmatrix} \Sigma_B & \\ & \sigma I_{t_B} \end{bmatrix} \begin{bmatrix} V_{Bk_B} & V_{Bl} \end{bmatrix}^T,
\end{aligned}
$$

where $U_{Ck_C} \in \mathbb{R}^{d\times k_C}$, $U_{Cl} \in \mathbb{R}^{d\times t_C}$, $U_{Dk_D} \in \mathbb{R}^{d\times k_D}$, $U_{Dl} \in \mathbb{R}^{d\times t_D}$, $U_{Bk_B} \in \mathbb{R}^{d\times k_B}$, $U_{Bl} \in \mathbb{R}^{d\times t_B}$, $t_C = rank(C) - k_C$, $t_D = rank(D) - k_D$, $t_B = rank(B) - k_B$.

Then,

$$
\begin{aligned}
&\begin{bmatrix} V_C & & \\ & V_D & \\ & & V_B \end{bmatrix}^T X_F \begin{bmatrix} V_C & & \\ & V_D & \\ & & V_B \end{bmatrix} \\
= &\begin{bmatrix} \begin{bmatrix} \Lambda'_C & \\ & 0 \end{bmatrix} & V_C^T C^T D V_D & V_C^T C^T B V_B \\ V_D^T D^T C V_C & \begin{bmatrix} \Lambda'_D & \\ & 0 \end{bmatrix} & V_D^T D^T B V_B \\ V_B^T B^T C V_C & V_B^T B^T D V_D & \begin{bmatrix} \Lambda'_B & \\ & 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \Lambda'_C & 0 & S_1 & S_2 & S_5 & S_6 \\ 0 & 0 & S_3 & S_4 & S_7 & S_8 \\ S_1^T & S_3^T & \Lambda'_D & 0 & S_9 & S_{10} \\ S_2^T & S_4^T & 0 & 0 & S_{11} & S_{12} \\ S_5^T & S_7^T & S_9^T & S_{11}^T & \Lambda'_B & 0 \\ S_6^T & S_8^T & S_{10}^T & S_{12}^T & 0 & 0 \end{bmatrix}.
\end{aligned}
$$

Because X_F is symmetric positive semidefinite and $rank(X_F) = k$, according the determination condition of a symmetric positive semidefinite matrix, there exists a k-order principal minor is positive, and the determinant of arbitrary $(k+1)$-order principal minor is zero. It follows that $S_2 = 0$, $S_3 = 0$, $S_4 = 0$, $S_6 = 0$, $S_7 = 0$, $S_8 = 0$, $S_{10} = 0$, $S_{11} = 0$, $S_{12} = 0$, and $k_C + k_D + k_B = k$.

Thus, we have

$$
\begin{aligned}
\begin{bmatrix} U_{Ck_C}\Sigma_C & \sigma U_{Cl} \end{bmatrix}^T \begin{bmatrix} U_{Dk_D}\Sigma_D & \sigma U_{Dl} \end{bmatrix} &= \begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix}, \\
\begin{bmatrix} U_{Ck_C}\Sigma_C & \sigma U_{Cl} \end{bmatrix}^T \begin{bmatrix} U_{Uk_B}\Sigma_B & \sigma U_{Bl} \end{bmatrix} &= \begin{bmatrix} S_5 & 0 \\ 0 & 0 \end{bmatrix}, \\
\begin{bmatrix} U_{Dk_D}\Sigma_C & \sigma U_{Dl} \end{bmatrix}^T \begin{bmatrix} U_{Bk_B}\Sigma_B & \sigma U_{Bl} \end{bmatrix} &= \begin{bmatrix} S_9 & 0 \\ 0 & 0 \end{bmatrix},
\end{aligned}
$$

which lead to

$$
\begin{aligned}
&U_{Ck_C} \perp U_{Dl}, U_{Cl} \perp U_{Dk_D}, U_{Cl} \perp U_{Dl}, \\
&U_{Ck_C} \perp U_{Bl}, U_{Cl} \perp U_{Bk_B}, U_{Cl} \perp U_{Bl}, \\
&U_{Dk_D} \perp U_{Bl}, U_{Dl} \perp U_{Bk_B}, U_{Dl} \perp U_{Bl}.
\end{aligned}
$$

Let $\hat{U}_A$ be an orthonormal basis of

$$
\mathcal{L}\left(\begin{bmatrix} U_{Ck_C} & U_{Dk_D} \end{bmatrix}\right) \cap \mathcal{L}^{\perp}\left(\begin{bmatrix} U_{Cl} & U_{Dl} \end{bmatrix}\right),
$$

where $\mathcal{L}(\cdot)$ denotes the column space of a matrix, and $\mathcal{L}^{\perp}(\cdot)$ denotes the orthonormal complement.

Obviously, $rank(\hat{U}_A) \leq k$, so in order to construct the first k eigen vectors, we append a few mutually orthogonal vectors from U_{Cl} and U_{Dl} to $\hat{U}_A$. Divide U_{Cl} and U_{Dl} into two parts, $U_{Cl} = \begin{bmatrix} U_{Cl}^{A1} & U_{Cl}^{A2} \end{bmatrix}$, $U_{Dl} = \begin{bmatrix} U_{Dl}^{A1} & U_{Dl}^{A2} \end{bmatrix}$, satisfying that the first k eigen vectors of A are

$$U_{Ak} = \begin{bmatrix} \hat{U}_A & U_{Cl}^{A1} & U_{Dl}^{A1} \end{bmatrix}.$$

Then, we have

$$AA^T = \hat{A}\hat{A}^T + \sigma^2 U_{Cl}^{A2} U_{Cl}^{A2^T} + \sigma^2 U_{Dl}^{A2} U_{Dl}^{A2^T},$$

which leads to

$$EE^T = \hat{E}\hat{E}^T + \sigma^2 U_{Cl}^{A2} U_{Cl}^{A2^T} + \sigma^2 U_{Dl}^{A2} U_{Dl}^{A2^T},$$

Similarly, let $\hat{U}_E$ be an orthonormal basis of $\mathcal{L}\left(\begin{bmatrix} U_{Ck_C} & U_{Bk_B} \end{bmatrix}\right) \mathcal{L}^{\perp}\left(\begin{bmatrix} U_{Cl} & U_{Bl} \end{bmatrix}\right)$, divide U_{Cl} and U_{Bl} into two parts, $U_{Cl} = \begin{bmatrix} U_{Cl}^{E1} & U_{Cl}^{E2} \end{bmatrix}$, $U_{Bl} = \begin{bmatrix} U_{Bl}^{E1} & U_{Bl}^{E2} \end{bmatrix}$, then the first k singular vectors of E are

$$U_{Ek} = \begin{bmatrix} \hat{U}_E & U_{Cl}^{E1} & U_{Bl}^{E1} \end{bmatrix}.$$

Assuming that the intersection of of U_{Cl}^{A2} and U_{Cl}^{E1} is U_{Cl}^{I} (null or non-null), it can be written that $U_{Cl}^{A2} = \begin{bmatrix} U_{Cl}^{I} & U_{Cl}^{A} \end{bmatrix}$, $U_{Cl}^{E1} = \begin{bmatrix} U_{Cl}^{I} & U_{Cl}^{E} \end{bmatrix}$. Let $rank(U_{Cl}^{C}) = k_s$, the following holds

$$\begin{aligned}
& U_{Ek}^T U_{Cl}^{A2} U_{Cl}^{A2^T} U_{Ek} \\
= & \begin{bmatrix} \hat{U}_E & U_{Cl}^{E1} & U_{Bl}^{E1} \end{bmatrix}^T U_{Cl}^{A2} U_{Cl}^{A2^T} \begin{bmatrix} \hat{U}_E & U_{Cl}^{E1} & U_{Bl}^{E1} \end{bmatrix} \\
= & \begin{bmatrix} 0 & \begin{bmatrix} U_{Cl}^{I^T} \\ U_{Cl}^{A^T} \end{bmatrix} \begin{bmatrix} U_{Cl}^{I} & U_{Cl}^{E} \end{bmatrix} & 0 \end{bmatrix}^T \begin{bmatrix} 0 & \begin{bmatrix} U_{Cl}^{I^T} \\ U_{Cl}^{A^T} \end{bmatrix} \begin{bmatrix} U_{Cl}^{I} & U_{Cl}^{E} \end{bmatrix} & 0 \end{bmatrix} \\
= & \begin{bmatrix} 0 & & \\ & I_{k_s} & \\ & & 0 \end{bmatrix}_{k \times k},
\end{aligned}$$

$$\begin{aligned}
& U_{Ek}^T U_{Dl}^{A2} U_{Dl}^{A2^T} U_{Ek} \\
= & \begin{bmatrix} \hat{U}_E & U_{Cl}^{E1} & U_{Bl}^{E1} \end{bmatrix}^T U_{Dl}^{A2} U_{Dl}^{A2^T} \begin{bmatrix} \hat{U}_E & U_{Cl}^{E1} & U_{Bl}^{E1} \end{bmatrix} = 0_{k \times k}.
\end{aligned}$$

So

$$U_{Ek}^T EE^T U_{Ek} = U_{Ek}^T \hat{E}\hat{E}^T U_{Ek} + \sigma^2 \begin{bmatrix} 0 & & \\ & I_{k_s} & \\ & & 0 \end{bmatrix}_{k \times k}.$$

Finally, we can conclude that

$$best_k(S_E) = best_k(\hat{S_E}).$$

completing the proof.

□