Supplementary Material

Mourad Khayati, Ines Arous, Philippe Cudré-Mauroux
eXascale Infolab, University of Fribourg
Fribourg, Switzerland
{firstname.lastname}@unifr.ch

I. Background

A. Centroid Decomposition (CD)

Example 1 (Centroid Decomposition): To illustrate the application of the CD algorithm, consider the input matrix, $\mathbf{X}$, that contains two time series of five elements each:

$$\mathbf{X} = \begin{bmatrix} 5 & 1 \\ -10 & 5 \\ -9 & 4 \\ 4 & 6 \\ 2 & -4 \end{bmatrix}$$

Among all sign vectors, the sign vector that maximizes $\|\mathbf{X}^T \cdot Z\|$ is $Z_1 = \{-1, 1, 1, -1, -1\}^T$. Z_1 is used to compute the first column of $\mathbf{R}$ (and $\mathbf{L}$) during iteration 1 as follows:

$$R_{*1} = \frac{\mathbf{X}^T \cdot Z_1}{\|\mathbf{X}^T \cdot Z_1\|} = \begin{bmatrix} -0.98 \\ 0.19 \end{bmatrix}; \quad L_{*1} = \mathbf{X} \cdot R_{*1} = \begin{bmatrix} -4.7 \\ 10.78 \\ 9.6 \\ -2.74 \\ -2.74 \end{bmatrix}$$

Similarly, the second column of $\mathbf{R}$ (and $\mathbf{L}$) are respectively computed using Z_2 which is derived from $\mathbf{X} - L_{*1} \cdot R_{*1}^T$. The resulting decomposition produced by CD is (only two decimals are shown):

$$\mathbf{X} = \begin{bmatrix} 5 & 1 \\ -10 & 5 \\ -9 & 4 \\ 4 & 6 \\ 2 & -4 \end{bmatrix} = \underbrace{\begin{bmatrix} -4.7 & 1.96 \\ 10.78 & 2.94 \\ 9.6 & 2.15 \\ -2.74 & 6.66 \\ -2.74 & -3.53 \end{bmatrix}}_{\mathbf{L}} \cdot \underbrace{\begin{bmatrix} -0.98 & 0.19 \\ 0.19 & -0.98 \end{bmatrix}}_{\mathbf{R}^T}$$

II. Anticipatory Sign Vector

Lemma 1 (Weight vectors are incremental): Let $Z^{(k)}$ be Z at iteration k, P the set of positions of the elements flipped in $Z^{(k)}$ and let v_i be the i-th weight value in V. For any two consecutive iterations of sign vectors, the weight vectors are linearly dependent, i.e.,

$$v_i^{(k+1)} = v_i^{(k)} - 2 \times \sum_{p \in P \setminus \{i\}} (X_{i*} \cdot X_{p*}^T)$$

Proof: By definition of the weight vector, we have:

$$V^{(k)} = diag^{=0}(\mathbf{X} \cdot \mathbf{X}^T) \cdot Z^{(k)}$$
$$V^{(k+1)} = diag^{=0}(\mathbf{X} \cdot \mathbf{X}^T) \cdot Z^{(k+1)} \quad (1)$$

Let U be a vector with the same length as $Z^{(k)}$ where for each $p \in P$, $U_p = 1$ and all other elements are 0. Using U we compute $Z^{(k+1)}$ as follows

$$Z^{(k+1)} = Z^{(k)} - 2 \times U \quad (2)$$

Putting (2) into (1) we get

$$\begin{aligned} V^{(k+1)} &= diag^{=0}(\mathbf{X} \cdot \mathbf{X}^T) \cdot (Z^{(k)} - 2 \times U) \\ &= diag^{=0}(\mathbf{X} \cdot \mathbf{X}^T) \cdot Z^{(k)} - \\ &\qquad 2 \times diag^{=0}(\mathbf{X} \cdot \mathbf{X}^T) \cdot U \\ &= V^{(k)} - 2 \times diag^{=0}(\mathbf{X} \cdot \mathbf{X}^T) \cdot U \quad (3) \end{aligned}$$

Let $col(\mathbf{X}, p)$ be an auxiliary function that returns the p-th column of $\mathbf{X}$. Then, from (3) we get

$$\begin{aligned} V^{(k+1)} &= V^{(k)} - 2 \times \sum_{p \in P} col(diag^{=0}(\mathbf{X} \cdot \mathbf{X}^T), p) \\ &= V^{(k)} - 2 \times \sum_{p \in P} \begin{bmatrix} X_{1*} \cdot X_{p*}^T \\ X_{2*} \cdot X_{p*}^T \\ \vdots \\ X_{n*} \cdot X_{p*}^T \end{bmatrix} \end{aligned}$$

Thus, $\forall i \in [1, n]$ we have

$$v_i^{(k+1)} = v_i^{(k)} - 2 \times \sum_{p \in P \setminus \{i\}} (X_{i*} \cdot X_{p*}^T) \quad (4)$$

■

In the particular case where only one sign is flipped, $\forall i \in [1, n] \setminus \{p\}$, Lemma 1 can be rewritten as follows:

$$v_i^{(k+1)} = v_i^{(k)} - 2 \times (X_{i*} \cdot X_{p*}^T) \quad (5)$$

with $v_p^{(k+1)} = v_p^{(k)}$

Example 2: To illustrate the incremental computation of the weight vectors, consider the input matrix of our running example, i.e.,

$$\mathbf{X} = \begin{bmatrix} 5 & 1 \\ -10 & 5 \\ -9 & 4 \\ 4 & 6 \\ 2 & -4 \end{bmatrix}.$$

For the sake of simplicity, we illustrate the case where only one sign flip is performed. First, Z is initialized with 1s, i.e., $Z^{(1)} = \{1, 1, 1, 1, 1\}^T$ and the initial weight vector is computed iteratively to get $V^{(1)} = \{-54, 15, 23, -12, -84\}^T$. Three elements of $Z^{(1)}$ have a different sign from the corresponding elements in $V^{(1)}$ and among them the element in the 5th position has the highest absolute value. Using $p = 5$, the next weight vector is incrementally computed (using (5)) as follows

$$\begin{aligned}
v_1 &= \text{-}54 - 2 \times (\begin{bmatrix}5 & 1\end{bmatrix} \times \begin{bmatrix}2\\ \text{-}4\end{bmatrix}) = \text{-}66\\
v_2 &= 15 - 2 \times (\begin{bmatrix}\text{-}10 & 5\end{bmatrix} \times \begin{bmatrix}2\\ \text{-}4\end{bmatrix}) = 95\\
v_3 &= 23 - 2 \times (\begin{bmatrix}\text{-}9 & 4\end{bmatrix} \times \begin{bmatrix}2\\ \text{-}4\end{bmatrix}) = 91\\
v_4 &= \text{-}12 - 2 \times (\begin{bmatrix}4 & 6\end{bmatrix} \times \begin{bmatrix}2\\ \text{-}4\end{bmatrix}) = 20\\
v_5 &= \text{-}84
\end{aligned}$$

i.e.,

$$Z^{(2)} = \begin{bmatrix}1\\1\\1\\1\\\text{-}1\end{bmatrix} \quad \text{and} \quad V^{(2)} = \begin{bmatrix}\text{-}66\\95\\91\\20\\\text{-}84\end{bmatrix}.$$

III. Incremental Centroid Decomposition (InCD)

Lemma 2 (Matrix Similarity): Let $\widetilde{\mathbf{X}}$ be the matrix resulting from incrementing an $n \times m$ matrix $\mathbf{X}$ with an $r \times m$ matrix $\Delta\mathbf{X}$ and let A_{i*} be the i-th row of $\Delta\mathbf{X}$. Let also $Z \in \{\text{-}1, 1\}^n$ and $\tilde{Z} \in \{\text{-}1, 1\}^{n+1}$ be two sign vectors. Then, the following holds:

$$\max\|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}\| = \max\|\mathbf{X}^T \cdot Z\| + \sum_{i=1}^{r} \|A_{i*}\|$$

Proof: We assume that $\mathbf{X}^T \cdot Z$ and the added rows are linearly dependent. However, in practice, the similarity holds for arbitrary row updates (see the correctness experiment in Table I).

Let $\mathbf{0}$ be a matrix of zeros, $\mathbf{Y}$ be an $(n+r) \times m$ matrix consisting of $\mathbf{X}$ appended with $\mathbf{0}_{r \times m}$ and let $\mathbf{M}$ be an $(n + r) \times m$ matrix consisting of $\mathbf{0}_{n \times m}$ incremented with $\Delta\mathbf{X}$. Then, we have the following:

$$\begin{aligned}
\widetilde{\mathbf{X}} &= \begin{bmatrix}\mathbf{X}\\ \Delta\mathbf{X}\end{bmatrix} = \begin{bmatrix}\mathbf{X}\\ \mathbf{0}_{r \times m}\end{bmatrix} + \begin{bmatrix}\mathbf{0}_{n \times m}\\ \Delta\mathbf{X}\end{bmatrix}\\
&= \begin{bmatrix}\mathbf{X}\\ 0_{1*}\\ \vdots\\ 0_{r*}\end{bmatrix} + \begin{bmatrix}\mathbf{0}_{n \times m}\\ A_{1*}\\ \vdots\\ A_{r*}\end{bmatrix}\\
&= \mathbf{Y} + \mathbf{M}
\end{aligned} \tag{6}$$

By transposing both sides of (6), multiplying each side by $\tilde{Z}$ and normalizing, we get

$$\begin{aligned}
\|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}\| &= \|(\mathbf{Y} + \mathbf{M})^T \cdot \tilde{Z}\|\\
&= \|\mathbf{Y}^T \cdot \tilde{Z} + \mathbf{M}^T \cdot \tilde{Z}\|
\end{aligned} \tag{7}$$

The computation of the right hand side of (7) gives

$$\begin{aligned}
\mathbf{Y}^T \cdot \tilde{Z} &= \begin{bmatrix}\mathbf{X}^T 0_{1*}^T \cdots 0_{r*}^T\end{bmatrix} \cdot \begin{bmatrix}Z\\ \tilde{z}_{n+1}\\ \vdots\\ \tilde{z}_{n+r}\end{bmatrix}\\
&= \mathbf{X}^T \cdot Z + \tilde{z}_{n+1} \times 0_{1*}^T + \cdots + \tilde{z}_{n+r} \times 0_{r*}^T\\
&= \mathbf{X}^T \cdot Z
\end{aligned} \tag{8}$$

and

$$\begin{aligned}
\mathbf{M}^T \cdot \tilde{Z} &= \begin{bmatrix}\mathbf{0}_{n \times m}\\ A_{1*}\\ \vdots\\ A_{r*}\end{bmatrix}^T \cdot \begin{bmatrix}\tilde{z}_1\\ \vdots\\ \tilde{z}_{n+r}\end{bmatrix}\\
&= \sum_{i=1}^{n} \tilde{z}_i \times 0_{i*} + \tilde{z}_{n+1} \times A_{1*}^T + \cdots + \tilde{z}_{n+r} \times A_{r*}^T\\
&= \tilde{z}_{n+1} \times A_{1*}^T + \cdots + \tilde{z}_{n+r} \times A_{r*}^T\ .
\end{aligned} \tag{9}$$

Putting (8) and (9) into (7) gives

$$\|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}\| = \|\mathbf{X}^T \cdot Z + \tilde{z}_{n+1} \times A_{i*}^T + \cdots + \tilde{z}_{n+r} \times A_{r*}^T\| \tag{10}$$

Equation (10) is valid for all $\tilde{Z}$ and Z including the maximizing one, i.e., Z_{max}. It follows:

$$\|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}_{max}\| = \|\mathbf{X}^T \cdot Z_{max} + \tilde{z}_{n+1} \times A_{i*}^T + \cdots + \tilde{z}_{n+r} \times A_{r*}^T\| \tag{11}$$

Since $\mathbf{X}^T \cdot Z$ and A_{i*}, $\forall i \in \{1, \cdots, r\}$, are linearly dependent, it follows:

$$\|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}_{max}\| = \|\mathbf{X}^T \cdot Z_{max}\| + \sum_{i=1}^{r} \|\tilde{z}_{n+i} \times A_{i*}^T\| \tag{12}$$

Since $\tilde{Z}_{max}$ is the maximizing vector of $\|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}\|$, and Z_{max} is the maximizing sign vector of $\|\mathbf{X}^T \cdot Z\|$, then we get

$$\max\|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}\| = \max\|\mathbf{X}^T \cdot Z\| + \sum_{i=1}^{r} \|\tilde{z}_{n+i} \times A_{i*}^T\|$$

We have $z_{n+i} = \pm 1$, hence $\forall i \in \{1, \cdots, r\}$, $\|\tilde{z}_{n+i} \times A_{i*}^T\| = \|A_{i*}^T\|$ and $\|A_{i*}^T\| = \|A_{i*}\|$. Therefore, we have

$$\max\|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}\| = \max\|\mathbf{X}^T \cdot Z\| + \sum_{i=1}^{r} \|A_{i*}\|$$

which concludes the proof. ■

Theorem 1 (Correctness): Let $\widetilde{\mathbf{X}}$ be the resulting matrix of incrementing an input matrix $\mathbf{X}$ with $\Delta\mathbf{X}$. Then, InCD returns the sign vector, $\tilde{Z}$, that maximizes $\|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}\|$.

Proof: Using $Z^{(1)}$ as Z at the first iteration of the algorithm, we introduce the two following vectors. Let $\tilde{Z}$ be the resulting sign vector obtained by batch CD (i.e., $\tilde{Z}^{(1)} = [1, \ldots, 1]$) and let $\tilde{Z}_c$ be the resulting sign vector obtained by InCD with (i.e., $\tilde{Z}_c^{(1)} = [Z_c, 1, \ldots, 1]$; Z_c is the cached sign vector). Proving the correctness of InCD boils down to proving that $\underset{\tilde{Z}_c \in \{\text{-}1,1\}^{(n+r)}}{arg\ max} \|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}_c\| \equiv \underset{\tilde{Z} \in \{\text{-}1,1\}^{(n+r)}}{arg\ max} \|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}\|$.

Let $\mathbf{I}$ be an identity matrix, $\mathbf{D}$ be a diagonal matrix containing $\tilde{Z}_c^{(1)}$, i.e., $\mathbf{D} = diag(\tilde{Z}_c^{(1)})$, and let $\widetilde{\mathbf{X}}_D$ be an $(n+r) \times m$ matrix s.t. $\widetilde{\mathbf{X}}_D^T = \widetilde{\mathbf{X}}^T \cdot \mathbf{D}$. Let also $\tilde{Z}_D$ be an $(n+r)$ sign vector s.t $\tilde{Z}_D = \mathbf{D} \cdot \tilde{Z}_c$.

First, we prove the following:

$$\underset{\tilde{Z}_c \in \{-1,1\}^{(n+r)}}{arg\ max} \|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}_c\| \equiv \underset{\tilde{Z}_D \in \{-1,1\}^{(n+r)}}{arg\ max} \|\widetilde{\mathbf{X}}_D^T \cdot \tilde{Z}_D\|$$

Since $\mathbf{D}$ is a signature matrix where the diagonal elements are +1 or -1, then $\mathbf{D} \cdot \mathbf{D} = \mathbf{I}$. It follows:

$$\underset{\tilde{Z}_c \in \{-1,1\}^{(n+r)}}{arg\ max} \|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}_c\| \equiv \underset{\tilde{Z}_c \in \{-1,1\}^{(n+r)}}{arg\ max} \|\widetilde{\mathbf{X}}^T \cdot \mathbf{D} \cdot \mathbf{D} \cdot \tilde{Z}_c\| \tag{13}$$

By definition of $\mathbf{D}$, we have $d_{ii} \times z_i^{(1)} = 1, \forall i \in \{1, \cdots, (n+r)\}$ where $z_i^{(1)} \in \tilde{Z}_c^{(1)}$ which yields $\mathbf{D} \cdot \tilde{Z}_c^{(1)} = [1, \ldots, 1]$. Since $\tilde{Z}_D^{(1)} = \mathbf{D} \cdot \tilde{Z}_c^{(1)}$, we replace the argument $\tilde{Z}_c$ by $\tilde{Z}_D$ and get

$$\begin{aligned}\underset{\tilde{Z}_c \in \{-1,1\}^{(n+r)}}{arg\ max} \|\widetilde{\mathbf{X}}^T \cdot \mathbf{D} \cdot \mathbf{D} \cdot \tilde{Z}_c\| &\equiv \underset{\tilde{Z}_D \in \{-1,1\}^{(n+r)}}{arg\ max} \|(\widetilde{\mathbf{X}}^T \cdot \mathbf{D}) \cdot \tilde{Z}_D\| \\ &\equiv \underset{\tilde{Z}_D \in \{-1,1\}^{(n+r)}}{arg\ max} \|\widetilde{\mathbf{X}}_D^T \cdot \tilde{Z}_D\|\end{aligned} \tag{14}$$

Putting (13) into (14), we get

$$\underset{\tilde{Z}_c \in \{-1,1\}^{(n+r)}}{arg\ max} \|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}_c\| \equiv \underset{\tilde{Z}_D \in \{-1,1\}^{(n+r)}}{arg\ max} \|\widetilde{\mathbf{X}}_D^T \cdot \tilde{Z}_D\| \tag{15}$$

Next, we prove the following

$$\underset{\tilde{Z} \in \{-1,1\}^{(n+r)}}{arg\ max} \|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}\| \equiv \underset{\tilde{Z}_D \in \{-1,1\}^{(n+r)}}{arg\ max} \|\widetilde{\mathbf{X}}_D^T \cdot \tilde{Z}_D\|$$

By definition of $\widetilde{\mathbf{X}}_D$ we have

$$\begin{aligned}\widetilde{\mathbf{X}}_D^T &= \widetilde{\mathbf{X}}^T \cdot \mathbf{D} \\ &= \begin{bmatrix} \mathbf{X} \\ \Delta\mathbf{X} \end{bmatrix}^T \cdot \begin{bmatrix} diag(Z_c) & \mathbf{0}_{n\times r} \\ \mathbf{0}_{r\times n} & \mathbf{I}_{r\times r} \end{bmatrix} \\ &= \begin{bmatrix} (\mathbf{X}^T \cdot diag(Z_c)) & \Delta\mathbf{X}^T \end{bmatrix}\end{aligned} \tag{16}$$

From (16), we can see that $\widetilde{\mathbf{X}}_D$ is $diag(Z_c)^T \cdot \mathbf{X}$ incremented with $\Delta\mathbf{X}$. By applying Lemma 2 on $\widetilde{\mathbf{X}}_D$ we get

$$\begin{aligned}\max\|\widetilde{\mathbf{X}}_D^T \cdot \tilde{Z}_D\| &= \max\|\mathbf{X}^T \cdot diag(Z_c) \cdot Z\| + \sum_{i=1}^{r} \|A_{i*}\| \\ &= \max\|\mathbf{X}^T \cdot Z\| + \sum_{i=1}^{r} \|A_{i*}\|\end{aligned} \tag{17}$$

Using (17) and Lemma 2, we get

$$\max\|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}\| = \max\|\widetilde{\mathbf{X}}_D^T \cdot \tilde{Z}_D\| \tag{18}$$

Since the two equations in (18) use the same initial maximizing sign vector that contains only 1s. Thus, (18) can be rewritten as follows:

$$\underset{\tilde{Z} \in \{-1,1\}^n}{arg\ max} \|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}\| = \underset{\tilde{Z}_D \in \{-1,1\}^n}{arg\ max} \|\widetilde{\mathbf{X}}_D^T \cdot \tilde{Z}_D\| \tag{19}$$

By transitivity of (15) and (19), we get

$$\underset{\tilde{Z}_c \in \{-1,1\}^n}{arg\ max} \|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}_c\| \equiv \underset{\tilde{Z} \in \{-1,1\}^n}{arg\ max} \|\widetilde{\mathbf{X}}^T \cdot \tilde{Z}\|$$

Therefore, InCD computes the same maximizing sign vector as the batch CD, which concludes the proof. ∎

IV. Experimental Evaluation

1) Correctness: In Table. I, we evaluate the correctness of InCD by comparing the centroid value, $\|\mathbf{X}^T \cdot Z\|$, computed by InCD against the one computed by the batch SSV technique. We randomly append r rows to a matrix containing the maximum number of time series (columns) per dataset each of $1k$ values, and compute the two centroid values for different r values. The results show that for all datasets InCD computes the same centroid value as SSV and thus, the correct maximizing sign vector. The result of this experiment confirms the correctness proof of Theorem 1.

TABLE I
Centroid values of InCD vs. Batch CD.

$\|\mathbf{X}^T \cdot Z\|$ / dataset	$r = 10$		$r = 20$		$r = 30$	
	InCD	SSV	InCD	SSV	InCD	SSV
BAFU	2520	2520	2539	2539	2558	2558
MeteoSwiss	1562	1562	1566	1566	1563	1563
Gas	2508	2508	2529	2529	2547	2547
Temperature	1144	1144	1158	1158	1172	1172