Clustering by Connections in the Input Matrix

Jonathan McFadden
University of Washington Tacoma
mcfaddja@uw.edu

Summer 2016

Contents

Part I

Definitions and Descriptions

Chapter 1

Definitions

1.1 Input Matrix

Let $\bar{\bar{\mathbf{L}}} \in \mathbb{R}^{m \times n}$ be a matrix which represents how m terms describe a collection of n objects. We can express this formally as

$$\left(\bar{\bar{\mathbf{L}}}\right)_{ij} \equiv \begin{cases} 1 & \text{if the } i\text{th term describes the } j\text{th object.} \\ 0 & \text{otherwise} \end{cases} \tag{1.1.1}$$

We may also use the alternative definition

$$\left(\bar{\bar{\mathbf{L}}}\right)_{ij} \equiv \begin{cases} x & \text{the } i\text{th term describes the } j\text{th object } x \text{ times} \\ 0 & \text{otherwise} \end{cases} \tag{1.1.1a}$$

In both definitions for the $\left(\bar{\bar{\mathbf{L}}}\right)_{ij}$, we require that $i \in [1, m] \subset \mathbb{Z}^+$ and that $j \in [1, n] \subset \mathbb{Z}^+$.

1.2 SVD Matrices

We will be using both the full SVD of $\bar{\bar{\mathbf{L}}}$ and the κth order approximate SVD of $\bar{\bar{\mathbf{L}}}$ where the order of the approximate SVD, κ, must satisfy the relation

$$\kappa < r$$

with r defined as

$$r = \text{rank}\left[\bar{\bar{\mathbf{L}}}\right] \tag{1.2.1}$$

1.2.1 Full SVD

Let the full singular value decomposition of $\bar{\bar{\mathbf{L}}}$ be given by

$$\bar{\bar{\mathbf{L}}} = \bar{\bar{\mathbf{U}}}\,\bar{\bar{\mathbf{S}}}\,\bar{\bar{\mathbf{V}}}^{\intercal}$$

where $\bar{\bar{\mathbf{U}}} \in \mathbb{R}^{m\times r}$, $\bar{\bar{\mathbf{S}}} \in \mathbb{R}^{r\times r}$, and $\bar{\bar{\mathbf{V}}} \in \mathbb{R}^{n\times r}$ with r as defined as above in (1.2.1). We call $\bar{\bar{\mathbf{U}}}$ the "*Row Matrix*" of $\bar{\bar{\mathbf{L}}}$, $\bar{\bar{\mathbf{S}}}$ the "*Singular Value Matrix*" of $\bar{\bar{\mathbf{L}}}$, and $\bar{\bar{\mathbf{V}}}$ the "*Column Matrix*" of $\bar{\bar{\mathbf{L}}}$.

1.2.2 Approximate SVD

Let the κth order approximate singular value decomposition of $\bar{\bar{\mathbf{L}}}$ be expressed as

$$\bar{\bar{\mathbf{L}}} = \bar{\bar{\mathbf{U}}}_{\kappa}\,\bar{\bar{\mathbf{S}}}_{\kappa}\,\bar{\bar{\mathbf{V}}}_{\kappa}^{\intercal}$$

where $\bar{\bar{\mathbf{U}}}_{\kappa} \in \mathbb{R}^{m\times\kappa}$, $\bar{\bar{\mathbf{S}}}_{\kappa} \in \mathbb{R}^{\kappa\times\kappa}$, and $\bar{\bar{\mathbf{V}}}_{\kappa} \in \mathbb{R}^{n\times\kappa}$ with r as defined as above in (1.2.1). We call $\bar{\bar{\mathbf{U}}}_{\kappa}$ the "*Approximate Row Matrix*" or "*Reduced Row Matrix*" of $\bar{\bar{\mathbf{L}}}$, $\bar{\bar{\mathbf{S}}}_{\kappa}$ the "*Approximate Singular Value Matrix*" or "*Reduced Singular Value Matrix*" of $\bar{\bar{\mathbf{L}}}$, and $\bar{\bar{\mathbf{V}}}_{\kappa}$ the "*Approximate Column Matrix*" or "*Reduced Column Matrix*" of $\bar{\bar{\mathbf{L}}}$.

1.3 Useful Functions

The following function will prove useful for the alternative definitions of the *Connection Matrices* given below. This function indicates if the ith element of $\bar{\bar{\mathbf{L}}}$ can be connected to the jth element of $\bar{\bar{\mathbf{L}}}$ and is denoted by $\delta^{\star}\left(\bar{\bar{\mathbf{L}}}, i, j\right)$. We define this function according to the expression

$$\delta^{\star}\left(\bar{\bar{\mathbf{L}}}, i, j\right) \equiv \begin{cases} 1 & \text{if } \bar{\bar{\mathbf{L}}}_{ij} \neq 0 \\ 0 & \text{otherwise} \end{cases} \tag{1.3.1}$$

with $i \in [1, m] \subset \mathbb{Z}^{+}$ and that $j \in [1, n] \subset \mathbb{Z}^{+}$.

1.4 Connection Matrices

We will create two "*Connection Matrices*" from $\bar{\bar{\mathbf{L}}}$. These matrices represent the connections among all of the row elements of $\bar{\bar{\mathbf{L}}}$ or among all of the column elements of $\bar{\bar{\mathbf{L}}}$. They will provide the information required for clustering either the rows of $\bar{\bar{\mathbf{L}}}$ or for clustering the columns of $\bar{\bar{\mathbf{L}}}$.

1.4.1 Row Connection Matrix

The first of the "*Connection Matrices*" that we will create from $\bar{\bar{\mathbf{L}}}$ is the "*Row Connection Matrix*". This matrix will be used to provide information for clustering the rows of $\bar{\bar{\mathbf{L}}}$. We will denote this matrix by $\bar{\bar{\mathbf{A}}}_{U} \in \mathbb{R}^{m \times m}$ and define it, in terms of $\bar{\bar{\mathbf{L}}}$, according to the expression

$$\bar{\bar{\mathbf{A}}}_{U} = \bar{\bar{\mathbf{L}}}\,\bar{\bar{\mathbf{L}}}^{\intercal} \tag{1.4.1}$$

We may also use the alternate definition

$$\left(\bar{\bar{\mathbf{A}}}_U\right)_{ij} = \sum_{k=1}^{n} \left\{ \delta^{\star}\left(\bar{\bar{\mathbf{L}}}, i, k\right) \bar{\bar{\mathbf{L}}}_{jk} \right\} \tag{1.4.1a}$$

where $i, j \in [1, m] \subset \mathbb{Z}^{+}$.

1.4.2 Column Connection Matrix

The second "*Connection Matrix*" to be created from $\bar{\bar{\mathbf{L}}}$ is the "*Column Connection Matrix*". This matrix will be used to provide information for clustering the columns of $\bar{\bar{\mathbf{L}}}$. We will denote this matrix by $\bar{\bar{\mathbf{A}}}_V \in \mathbb{R}^{n \times n}$ and define it, in terms of $\bar{\bar{\mathbf{L}}}$, according to the expression

$$\bar{\bar{\mathbf{A}}}_V = \bar{\bar{\mathbf{L}}}^{\intercal} \bar{\bar{\mathbf{L}}} \tag{1.4.2}$$

We may also use the alternate definition

$$\left(\bar{\bar{\mathbf{A}}}_V\right)_{ij} = \sum_{k=1}^{m} \left\{ \delta^{\star}\left(\bar{\bar{\mathbf{L}}}, k, i\right) \bar{\bar{\mathbf{L}}}_{kj} \right\} \tag{1.4.2a}$$

where $i, j \in [1, n] \subset \mathbb{Z}^{+}$.

1.5 Relation Sets

We will now define two sets of 2-tuples. The first will describe the relation or relations each row element has with all the other row elements; while the second will describe the relation or relations each column element has with all the other column elements. One of the elements of the 2-tuple represents the other rows or columns that are related to a given row or column. The other element of the 2-tuple describes the strength of those relations.

1.5.1 Row Relation Set

Let $\mathbb{U}$ be a set of m 2-tuples that can be expressed as

$$\mathbb{U} = \Big\{ \{\mathbb{U}_1, \mathbb{M}_1\}, \{\mathbb{U}_2, \mathbb{M}_2\}, \dots, \{\mathbb{U}_m, \mathbb{M}_m\} \Big\}$$

where the $\{\mathbb{U}_i, \mathbb{M}_i\}$ are 2-tuples where the first element of the tuple, $\mathbb{U}_i$, represents which rows in $\bar{\bar{\mathbf{L}}}$ are connected to the ith row of $\bar{\bar{\mathbf{L}}}$; while the second element of the tuple, $\mathbb{M}_i$, describes the strength of those relations, with i such that $i \in [1, m] \subset \mathbb{Z}^+$. We now define $\mathbb{U}_i$ and $\mathbb{M}_i$ concurrently as

$$\mathbb{U}_i \equiv \left\{ j \;\middle|\; j \in [1, m] \subset \mathbb{Z}^+ \text{ and } \left(\bar{\bar{\mathbf{A}}}_U\right)_{ij} \neq 0 \right\} \tag{1.5.1}$$

and

$$\mathbb{M}_i \equiv \left\{ \left(\bar{\bar{\mathbf{A}}}_U\right)_{ij} \;\middle|\; j \in [1, m] \subset \mathbb{Z}^+ \text{ and } \left(\bar{\bar{\mathbf{A}}}_U\right)_{ij} \neq 0 \right\} \tag{1.5.2}$$

respectively. These definitions require that the expression

$$\big|\mathbb{U}_i\big| = \big|\mathbb{M}_i\big|$$

holds for all $i \in [1, m] \subset \mathbb{Z}^+$. Additionally, we have defined $\mathbb{U}_i$ and $\mathbb{M}_i$ such that the strength of the connection to the lth element in $\mathbb{U}_i$ is represented by the lth element of $\mathbb{M}_i$.

1.5.2 Column Relation Set

Let $\mathbb{V}$ be a set of n 2-tuples that can be expressed as

$$\mathbb{V} = \Big\{ \{\mathbb{V}_1, \mathbb{N}_1\}, \{\mathbb{V}_2, \mathbb{N}_2\}, \ldots, \{\mathbb{V}_n, \mathbb{N}_n\} \Big\}$$

where the $\{\mathbb{V}_j, \mathbb{N}_j\}$ are 2-tuples where the first element of the tuple, $\mathbb{V}_j$, represents which columns in $\bar{\bar{\mathbf{L}}}$ are connected to the jth column of $\bar{\bar{\mathbf{L}}}$; while the second element of the tuple, $\mathbb{N}_j$, describes the strength of those relations, with j such that $j \in [1, n] \subset \mathbb{Z}^+$. We now define $\mathbb{V}_j$ and $\mathbb{N}_j$ concurrently as

$$\mathbb{V}_j \equiv \left\{ i \;\middle|\; i \in [1, n] \subset \mathbb{Z}^+ \text{ and } \left(\bar{\bar{\mathbf{A}}}_V\right)_{ji} \neq 0 \right\} \tag{1.5.3}$$

and

$$\mathbb{N}_j \equiv \left\{ \left(\bar{\bar{\mathbf{A}}}_V\right)_{ji} \;\middle|\; i \in [1, n] \subset \mathbb{Z}^+ \text{ and } \left(\bar{\bar{\mathbf{A}}}_V\right)_{ji} \neq 0 \right\} \tag{1.5.4}$$

respectively. These definitions require that the expression

$$\left|\mathbb{V}_j\right| = \left|\mathbb{N}_j\right|$$

holds for all $j \in [1, n] \subset \mathbb{Z}^+$. Additionally, we have defined $\mathbb{V}_j$ and $\mathbb{N}_j$ such that the strength of the connection to the lth element in $\mathbb{V}_j$ is represented by the lth element of $\mathbb{N}_j$.

1.6 Cluster Sets

We now define two sets of sets with the first set of sets representing the row clusters and the second set of sets representing the column clusters. Each set of sets is composed of sets which represent the members of each row or column cluster.

1.6.1 Row Cluster Set

The "*Row Cluster Set*" is a set of $\mathfrak{u} \in \mathbb{Z}^+$ sets with each component set representing one of the $\mathfrak{u}$ row clusters. We will denote this set of sets as $\mathbb{U}^\star$ and formally describe it as

$$\mathbb{U}^\star = \left\{ \mathbb{U}_1^\star, \mathbb{U}_2^\star, \ldots, \mathbb{U}_{\mathfrak{u}}^\star \right\}$$

where the $\mathbb{U}_\mu^\star$ are sets which contain the members of the μth row cluster, with $\mu \in [1, \mathfrak{u}] \subset \mathbb{Z}^+$. Each of the $\mathbb{U}_\mu^\star$ is constructed from the $\{\mathbb{U}_i, \mathbb{M}_i\}$ of its (the $\mathbb{U}_\mu^\star$) constituent elements via a process we will describe later.

1.6.2 Column Cluster Set

The "*Column Cluster Set*" is a set of $\mathfrak{v} \in \mathbb{Z}^+$ sets with each component set representing one of the $\mathfrak{v}$ column clusters. We will denote this set of sets as $\mathbb{V}^\star$ and formally describe it as

$$\mathbb{V}^\star = \left\{ \mathbb{V}_1^\star, \mathbb{V}_2^\star, \ldots, \mathbb{V}_{\mathfrak{v}}^\star \right\}$$

where the $\mathbb{V}_\nu^\star$ are sets which contain the members of the νth column cluster, with $\nu \in [1, \mathfrak{v}] \subset \mathbb{Z}^+$. Each of the $\mathbb{V}_\nu^\star$ is constructed from the $\{\mathbb{V}_j, \mathbb{N}_j\}$ of its (the $\mathbb{V}_\nu^\star$) constituent elements via a process we will describe later.

Chapter 2

Disjoint Sets

When $\bar{\bar{\mathbf{L}}}$ consists of a series of disjoint sets, our clustering algorithm proceeds according to the following simple steps

1. Compute $\bar{\bar{\mathbf{A}}}_U$
2. Construct $\mathbb{U} = \Big\{ \{\mathbb{U}_1, \mathbb{M}_1\}, \{\mathbb{U}_2, \mathbb{M}_2\}, \ldots, \{\mathbb{U}_m, \mathbb{M}_m\} \Big\}$
3. Complete the $\big\{\mathbb{U}_i, \mathbb{M}_i\big\} \in \mathbb{U}$
4. Construct the $\mathbb{U}^\star$ set of sets
5. Compute $\bar{\bar{\mathbf{A}}}_V$
6. Construct $\mathbb{V} = \Big\{ \{\mathbb{V}_1, \mathbb{N}_1\}, \{\mathbb{V}_2, \mathbb{N}_2\}, \ldots, \{\mathbb{V}_n, \mathbb{N}_n\} \Big\}$
7. Complete the $\big\{\mathbb{V}_j, \mathbb{N}_j\big\} \in \mathbb{V}$
8. Construct the $\mathbb{V}^\star$ set of sets
9. Reorder $\bar{\bar{\mathbf{L}}}$ in the a new, temporary matrix, $\bar{\bar{\mathbf{L}}}_0 \in \mathbb{R}^{m \times n}$.
10. Reorder the temporary matrix, $\bar{\bar{\mathbf{L}}}_0$ into the final matrix $\bar{\bar{\mathbf{L}}}^\star \in \mathbb{R}^{m \times n}$.

The description of our algorithm for the case when $\bar{\bar{\mathbf{L}}}$ is composed of non-disjoint sets is given later in chapter 3.

2.1 Compute $\bar{\bar{\mathbf{A}}}_U$

We use the expression

$$\bar{\bar{\mathbf{A}}}_U = \bar{\bar{\mathbf{L}}}\,\bar{\bar{\mathbf{L}}}^{\intercal} \tag{1.4.1}$$

to compute $\bar{\bar{\mathbf{A}}}_U \in \mathbb{R}^{m \times m}$.

If a different weighting of the relations between the elements of $\bar{\bar{\mathbf{L}}}$ is desired, we can use the alternative expression

$$\left(\bar{\bar{\mathbf{A}}}_U\right)_{ij} = \sum_{k=1}^{n} \left\{ \delta^{\star}\left(\bar{\bar{\mathbf{L}}}, i, k\right)\, \bar{\bar{\mathbf{L}}}_{jk} \right\} \tag{1.4.1a}$$

where $i, j \in [1, m] \subset \mathbb{Z}^{+}$.

2.2 Construct $\mathbb{U}$

We initially construct $\mathbb{U}$ by following the sub-routine (sub algorithm) given below

Data: Connection Matrix for rows, $\bar{\bar{\mathbf{A}}}_U$, and the number of rows in $\bar{\bar{\mathbf{L}}}$, m.
Result: Initial value for each of the $\left\{\mathbb{U}_i, \mathbb{M}_i\right\} \in \mathbb{U}$.
begin
  for $i = 1 : m$ **do**
    $\mathbb{U}_i = \emptyset$; $\mathbb{M}_i = \emptyset$;
    for $j = 1 : m$ **do**
      if $\left(\bar{\bar{\mathbf{A}}}_U\right)_{ij} \neq 0$ **then**
        $\mathbb{U}_i = \mathbb{U}_i \bigcup \{j\}$;
        $\mathbb{M}_i = \mathbb{M}_i \bigcup \left\{ \left(\bar{\bar{\mathbf{A}}}_U\right)_{ij} \right\}$;
      end
    end
    $\mathbb{U}[i, 1] = \mathbb{U}_i$; $\mathbb{U}[i, 2] = \mathbb{M}_i$;
  end
end

Algorithm 1: Computing the initial value for each of the $\left\{\mathbb{U}_i, \mathbb{M}_i\right\} \in \mathbb{U}$.

After the completion of this sub-routine, we will make a copy of this initial state of $\mathbb{U}$. We denote this copy of this initial state as $\mathbb{U}_0$.

2.3 Complete the $\mathbb{U}$ Sets

The elements of $\mathbb{U}$ must now be "*completed*" so that they include any indirectly related elements. Taking advantage of the disjointedness of $\bar{\bar{\mathbf{L}}}$, the "*completion*" of the elements in $\mathbb{U}$ can be accomplished using the simple subroutine below

Data: Initial values for the elements of $\mathbb{U}$, $\{\mathbb{U}_i, \mathbb{M}_i\}$, and the number of rows in $\bar{\bar{\mathbf{L}}}$, m.
Result: The final values for each of the $\{\mathbb{U}_i, \mathbb{M}_i\} \in \mathbb{U}$.

```
begin
    boolean isChanged = true ;
    while isChanged do
        isChanged = false ;
        for  i = 1 : m do
            U_i = U[i,1]; M_i = U[i,2] ;
            for  j = 1 : m do
                U_j = U[j,1]; M_j = U[j,2] ;
                if  U_i ≠ U_j && U_i ⋂ U_j ≠ ∅ then
                    U_i = U_i ⋃ {U_j  U_i} ;
                    M_i = M_i ⋃ {M_j  M_i} ;
                    U_j = U_j ⋃ {U_i  U_j} ;
                    M_j = M_j ⋃ {M_i  M_j} ;
                    isChanged = true ;
                end
                if  isChanged then
                    U[j,1] = U_j; U[j,2] = M_j ;
                end
            end
            if  isChanged then
                U[i,1] = U_i; U[i,2] = M_i ;
            end
        end
    end
end
```

Algorithm 2: Computing the final value for each $\{\mathbb{U}_i, \mathbb{M}_i\} \in \mathbb{U}$.

2.4 Construct the $\mathbb{U}^{\star}$ Sets

To construct the sets of $\mathbb{U}^{\star}$, we look for unique $\mathbb{U}_i \in \mathbb{U}$ and then store each unique $\mathbb{U}_i$ as its own set in $\mathbb{U}^{\star}$. We accomplish this by using the following subroutine

```
Data: The final values for the elements of 𝕌, {𝕌_i, 𝕄_i}, and the number of rows in L̄̄, m.
Result: The sets of 𝕌*, with each set in 𝕌* representing a row cluster and containing its elements.
begin
    int nClust = 1; 𝕌* [nClust] = 𝕌 [1, 1] ;
    for i = 2 : m do
        boolean isNew = true; 𝕌_i = 𝕌 [i, 1] ;
        for j = 1 : nClust do
            𝕌*_j = 𝕌* [j] ;
            if 𝕌*_j = 𝕌_i then
                isNew = false ;
                break ;
            end
        end
        if isNew then
            nClust ++ ;
            𝕌* [nClust] = 𝕌_i ;
        end
    end
end
```

Algorithm 3: Compute the set of all row clusters sets, $\mathbb{U}^{\star}$.

2.5 Compute $\bar{\bar{\mathbf{A}}}_V$

We use the expression

$$\bar{\bar{\mathbf{A}}}_V = \bar{\bar{\mathbf{L}}}^{\intercal}\,\bar{\bar{\mathbf{L}}} \tag{1.4.2}$$

to compute $\bar{\bar{\mathbf{A}}}_V \in \mathbb{R}^{n \times n}$.

If a different weighting of the relations between the elements of $\bar{\bar{\mathbf{L}}}$ is desired, we can use the alternative expression

$$\left(\bar{\bar{\mathbf{A}}}_V\right)_{ij} = \sum_{k=1}^{m} \left\{ \delta^{\star}\left(\bar{\bar{\mathbf{L}}}, k, i\right) \bar{\bar{\mathbf{L}}}_{kj} \right\} \tag{1.4.2a}$$

where $i, j \in [1, n] \subset \mathbb{Z}^{+}$.

2.6 Construct $\mathbb{V}$

We initially construct $\mathbb{V}$ by following the sub-routine (sub algorithm) given below

Data: Connection Matrix for columns, $\bar{\bar{\mathbf{A}}}_V$, and the number of columns in $\bar{\bar{\mathbf{L}}}$, n.
Result: Initial value for each of the $\left\{\mathbb{V}_j, \mathbb{N}_j\right\} \in \mathbb{V}$.

```
begin
    for j = 1 : n do
        V_j = ∅; N_j = ∅ ;
        for i = 1 : n do
            if (A_V)_{ji} ≠ 0 then
                V_j = V_j ⋃ {i} ;
                N_j = N_j ⋃ { (A_V)_{ji} } ;
            end
        end
        V[j,1] = V_j; V[j,2] = N_j;
    end
end
```

Algorithm 4: Computing the initial value for each of the $\left\{\mathbb{V}_j, \mathbb{N}_j\right\} \in \mathbb{V}$.

After the completion of this sub-routine, we will make a copy of this initial state of $\mathbb{V}$. We denote this copy of this initial state as $\mathbb{V}_0$.

2.7 Complete the $\mathbb{V}$ Sets

The elements of $\mathbb{V}$ must now be "*completed*" so that they include any indirectly related elements. Taking advantage of the disjointedness of $\bar{\bar{\mathbf{L}}}$, the "*completion*" of the elements in $\mathbb{V}$ can be accomplished using the simple subroutine below

Data: Initial values for the elements of $\mathbb{V}$, $\{\mathbb{V}_j, \mathbb{N}_j\}$, and the number of columns in $\bar{\bar{\mathbf{L}}}$, n.
Result: The final values for each of the $\{\mathbb{V}_j, \mathbb{N}_j\} \in \mathbb{V}$.

```
begin
    boolean isChanged = true ;
    while isChanged do
        isChanged = false ;
        for j = 1 : n do
            V_j = V[j,1]; N_j = V[j,2] ;
            for i = 1 : n do
                V_i = V[i,1]; N_i = V[i,2] ;
                if V_j ≠ V_i && V_j ⋂ V_i ≠ ∅ then
                    V_j = V_j ⋃ {V_i V_j} ;
                    N_j = N_j ⋃ {N_i N_j} ;
                    V_i = V_i ⋃ {V_j V_i} ;
                    N_i = N_i ⋃ {N_j N_i} ;
                    isChanged = true ;
                end
                if isChanged then
                    V[i,1] = V_i; V[i,2] = N_i ;
                end
            end
            if isChanged then
                V[j,1] = V_j; V[j,2] = N_j ;
            end
        end
    end
end
```

Algorithm 5: Computing the final value for each $\{\mathbb{V}_j, \mathbb{N}_j\} \in \mathbb{V}$.

2.8 Construct the $\mathbb{V}^{\star}$ Sets

To construct the sets of $\mathbb{V}^{\star}$, we look for unique $\mathbb{V}_i \in \mathbb{V}$ and then store each unique $\mathbb{V}_i$ as its own set in $\mathbb{V}^{\star}$. We accomplish this by using the following subroutine

Data: The final values for the elements of $\mathbb{V}$, $\left\{\mathbb{V}_i, \mathbb{M}_i\right\}$, and the number of columns in $\bar{\bar{\mathbf{L}}}$, n.
Result: The sets of $\mathbb{V}^{\star}$, with each set in $\mathbb{V}^{\star}$ representing a column cluster and containing its elements.

```
begin
    int nClust = 1; V*[nClust] = V[1,1] ;
    for i = 2 : n do
        boolean isNew = true; V_i = V[i,1] ;
        for j = 1 : nClust do
            V*_j = V*[j] ;
            if V*_j = V_i then
                isNew = false ;
                break ;
            end
        end
        if isNew then
            nClust ++ ;
            V*[nClust] = V_i ;
        end
    end
end
```

Algorithm 6: Compute the set of all column clusters sets, $\mathbb{V}^{\star}$.