Sparse Derivative Matrix Determination with Pattern Graphs

Shahadat Hossain

1Department of Mathematics and Computer Science
University of Lethbridge, Canada

Euro AD Workshop 2006
(Aachen, Dec 07-08, 2006)

(Joint work with Trond Steihaug)

1
Outline

1. Introduction and Background
   - The Curtis, Powell, and Reid (CPR) Method
   - Orthogonal Partitioning

2. Graph Problems
   - Partitioning and Coloring
   - Coleman and Moré, 1983
   - Optimal Direct Methods

3. Symmetry + Sparsity and even more savings
   - Direct Determination
   - Bidirectional Determination

4. A Unified Approach
   - Sparsity Pattern Graph
   - Direct Methods via Sparsity Pattern Graph

5. Concluding Remarks

Shahadat Hossain
Given a nonlinear function

\[ F : \mathbb{R}^n \rightarrow \mathbb{R}^m \]

column \( j \) of the Jacobian matrix at \( x \)

\[ F_j' = \frac{\partial}{\partial x_j} F(x) \]

can be computed by AD or approximated by a forward difference:

\[ \frac{\partial}{\partial x_j} F(x) \approx \frac{1}{\varepsilon} \{ F(x + \varepsilon e_j) - F(x) \} \]
The sparsity pattern of the Jacobian matrix is known a priori and independent of the actual values of $x$. (Or can be computed as in AD)

If we need one or more components of $F$ at $x$ we need to compute the whole vector $F(x)$

- It is more efficient to evaluate the vector $F(x)$ than to evaluate each component of $F(x)$ separately: common sub-expressions are evaluated only once
- $F$ is a computer subroutine that returns the vector $F(x)$
Two columns are **structurally orthogonal** if they do not contain nonzeros in the same row position.

**Partition the columns into structurally orthogonal groups**
The Problem

\[
\frac{\partial F(x + ts)}{\partial t} \bigg|_{t=0} = F'(x)s \approx As = \frac{1}{\varepsilon} [F(x + \varepsilon s) - F(x)] \equiv b
\]

Forward difference (one extra function evaluation) gives \( As = b \) where \( b \) is the finite difference approximation.

Algorithmic Differentiation (AD) gives \( b = F'(x)s \).

**Formulation of the problem**
Obtain vectors \( s_1, \cdots, s_p \) such that the matrix vector product

\[
b_i \equiv As_i, \ i = 1, \cdots, p \text{ or } B \equiv AS
\]

determine the \( m \times n \) matrix \( A \) uniquely.
Main Steps in Computing \( A \) (Notations)

- \( A: m \times n \) matrix to be determined
- \( \rho_i: \) Number of nonzero entries in row \( i \) of \( A \)
- \( v_i: \) vector of column indices of nonzero entries in row \( i \) of \( A \)
- \( S: n \times p \) “seed” matrix

Obtain \( B = AS \) \((p \) matrix-vector products\).
Main Steps in Computing $A$ (Procedure)

Assume that $\rho_i \leq p$ for all $i$

1. **Seeding or Compression.** Obtain a suitable seed matrix $S \in \mathbb{R}^{n \times p}$ any square submatrix of which is *numerically well-conditioned* and *easy to solve*

2. **Harvesting or Reconstruction.** Determine the nonzero elements of $A$ row-by-row:
   a. Identify the reduced seed matrix $\hat{S}_i \in \mathbb{R}^{\rho_i \times p}$ for $A(i, v_i)$ (the $\rho_i$ nonzeros in row $i$ of $A$)

   $$\hat{S}_i = S(v_i, :)$$

   b. Solve for unknown elements $a_{ik} \neq 0$ of $A(i, :)$

   $$\hat{S}_i^T A(i, v_i)^T = B(i, :)^T$$
Harvesting row $i$ of $A$

$$\begin{pmatrix} A \\ \hat{S} \\ B \end{pmatrix} = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix}$$

Harvesting the unknown entries with column indices $k_1$, $k_2$, and $k_3$ in row $i$ of $A$

$$\hat{S}^T i A(i, v_i)^T = B(i, :)^T.$$
Let $\alpha = A(i, v_i)^T \in R^{p_i}$ and $\beta = B(i, :)^T \in R^p$. Then the unknown elements satisfy the overdetermined system ($\rho_i \leq p$)

$$\hat{S}^T \alpha = \beta$$

Let $\rho_i = p$. If $\hat{S}$

- is a permutation matrix then we have **direct determination**
- can be permuted to a triangular matrix then we have **determination by substitution**
- is a general nonsingular matrix we have **determination by elimination**.

**Fact**: The minimal number of matrix-vector products $p$ for any method is

$$\rho = \max_i \rho_i = p$$
Graph Concepts
$A \in \mathbb{R}^{m \times n}$, $G(A) = (V, E)$ $V = \{A(:, 1), \ldots, A(:, n)\}$
$E = \{\{A(:, i), A(:, j)\}: A(:, i) \not\perp_{S} A(:, j)\}$.

A **p-coloring** of the vertices of $G$ is a function $\phi: V \rightarrow \{1, 2, \ldots, p\}$ such that
$\{u, v\} \in E \Rightarrow \phi(u) \neq \phi(v)$. The **chromatic number** $\chi(G(A))$ is the smallest $p$ for which $G(A)$ has a p-coloring.

**Column partition**
A *partition* of the columns of $A$ is a division of columns into groups $C_1, C_2, \ldots, C_p$ such that each column belongs to one and only one group.

**Consistent partition**
A column partition where each group consists of *structurally orthogonal* columns is called a consistent (with direct determination) partition.
Coloring the Column Intersection Graph

Figure: A Sparse matrix and its column intersection graph.

Shahadat Hossain
Consistent partitioning and coloring (Coleman and Moré[1983])

- $\phi$ is a coloring of $G(A)$ if and only if $\phi$ induces a consistent partition of the columns of $A$.

- Coloring $G(A)$ is as hard as coloring a general graph.

- The CPR method is a greedy coloring method.

Consider vertices in their given order $v_1, \cdots, v_n$.

\[
\text{for } k = 1, \cdots, n \\
\quad \text{Assign vertex } v_k \text{ the smallest possible color}
\]

Ordering of the vertices affects the coloring.

- Numerical testing (DSM code) indicated $p$ close to $\rho = \max_i \rho_i$ (or largest identified clique).

- Two issues: Heuristic v.s. exact coloring (finding $\chi(G(A))$) and is this the best we can do (in terms of $p$)?
A Conjecture and a (Counter) Example

**Conjecture:** The chromatic number of the intersection graph is the minimal number of “function evaluations” (groups) needed to compute the Jacobian.

**Counter Example (Eisenstat)**

\[
A = \begin{pmatrix}
  a_{11} & 0 & 0 & a_{14} & 0 & 0 \\
  0 & a_{22} & 0 & 0 & a_{25} & 0 \\
  0 & 0 & a_{33} & 0 & 0 & a_{36} \\
  a_{41} & a_{42} & a_{43} & 0 & 0 & 0 \\
  a_{51} & 0 & 0 & 0 & a_{55} & a_{56} \\
  0 & a_{62} & 0 & a_{64} & 0 & a_{66} \\
  0 & 0 & a_{73} & a_{74} & a_{75} & 0
\end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}
\]

\(G(A)\) is complete so \(S = I\) and \(p = 6\).

**OR**

1. estimate the first 3 rows \((n/2)\) of \(A \equiv A_1\) using 2 matrix-vector products
2. estimate the last 4 \((n/2 + 1)\) rows of \(A \equiv A_2\) using 3 \(n\) matrix-vector products

yields \(p = 5 (n/2 + 2)\) matrix-vector products to completely determine \(A\) directly.
The Eisenstat Counter Example

Seedmatrix

\[ AS = \begin{pmatrix}
    a_{11} & 0 & 0 & a_{14} & 0 & 0 \\
    0 & a_{22} & 0 & 0 & a_{25} & 0 \\
    0 & 0 & a_{33} & 0 & 0 & a_{36} \\
    a_{41} & a_{42} & a_{43} & 0 & 0 & 0 \\
    a_{51} & 0 & 0 & 0 & a_{55} & a_{56} \\
    0 & a_{62} & 0 & a_{64} & 0 & a_{66} \\
    0 & 0 & a_{73} & a_{74} & a_{75} & 0
\end{pmatrix} \begin{pmatrix}
    1 & 0 & 1 & 0 & 0 \\
    1 & 0 & 0 & 1 & 0 \\
    1 & 0 & 0 & 0 & 1 \\
    0 & 1 & 1 & 0 & 0 \\
    0 & 1 & 0 & 1 & 0 \\
    0 & 1 & 0 & 0 & 1
\end{pmatrix} \]

Compression

\[ B = AS = \begin{pmatrix}
    a_{11} & a_{14} & a_{11} + a_{14} & 0 & 0 \\
    a_{22} & a_{25} & 0 & a_{22} + a_{25} & 0 \\
    a_{33} & a_{36} & 0 & 0 & a_{33} + a_{36} \\
    a_{41} + a_{42} + a_{43} & a_{45} + a_{56} & a_{51} & a_{42} & a_{43} \\
    a_{51} & a_{55} + a_{56} & a_{51} & a_{55} & a_{56} \\
    a_{62} & a_{64} + a_{66} & a_{64} & a_{62} & a_{66} \\
    a_{73} & a_{74} + a_{75} & a_{74} & a_{75} & a_{73}
\end{pmatrix} \]
The Eisenstat Counter Example (Reconstruction)

\[ \hat{S}^T \alpha = \beta \]

For example the nonzero elements in row 5 is determined in the reduced linear system

\[ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{pmatrix} \]

where

\[ \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{pmatrix} = \begin{pmatrix} a_{51} \\ a_{55} + a_{56} \\ a_{51} \\ a_{55} \\ a_{56} \end{pmatrix} \]

which gives

\[ \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} a_{51} \\ a_{55} \\ a_{56} \end{pmatrix}. \]
Column Segments Method (Hossain and Steihaug[1998,2003]

Let Π be a row partition of $A$ yielding $A(w_1, :), A(w_2, :), \ldots, A(w_q, :)$, where $w_i$ contains the row indices that constitute block $i$. $A(w_i, j), i = 1, \ldots, q$, and $j = 1, \ldots, n$ are called column segments.

**Structurally orthogonal column segments**

- (same column) $A(w_i, j)$ and $A(w_k, j)$ are structurally orthogonal.
- (same row block) $A(w_i, j)$ and $A(w_i, l)$ are structurally orthogonal if they do not have nonzero entries in the same row position.
- (different) $A(w_i, j)$ and $A(w_k, l)$, $i \neq k$ and $j \neq l$ are structurally orthogonal if
  - $A(w_i, j)$ and $A(w_i, l)$ are structurally orthogonal and
  - $A(w_k, j)$ and $A(w_k, l)$ are structurally orthogonal.
Matrix $A$ is partitioned into $q$ blocks.
Definition

Given matrix $A$ and row $q$-partition $\Pi$, the \textit{column-segment graph} associated with $A$ under partition $\Pi$ is a graph $G_\Pi(A) = (V, E)$ where the vertex $v_{ij} \in V$ corresponds to the column segment $A(w_i, j)$ not identically 0, and $\{v_{ij}, v_{kl}\} \in E$ if and only if column segments $A(w_{\tilde{i}}, j)$ and $A(w_{\tilde{k}}, l)$ are not structurally orthogonal.

Theorem

$\Phi$ is a coloring of $G_\Pi(A)$ if and only if $\Phi$ induces an orthogonal partition of the column segments of $A$. 
The element isolation graph (Newsam and Ramsdell, 1983) associated with \( A \in R^{m \times n} \) is denoted \( G_{I} = (V, E) \) where

\[
V = \{ a_{ij} \neq 0 : 1 \leq i \leq m, 1 \leq j \leq n \}
\]

and

\[
E = \{ \{ a_{ij}, a_{pq} \} \mid a_{ij} \text{ is not isolated from } a_{pq} \}.
\]

**Lemma**

Given a \( m \times n \) matrix \( A \) and a row \( m \)-partition \( \Pi_{m} \), \( \chi(G_{\Pi_{m}}(A)) = \chi(G_{I}(A)) \).

**Lemma**

\[
\chi(G_{\Pi_{m}}(A)) \leq \chi(G_{\Pi'(A)}) \leq \chi(G_{\Pi(A)}) \leq \chi(G_{\Pi_{1}}(A)) = \chi(G(A))
\]

where \( \Pi' \) is any refinement of \( \Pi \).
The minimal number of matrix-vector multiply in any direct determination method is $p = \chi(G_I(A))$.

**Proof.**

Consider the element isolation graph $\chi(G_I(A))$ of matrix $A$. The nonzero elements (column segments) that are directly determined in the product

$$ASE_k = Be_k, \ k = 1, 2, \ldots, q$$

are structurally orthogonal. Hence the vertices in $\chi(G_I(A))$ corresponding to those directly determined nonzero elements can be given color $k$. Since all the nonzero elements must be determined directly all vertices in $\chi(G_I(A))$ will receive a color. By Lemma 3 and 4 $p = \chi(G_I(A))$ is minimal.
The Eisenstat Example Revisited - Optimal Compression I

\[
\begin{pmatrix}
    a_{11} & 0 & 0 & a_{14} & 0 & 0 \\
    0 & a_{22} & 0 & 0 & a_{25} & 0 \\
    0 & 0 & a_{33} & 0 & 0 & a_{36} \\
    a_{41} & a_{42} & a_{43} & 0 & 0 & a_{45} \\
    a_{51} & 0 & 0 & 0 & a_{55} & a_{56} \\
    0 & a_{62} & 0 & a_{64} & 0 & a_{66} \\
    0 & 0 & a_{73} & a_{74} & a_{75} & 0
\end{pmatrix}
= 
\begin{pmatrix}
    0 & 0 & 1 & 1 \\
    0 & 1 & 0 & 1 \\
    1 & 0 & 0 & 1 \\
    1 & 1 & 0 & 0 \\
    1 & 0 & 1 & 0 \\
    0 & 1 & 1 & 0
\end{pmatrix}
\]

\[
\begin{pmatrix}
    a_{14} & a_{14} & a_{11} & a_{11} \\
    a_{25} & a_{22} & a_{25} & a_{22} \\
    a_{33} & a_{36} & a_{36} & a_{33} \\
    a_{43} & a_{42} & a_{41} & a_{43} + a_{41} \\
    a_{55} & a_{55} & a_{51} & a_{41} + a_{42} \\
    a_{64} & a_{62} + a_{64} + a_{66} & a_{66} & a_{62} \\
    a_{73} + a_{74} + a_{75} & a_{74} & a_{75} & a_{73}
\end{pmatrix}
\]
The Eisenstat Example Revisited - Reconstruction

\[
\begin{pmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 1 & 1 \\
1 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
\alpha_1 \\
\alpha_2 \\
\alpha_3
\end{pmatrix}
=
\begin{pmatrix}
\beta_1 \\
\beta_2 \\
\beta_3 \\
\beta_4
\end{pmatrix}
\]

\[
\begin{pmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 1 & 1 \\
1 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
a_{51} \\
a_{55} \\
a_{56}
\end{pmatrix}
=
\begin{pmatrix}
a_{55} \\
a_{56} \\
a_{51} + a_{55} + a_{56} \\
a_{51}
\end{pmatrix}
\]

We can reconstruct the matrix with \( p = 4 \) (v.s. 5).
Direct determination of a symmetric matrix (Hessian)
Let $G_s(A)$ be the adjacency graph associated with the symmetric matrix $A$. A mapping $\phi$ is a symmetric coloring for graph $G = (V, E)$ if following conditions are satisfied

1. $\{v_i, v_j\} \in E$ implies $\phi(v_i) \neq \phi(v_j)$.
2. For each path $v_i - v_j - v_k - v_l$ either $\phi(v_i) \neq \phi(v_k)$ or $\phi(v_j) \neq \phi(v_l)$.

**Theorem**

$\phi$ is a symmetric coloring for $G_s(A)$ if and only if $\phi$ induces a symmetrically consistent partition for the columns of $A$. 
Definition

A graph $G_{\sigma} = (V_{\sigma}, E_{\sigma})$ is called a symmetric completion of graph $G = (V, E)$ if $V_{\sigma} = V$ and

$$E_{\sigma} = E \cup \{e : e = \{v_i, v_k\} \text{ or } e = \{v_j, v_l\} \text{ for each path } v_i - v_j - v_k - v_l \in G\}$$

Theorem

*The symmetric chromatic number denoted* $\chi_{\sigma}$ *of graph* $G = (V, E)$ *is given by*

$$\chi_{\sigma}(G) = \min \{\chi(G_{\sigma}) : G_{\sigma} \text{ is a symmetric completion of } G\}$$

The symmetric $k$-coloring problem: Given a graph $G = (V, E)$ and an integer $k \geq 3$ does there exist a symmetric coloring for $G$ that uses only $k$ colors?

Theorem

*The symmetric k-coloring problem is NP-complete.*
Observations

- Element-wise sparsity (EI method) $\Rightarrow$ optimal (unidirectional) direct methods
- Symmetric matrix $\Rightarrow$ sparsity exploited in row and column directions automatically
- AD reverse mode: Given weight matrix $W^T \in \mathbb{R}^{pr \times m}$ one can compute $W^T A$ for the Jacobian matrix $A \in \mathbb{R}^{m \times n}$. Apply EI to $A$ and $A^T$ and obtain $S \in \mathbb{R}^{n \times pc}$ and $W^T \in \mathbb{R}^{pr \times m}$ and determine $A$ in $B = AS$ if $pc \leq pr$ or in $C^T = W^T A$, otherwise.

For matrices with few dense rows and columns neither forward nor reverse alone is effective
The Bipartition Problem

Given $A \in R^{m \times n}$ obtain vectors $w_1, \ldots, w_r \equiv W$ and $s_1, \ldots, s_p \equiv S$ such that for each $a_{ij} \neq 0$ there is an index $l_c$ such that $a_{ij} = b_{ilc}$ or that there is an index $l_r$ such that $a_{ij} = c_{jlr}$ where $W^T A = C^T$ and $AS = B$ and with $p = p_c + p_r$ minimized.
With $A \in R^{m \times n}$ we associate a bipartite graph $G_b(A) = (V_r \cup V_c, E)$ where with each row $i$ we associate a vertex $v_i \in V_r, i = 1, \ldots, m$ and with each column $j$ we associate a vertex $v_j \in V_c, j = 1, \ldots, n$ and the set of edges are defined as

$$E = \{\{v_i, v_j\} : a_{ij} \neq 0, v_i \in V_r, v_j \in V_c\}$$

W.L.O.G assume that the rows and columns are indexed as $1, \ldots, m, m+1, \ldots, m+n$ where first $m$ elements are row indices and the next $n$ elements are column indices.

**Definition**

A mapping $\phi_b : \{1, \ldots, m, m+1, \ldots, m+n\} \rightarrow \{1, \ldots, p\}$ is a path $p$-coloring of $G_b(A)$ if all of the following conditions are met:

1. $\phi_b(v_i) \neq \phi_b(v_j)$ for each $\{v_i, v_j\} \in E$.
2. $\{\phi_b(v_i) : v_i \in V_r\} \cap \{\phi_b(v_j) : v_j \in V_c\} = \emptyset$.
3. For each path $v_i - v_j - v_i' - v_j'$ either $\phi_b(v_i) \neq \phi_b(v_i')$ or $\phi_b(v_j) \neq \phi_b(v_j')$. 

Shahadat Hossain
Graph theoretic characterization of the partitioning problem is important. “Optimal” determination depends on the unraveling of the matrix sparsity at the “elemental” level.

Graph associated with sparse matrix keep changing
- Column intersection . . . (Column) Element Isolation
- Row intersection . . . (Row) Element Isolation
- Bipartite graph vs. Adjacency graph

Questions:
1. What is an optimal determination?
2. How are the different determination methods related?
Given \( A \in \mathbb{R}^{m \times n} \)
\( a_{i'j'} \neq 0 \) is a \textit{lateral neighbor} of \( a_{ij} \neq 0 \) if \( i = i' \) and, \( j' > j \) minimizes \( j' - j \) or \( j' < j \) minimizes \( j - j' \).
\( a_{i'j'} \neq 0 \) is a \textit{vertical neighbor} of \( a_{ij} \neq 0 \) if \( j = j' \) and, \( i' > i \) minimizes \( i' - i \) or \( i' < i \) minimizes \( i - i' \).

The \textit{sparsity-pattern graph} (or simply the pattern graph) associated with \( A \), \( G_P(A) = (V, E) \), where

\[
V = \{ v_{ij} : a_{ij} \neq 0, i = 1, 2, \ldots, m, j = 1, 2, \ldots, n \}
\]

and

\[
\{ v_{ij}, v_{i'j'} \} \in E \text{ if } a_{ij} \text{ and } a_{i'j'} \text{ are lateral or vertical neighbors.} \]
Figure: The Eisenstat (6 × 7) matrix and its associated sparsity-pattern graph
We have a path connecting vertices \( v_{ij} \neq v_{i'j'} \) denoted \( v_{ij} \sim^\ell v_{i'j'} \) if \( v_{ij} \equiv v_{i_0j_0}, v_{i_1j_1}, \ldots, v_{i_\ell j_\ell} \equiv v_{i'j'} \) is a sequence of vertices such that 
\[
\{v_{i_{k-1}j_{k-1}}, v_{i_kj_k}\} \in E, k = 1, 2, \ldots, \ell
\]
also \( v_{ij} \gtrsim^\ell v_{i'j'} \) denotes that the path in question is of length at least \( \ell \).

Let \( \Phi : V \mapsto \{1, 2, \ldots, p\} \) be a mapping such that for \( v_{ij} \in V \),
\[
v_{ij} \gtrsim^1 v_{ij'}, j \neq j' \text{ implies } \Phi(v_{ij}) \neq \Phi(v_{ij'})
\]
and
\[
v_{ij} \gtrsim^1 v_{ij'}, v_{i'j'} \gtrsim^1 v_{i'j''}, i \neq i', j \neq j' \text{ implies } \Phi(v_{ij}) \neq \Phi(v_{i'j''})
\]
Then the mapping \( \Phi \) is said to yield a column induced direct cover for the vertices of \( G_P(A) \).
Theorem

φ yields a column induced direct cover of $G_{\mathcal{P}}(A) = (V_{\mathcal{P}}, E_{\mathcal{P}})$ if and only if φ is a coloring of $G_{\mathcal{I}}(A) = (V_{\mathcal{I}}, E_{\mathcal{I}})$.

Corollary

A column induced direct cover for $G_{\mathcal{P}}(A)$ induces direct determination of the nonzero unknowns of A.
Let $\phi : V \leftrightarrow \{1, 2, \ldots, p\}$ be a mapping that satisfies the following:
for $v_{ij} \in V$,
\[ v_{ij} \geq 1 \sim v_{i'j}, i \neq i' \text{ implies } \phi(v_{ij}) \neq \phi(v_{i'j}), \]
and
\[ v_{ij} \geq 1 \sim v_{i'j} \geq 1 \sim v_{i''j'}, i \neq i', j \neq j' \text{ implies } \phi(v_{ij}) \neq \phi(v_{i'j'}). \]

Then the mapping $\phi$ is said to yield a row induced direct cover for the vertices of $G_P(A)$. 

**Corollary**

$\phi$ yields a row induced direct cover of $G_P(A)$ if and only if $\phi$ is a coloring of $G_I(A^T)$. 
**Figure:** EI coloring (column induced direct cover) of $G(A) (G_{PD}(A))$
Let $\Phi_\beta : V \mapsto \{1, 2, \ldots, p\}$ be a mapping that satisfies the following:

For each $v_{ij} \in V$

1. $v_{ij} \succeq^1 v_{ij'}, j \neq j'$ implies $\Phi_\beta(v_{ij}) \neq \Phi_\beta(v_{ij'})$

and

$v_{ij} \succeq^1 v_{ij'}, i \neq i', j \neq j'$ implies $\Phi_\beta(v_{ij}) \neq \Phi_\beta(v_{i'j'})$

OR

2. $v_{ij} \succeq^1 v_{i'j}, i \neq i'$ implies $\Phi_\beta(v_{ij}) \neq \Phi_\beta(v_{i'j})$

and

$v_{ij} \succeq^1 v_{i'j'}, i \neq i', j \neq j'$ implies $\Phi_\beta(v_{ij}) \neq \Phi_\beta(v_{i'j'})$

Then $\Phi_\beta$ yields a row-or-column induced direct cover for $A$. 

Shahadat Hossain
Let $S \in R^{n \times p_c}$ and $W \in R^{m \times p_r}$ be two seed matrices such that $AS = B$ and
$W^T A = C$ can be calculated. In a complete direct cover (Hossain and Steihaug, 1998) for $A$ for each $a_{ij} \neq 0$ there is a $l$ such that $a_{ij} = b_{il}$ or $a_{ij} = c_{lj}$.

**Theorem**

Φβ is a complete direct cover of A if and only if Φβ is a row-or-column induced direct cover for $G_\mathcal{P}(A)$. 
Figure: (a) EI coloring (column induced direct cover) of $G(A)$ ($G_P(A)$) (b) Complete Direct Cover (column-or-row induced direct cover) of $A$ ($G_P(A)$)
Exploiting known structural information in determining sparse derivative matrices leads to problems that are dependent on specific graph models.

Designing good heuristics require deep insight into the exploitable information.

Pattern graph provides a unifying tool express different sparse matrix determination problems in a transparent way while retaining a closer structural similarity to the underlying matrix.

Extend to Substitution Methods.

Need to explore the algorithmic implications for the problems under this model.