$$\begin{bmatrix}X_1\\X_2\end{bmatrix} \Big| X_{rest} \sim N^{-1}\left(h, \begin{bmatrix}J_{11} & J_{12}\\ J_{21} & J_{22}\end{bmatrix} = \tilde{J}\right)$$

$$\sim N(\mu, \tilde{\Sigma} = \tilde{J}^{-1})$$

$$\tilde{\Sigma}_{12} = 0 \longleftrightarrow \tilde{J}_{12} = 0$$

* Gaussian Graphical Models.

Covariance Form

$$\begin{bmatrix}x_1\\x_2\end{bmatrix} \sim N\left(\begin{bmatrix}\mu_1\\\mu_2\end{bmatrix}, \begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{bmatrix}\right)$$

Information Form

$$\begin{bmatrix}x_1\\x_2\end{bmatrix} \sim N^{-1}\left(\begin{bmatrix}h_1\\h_2\end{bmatrix}, \begin{bmatrix}J_{11} & J_{12}\\ J_{21} & J_{22}\end{bmatrix}\right)$$

Marginal

$$\boxed{x_1 \sim N(\mu_1, \Sigma_{11})}$$

$$x_1 \sim N^{-1}\left(h_1 - J_{12}\boxed{J_{22}^{-1}}h_2,\ J_{11} - J_{12}\boxed{J_{22}^{-1}}J_{21}\right)$$

Independence

$$X_1 \perp X_2 \Longleftrightarrow \Sigma_{12} = 0$$

uncorrelation → Independence

$$X_i \perp\!\!\!\perp X_j | X_{rest} \Longleftrightarrow J_{ij} = 0$$

Conditional Indep.

Conditioning

$$x_1 | x_2 \sim N\left(\mu_1 + \Sigma_{12}\boxed{\Sigma_{22}^{-1}}(x_2 - \mu_2),\ \underbrace{\Sigma_{11} - \Sigma_{12}\boxed{\Sigma_{22}^{-1}}\Sigma_{21}}_{\text{schur complement}} = S\right)$$

$$\boxed{x_1 | x_2 \sim N^{-1}(h_1 - J_{12}x_2, J_{11})}$$

$$\exp\left(-\frac{1}{2}\begin{bmatrix}x_1-\mu_1\\x_2-\mu_2\end{bmatrix}^T \underbrace{\begin{bmatrix}(\Sigma_{11} - S)^{-1}, & -S^{-1}\Sigma_{12}\Sigma_{22}^{-1}\\ -\Sigma_{22}^{-1}\Sigma_{21}S^{-1}, & \Sigma_{22}^{-1} + \Sigma_{22}^{-1}\Sigma_{21}S^{-1}\Sigma_{12}\Sigma_{22}^{-1}\end{bmatrix}}_{\Sigma^{-1}} \begin{bmatrix}x_1-\mu_1\\x_2-\mu_2\end{bmatrix}\right)$$

$$\begin{bmatrix}X_1\\X_2\\X_3\end{bmatrix} \sim N\left(\begin{bmatrix}\Sigma_{11} & \Sigma_{12} & \Sigma_{13}\\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23}\\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33}\end{bmatrix}\right)$$

Intact in Marginalizing

Marg

$$\begin{bmatrix}X_1\\X_2\end{bmatrix} \sim N\left(\begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{bmatrix}\right)$$

Cond

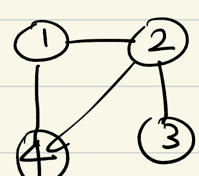
$$\begin{bmatrix}X_1\\X_2\end{bmatrix} \Big| X_3 \sim N(\text{Complex})$$

Def. Undirected Gaussian Graphical Models.

$$X \sim N^{-1}(h, J),\ P(x) = \frac{1}{Z}\prod_{(i,j)\in E} e^{-x_i^T J_{ij} x_j} \times \prod_{i\in V} e^{\left(-\frac{1}{2}x_i^T J_{ii} x_i + h_i^T x_i\right)}$$

$$E = \{(i,j) | J_{ij} \neq 0\}$$

$$J = \begin{bmatrix} & & 0 & \\ 0 & & & 0\\ & & 0 & \\ & & & \end{bmatrix}$$



G

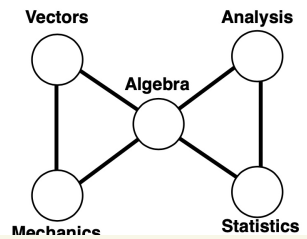
- Examination scores of 88 students in 5 subjects
- empirical information matrix (diagonal and above) covariance (below diagonal)

	Mechanics	Vectors	Algebra	Analysis	Statistics
Mechanics	5.24	-2.44	-2.74	0.01	-0.14
Vectors	0.33	10.43	-4.71	-0.79	-0.17
Algebra	0.23	0.28	26.95	-7.05	-4.70
Analysis	0.00	0.08	0.43	9.88	-2.02
Statistics	0.02	0.02	0.36	0.25	6.45

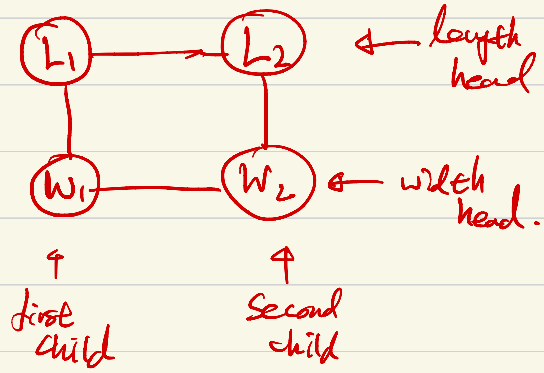
$J \Uparrow$

$\Sigma \Leftarrow$

G

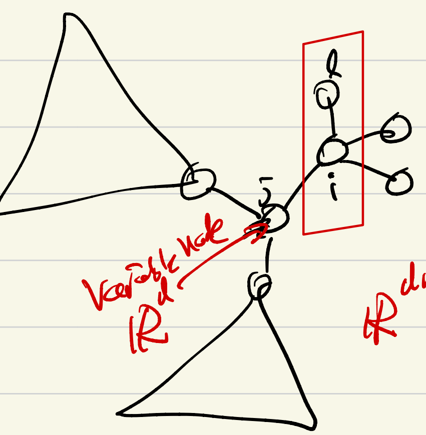


$L_1, L_2, w_1, w_2,$

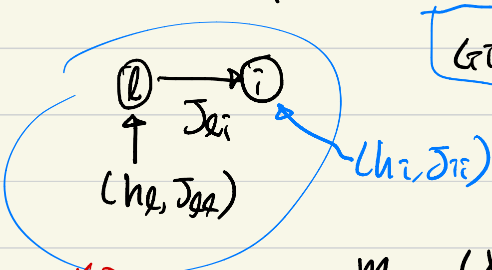


* Gaussian Belief Propagation.

in the same way as we derived BP for discrete G.M.



$\mathbb{R}^{dn} = X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$



Given: $\begin{bmatrix} x_l \\ x_i \end{bmatrix} \sim N^{-1}\left(\begin{bmatrix} h_l \\ 0 \end{bmatrix}, \begin{bmatrix} J_{ll} & J_{li} \\ J_{il} & 0 \end{bmatrix} \right)$, $x_l \in \mathbb{R}^d$, $\mathbb{R}^{d\times d}$

$$m_{l\to i}(x_i) \propto \frac{1}{Z} \exp\left(-\frac{1}{2} x_i^T J_{l\to i} x_i + h_{l\to i}^T x_i\right)$$

$\trianglerighteq (h_{l\to i}, J_{l\to i})$

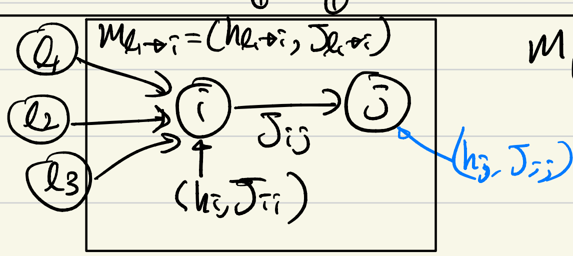
marginal $x_i \sim N^{-1}(h_{l\to i}, J_{l\to i})$, $h_{l\to i} \in \mathbb{R}^d$

$h_{l\to i} = -J_{il} J_{ll}^{-1} h_l$ $\qquad$ $J_{l\to i} = -J_{il} J_{ll}^{-1} J_{li}$

message: $\mathbb{R}^d \times \mathbb{R}^{d\times d}$

$2|E|$ messages

$\mathbb{R}^d \quad \mathbb{R}^{d\times d}$



$m_{i\to j}(x_j) \sim N^{-1}(h_{i\to j}, J_{i\to j})$

$$h_{i\to j} = -J_{ji}\left(J_{ii} + \sum_{l\in\partial i\setminus\{j\}} J_{l\to i}\right)^{-1}\left(h_i + \sum_{l\in\partial i\setminus j} h_{l\to i}\right)$$

$$J_{i\to j} = -J_{ji}\left(J_{ii} + \sum_{l\in\partial i\setminus j} J_{l\to i}\right)^{-1} J_{ij}$$

Decision/Marginal

$\hat{h}_i = h_i + \sum_{\ell \in \partial i} \hat{u}_{\ell \to i}$ $\qquad \hat{x}_i \sim N^{-1}(\hat{h}_i, \hat{J}_{ii})$

$\hat{J}_i = J_{ii} + \sum_{\ell \in \partial i} J_{\ell \to i}$ $\qquad N(\hat{J}_i^{-1} \hat{h}_i, \hat{J}_{ii}^{-1})$

$O(d^3)$

[T steps of BP takes $O(d^3 \cdot |E| \cdot T)$ ← $O(n)$, $\log(n)$

inverting $J \in \mathbb{R}^{dn \times dn}$ $\quad O(d^3 \cdot n^3)$

---

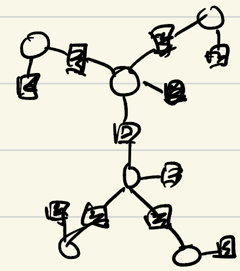
* Alternative version of GBP.

$$h_{i \to j} = h_i - \sum_{\ell \in \partial i \setminus j} J_{i\ell} \cdot J_{\ell \to i}^{-1} h_{\ell \to i}$$

$$J_{i \to j} = J_{ii} - \sum_{\ell \in \partial i \setminus j} J_{i\ell} J_{\ell \to i}^{-1} J_{\ell i}$$

$$\hat{h}_i = h_i - \sum_{\ell \in \partial i} J_{i\ell} J_{\ell \to i}^{-1} h_{\ell \to i}$$

$$\hat{J}_i = J_i - \sum_{\ell \in \partial i} J_{i\ell} J_{\ell \to i}^{-1} J_{\ell i}$$



$\Sigma \Rightarrow \int$

---

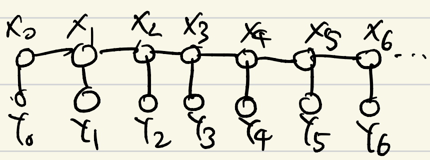
* Maximization is equivalent $(\hat{\mu}_1, \dots, \hat{\mu}_n) = \hat{\mu}$

* Q. $J = [\quad]$ How do we know $J \succ 0$

① If you have to check exactly $O((dn)^3)$

② $\exists$ sufficient conditions that imply $J \succ 0$

Gaussian Hidden Markov Models. = linear dynamical systems (Kalman Filtering)



state $X_i \in \mathbb{R}^d$

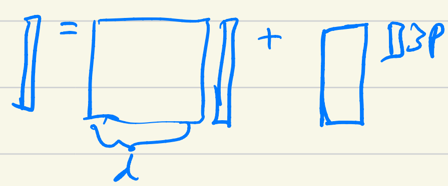
state transition matrix: $A \in \mathbb{R}^{d \times d}$

process noise [$V_t \in \mathbb{R}^p \sim N(0, V)$

$B \in \mathbb{R}^{d \times p}$

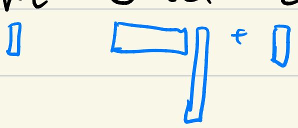
$$X_0 \sim N(0, \Sigma_0)$$

$$X_{t+1} = AX_t + BV_t$$



observation: $Y_t \in \mathbb{R}^{d'}$, noise $W_t \sim N(0, W)$.

$$Y_t = CX_t + W_t.$$



$$X_0 \sim N(0, \Sigma_0)$$
G.GM
$$X_{t+1} | X_t \sim N(AX_t, H = BVB^T)$$
$$Y_t | X_t \sim N(CX_t, W)$$

factorization:

$$P(X) = \frac{1}{Z} \underbrace{\exp(-\tfrac{1}{2} X_0^T \Sigma_0^{-1} X_0)}_{P(X_0)} \underbrace{\exp\left(-\tfrac{1}{2}(X_1 - AX_0)^T H^{-1} (X_1 - AX_0)\right)}_{P(X_1|X_0)} \cdot \exp \cdots$$
$$\underbrace{\exp\left(-\tfrac{1}{2}(Y_0 - CX_0)^T W^{-1} (Y_0 - CX_0)\right)}_{P_{Y_0|X_0}} \cdot \exp(\cdots$$

Information form:
$$= \frac{1}{Z} \prod_{i=0}^{n} \exp\left(-\tfrac{1}{2} X_i^T J_i X_i + h_i^T X_i\right) \cdot \prod_{i=0}^{n-1} \exp(-X_i^T J_{i,i+1} X_{i+1})$$

$h_i^T = Y_i^T W^{-1} C$, $\quad J_{i,i+1} = -A^T H^{-1}$

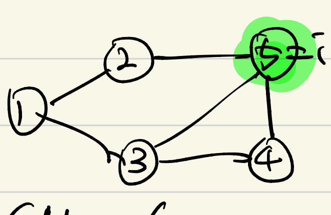
$$J_i = \begin{cases} \Sigma_0^{-1} + C^T W^{-1} C + A^T H^{-1} A, & i = 0 \\ H^{-1} + C^T W^{-1} C + A^T H^{-1} A, & 0 < i < n \\ H^{-1} + C^T W^{-1} C, & i = n \end{cases}$$

---

* Claim: If GBP converges (on a loopy graph)
then expectation is correct.
Formally, let $\hat{\mu}_i^{(l)} \triangleq (\hat{J}_i^{(l)})^{-1} \cdot \hat{h}_i^{(l)}$ ← l-steps GBP.

If $\hat{\mu}_i^{(\infty)} \triangleq \lim_{l \to \infty} \hat{\mu}_i^{(l)}$ exists then
$$\hat{\mu}_i^{\infty} = \mu_i$$
Truth.

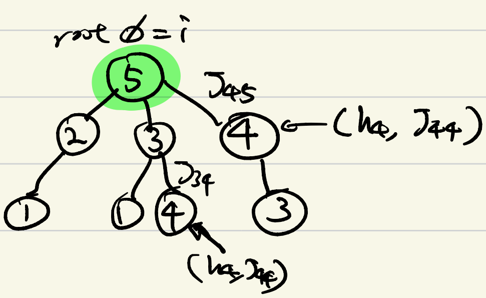
Def. Computation Tree $CT_G(i; l)$ is the tree of l-steps of non-reversing walks on G starting at node i



GM. G

$\Longrightarrow$

$C.T_G(i, l)$



$$\hat{\mu}_i^{(l)}(G) = \hat{\mu}_\phi^{(\infty)}(CT_G(i;l))$$

on G. Given (J, h)

$J \cdot \mu = h$. $\leftarrow$ we define μ

$\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix}$ satisfy $\begin{cases} J_i \mu_i + \sum_{l \in \partial i} J_{il} \mu_l = h_i \\ \vdots \end{cases}$

$\{\hat{\mu}_{\phi i}^{(l)}\}$ Satisfy

if Convergence.