Semidefinite Programming	**Lecture 22**
OR 6327 Spring 2012	**April 12, 2012**
Scribe: Sin–Shuen Cheung	

Today's lecture is on symmetric matrix completions and associated algorithms.
Recall: C and all A_i's are all zero on entries outside $E \subset N \times N$, and we assume that $jj \in E$ for all j's.

Theorem 1 (Grone, Johnson, Sa and Wolkowicz) *If the partial symmetric matrix $\bar{X}$ has a positive definite completion, it has a unique one $\hat{X} = \hat{X}(\bar{X})$ solving*

$$\begin{aligned} \min \quad & -\ln\det X \\ & x_{ij} = \bar{x}_{ij}, \quad \forall ij \in E; \end{aligned}$$

characterized by

$$\hat{x}_{ij} = \bar{x}_{ij}, \quad \forall ij \in E, \qquad (\hat{X}^{-1})_{ij} = 0, \quad \forall ij \notin E.$$

Example 1

$$\bar{X} = \begin{bmatrix} 5 & ? & -4 \\ ? & 10 & 6 \\ -4 & 6 & 4 \end{bmatrix}$$

If we put ξ in the 12 and 21 positions, the determinant is $-140 - 48\xi - 4\xi^2$, which is maximized at $\xi = -6$. So the corresponding solution is

$$\hat{X} = \begin{bmatrix} 5 & -6 & -4 \\ -6 & 10 & 6 \\ -4 & 6 & 4 \end{bmatrix}$$

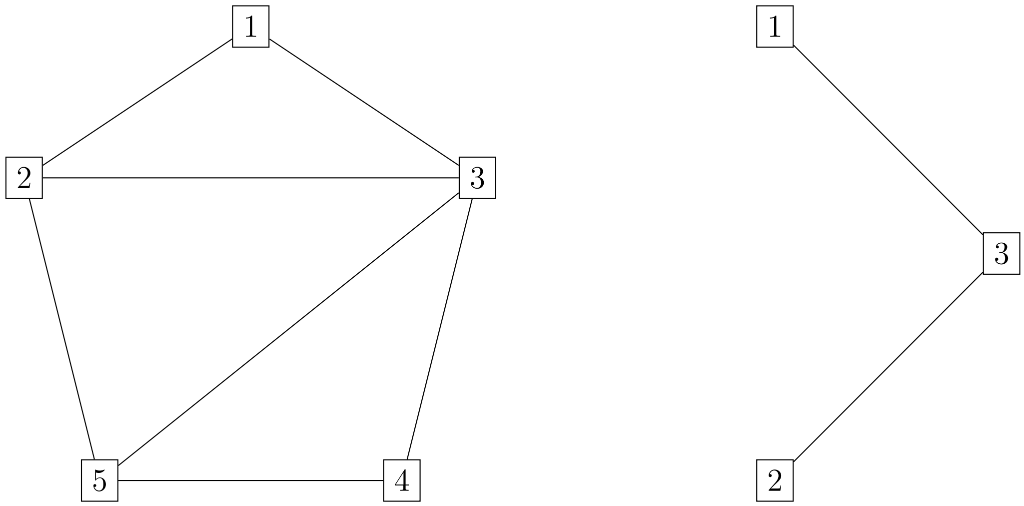
with

$$\hat{X}^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -\frac{3}{2} \\ 1 & -\frac{3}{2} & \frac{7}{2} \end{bmatrix}.$$

Let $E^\circ = E \backslash \{jj : j \in N\}$ and consider the graph $G = (N, E^\circ)$. We call G *chordal* if every circuit of G of length at least 4 has a chord. If G is chordal, there is an ordering of its nodes, called a perfect elimination (pe) order, such that if we delete the nodes in this order, the nodes adjacent to a node being deleted form a clique.

– If G is not chordal, we'll extend E to make it chordal.

– We'll assume that $1, \ldots, n$ is a pe order.



A chordal graph The chordal graph corresponding to our example

Figure 1: Chordal graphs

Suppose $Y \succ \mathbf{0}$ with sparsity pattern E. Then (N, E°) is chordal with pe order $1, \ldots, n$ *iff* Y has the Cholesky factorization LL^T where L has the same sparsity pattern E, i.e., there is no fill-in. In particular, if $\bar{X}$ is a partial symmetric matrix corresponding to E, then it has a completion $\hat{X}$ with $Y = \hat{X}^{-1}$ having such a sparse Cholesky factorization

$$LL^T = L_1 L_2 \ldots L_n L_n^T \ldots L_2^T L_1^T,$$

where L_j is the identity matrix except for its jth column, which is the jth column l_j of L and has $(l_j)_k \neq 0$ only for $k \geq j$, $kj \in E$. Let $M_j := L_j^{-1}$, with the same sparsity pattern. So,

$$\hat{X} = M_1^T M_2^T \ldots M_n^T M_n \ldots M_2 M_1.$$

Key fact: we can find all the M_j's from $\hat{X}$ and E.

Example 2 (Example 1 revisited)

$$\bar{X} = \begin{bmatrix} 5 & \xi & -4 \\ \xi & 10 & 6 \\ -4 & 6 & 4 \end{bmatrix}.$$

Consider the 33 *entry:*

$$4 = \bar{x}_{33} = \hat{x}_{33} = e_3^T M_1^T M_2^T M_3^T M_3 M_2 M_1 e_3 = e_3^T M_3^T M_3 e_3.$$

Since M_3 has the form of

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \mu \end{bmatrix},$$

we have that

$$M_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad L_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}.$$

Next, the 32 *entry:*

$$6 = \bar{x}_{32} = \hat{x}_{32} = e_3^T M_1^T M_2^T M_3^T M_3 M_2 M_1 e_2 = e_3^T M_3^T M_3 M_2 e_2 = 4 e_3^T M_2 e_2.$$

This gives

$$M_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \mu & 0 \\ 0 & \frac{3}{2} & 1 \end{bmatrix},$$

where μ is to be determined.
Then, the 22 *entry:*

$$\begin{aligned} 10 &= \bar{x}_{22} = \hat{x}_{22} = e_2^T M_1^T M_2^T M_3^T M_3 M_2 M_1 e_2 = e_2^T M_2^T M_3^T M_3 M_2 e_2 \\ &= (0, \mu, \tfrac{3}{2}) \begin{bmatrix} 1 & & \\ & 1 & \\ & & 4 \end{bmatrix} \begin{pmatrix} 0 \\ \mu \\ \frac{3}{2} \end{pmatrix} \\ &= \mu^2 + 9. \end{aligned}$$

It follows that $\mu = 1$. So,

$$M_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{3}{2} & 1 \end{bmatrix}.$$

Finally, look at the 21, 31 *and* 11 *positions. We don't know $\hat{x}_{21}$, but since* 21 *is* not *in E and we have a chordal graph, then $l_{21} = 0$. So $(L_1)_{21} = 0$. So, $(M_1)_{21} = 0$.*
For entry 31,

$$-4 = \bar{x}_{31} = \hat{x}_{31} = e_3^T M_1^T M_2^T M_3^T M_3 M_2 M_1 e_1 = 4 e_3^T M_2 M_1 e_1.$$

Hence, $e_3^T M_1 e_1 = -1$.
Continuing, we can compute $e_1^T M_1 e_1$ and hence we have M_1, M_2 and M_3.

An iteration of the matrix completion algorithm starts with the current partial iterate $(\bar{X}, y, S)$. Use the method above to compute $M_1, \ldots, M_n$. If $\bar{X} \bullet S$ (note that we can compute this since $s_{ij} = 0$ for $ij \notin E$) is no more than $\epsilon(\bar{X}_0 \bullet S_0)$, stop: we have an ϵ-optimal solution with $X = \hat{X}(\bar{X}) = M_1^T M_2^T \ldots M_n^T M_n \ldots M_2 M_1$. Otherwise, we need to compute the search direction. We use HKM. We compute

$$B_j = (M_1^T M_2^T \ldots M_n^T M_n \ldots M_2 M_1) A_j (L_S^{-T} L_S^{-1}),$$

where $L_S L_S^T$ is the (sparse) Cholesky factorization of S and then $A_i \bullet B_j$ for $i \geq j$.This gives us Δy and then we get ΔS, and finally $\Delta\bar{X}$, the entries in E of

$$\sigma\mu L_S^{-T} L_S^{-1} - \bar{X} - J((M_1^T M_2^T \ldots M_n^T M_n \ldots M_2 M_1)\Delta S(L_S^{-T} L_S^{-1})).$$

Finally, $\bar{X}_+ = \bar{X} + \alpha\Delta\bar{X}$, etc.

For long-step or adaptive algorithms, we need to search on α to ensure positive definiteness. For S, we only need the extreme eigenvalues of $L_S^{-1}\Delta S L_S^{-1}$, which is easy. For $\bar{X}$, it is not so easy. We want $\bar{X} + \alpha\Delta\bar{X}$ to have a positive definite completion.

Example 3 (Example 1 for the last time)

$$\bar{X} = \begin{bmatrix} 5 & ? & -4 \\ ? & 10 & 6 \\ -4 & 6 & 4 \end{bmatrix}$$

For $\bar{X}$ to have a positive definite completion, we clearly need $\bar{X}_{13,13}$ pd, $\bar{X}_{23,23}$ pd. It turns out that these conditions are sufficient.

In general, if G is chordal, it has only $O(n)$ maximal cliques, say $C_1, \ldots, C_l$.

Theorem 2 *$\bar{X}$ has a pd completion* iff *all $\bar{X}_{C_k,C_k}$'s are pd.*

Note that the matrix completion short-step path-following algorithm might not be polynomial, since $\hat{X}$ is updated at each iteration in an unconventional way. However, related interior-point methods are indeed polynomial, using potential reduction techniques. These methods decrease the potential function below by a constant at each iteration, where $\rho = n + \sqrt{n}$:

$$\phi_\rho(X, y, S) := \rho \ln X \bullet S + F(X) + F_*(S) = \underbrace{(\rho - n)\ln X \bullet S}_{\text{optimality}} + \underbrace{(n \ln X \bullet S + F(X) + F_*(S))}_{\text{close to central path}}.$$

Moving from "$\bar{X} + \alpha\Delta\bar{X}$" to $\hat{X}(\bar{X} + \alpha\Delta\bar{X})$ only decreases this potential function more.