

# Graph-Valued Regression

**Presenter: Xi Chen**

**Han Liu, Xi Chen, John Lafferty and Larry Wasserman**

**Machine Learning Department**

**Carnegie Mellon University**

# Graph-Valued Regression

**Multivariate Regression**  
(supervised)

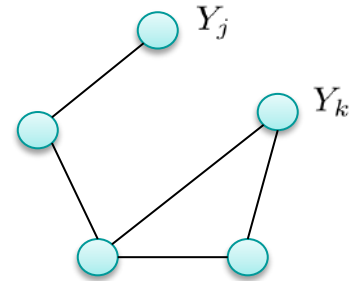
$$\mu(x) = E(Y|X = x)$$

$$Y \in \mathbb{R}^p \quad x \in \mathcal{X} = [0, 1]^d$$

**Undirected Graphical Model**  
(unsupervised)

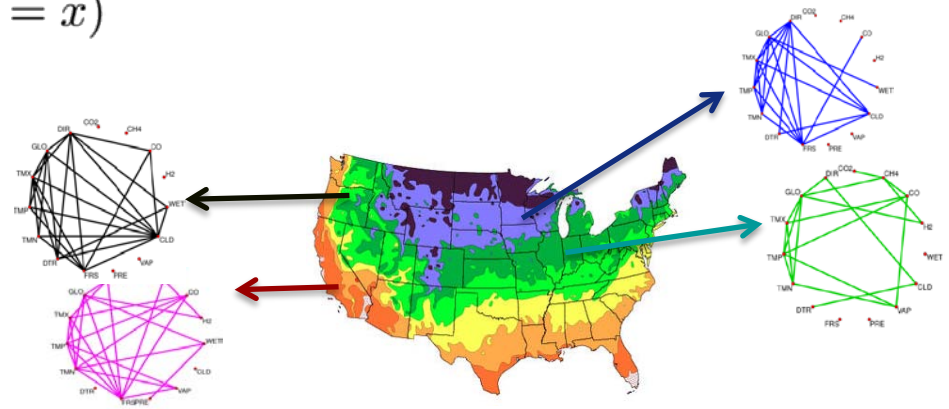
$$\text{Graph}(Y) = (V, E)$$

$$(j, k) \notin E \Leftrightarrow Y_j \perp\!\!\!\perp Y_k \mid \text{rest}$$



**Graph-Valued Regression:**  $\text{Graph}(Y|X = x)$

**Application: complex data analysis**



**Learn  $\text{Graph}(Y)$  via Graphical Lasso**

$$Y \sim N_p(\mu, \Sigma) \quad \Omega = \Sigma^{-1}$$

$$\Omega_{jk} = 0 \Leftrightarrow Y_j \perp\!\!\!\perp Y_k \mid \text{rest} \Leftrightarrow (j, k) \notin E$$

[Yuan & Lin 07]

$$\hat{\Omega} = \arg \min_{\Omega \succ 0} \left\{ \underbrace{\text{tr}(S\Omega) - \log |\Omega|}_{\text{- Log-likelihood}} + \lambda \underbrace{\|\Omega\|_1}_{\text{Sparsity}} \right\}$$

- Log-likelihood      Sparsity

S: Sample Covariance Matrix

# Go-CART (Graph Optimized CART)

## Co-CART (Graph-Optimized CART):

$$Y | X = x \sim N_p(\mu(x), \Sigma(x))$$

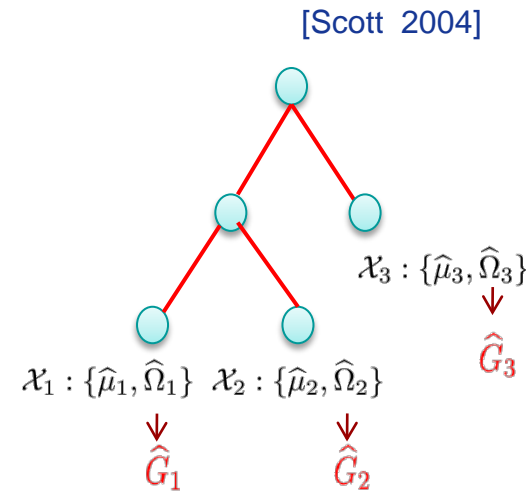
## Partition based estimator:

**Input data:**  $\{(x_1, y_1), \dots, (x_n, y_n)\}$  where  $x_i \in \mathcal{X} = [0, 1]^d$  and  $y_i \in \mathbb{R}^p$

1. **Dyadic Decision Tree based Partition:**  $\mathcal{X} = \bigcup_{j=1}^m \mathcal{X}_j$

2. **Estimate Graphs:**  $\hat{G}_j$  via *Graphical Lasso* based on  $\{y_i : x_i \in \mathcal{X}_j\}$

$$G(Y|X = x) = \hat{G}_j, \quad \forall x \in \mathcal{X}_j$$



## Co-CART Estimators:

1. Penalized Empirical Risk Minimization Estimator: Minimize **negative conditional log-likelihood** + **tree-induced penalty**
2. Held-out Risk Minimization Estimator: Minimize **held-out negative conditional log-likelihood**

## Statistical Property:

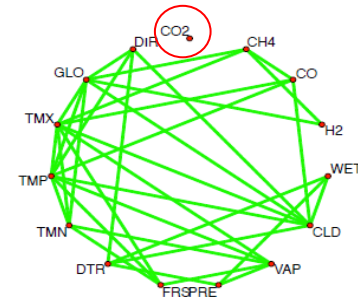
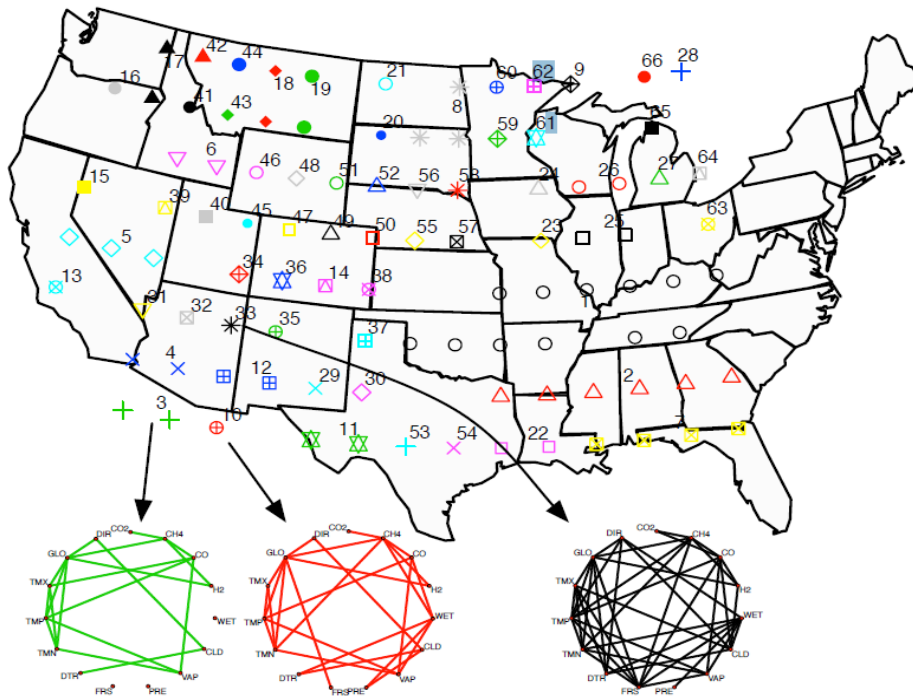
- Oracle inequality
- Tree partition consistency :  $\mathbb{P}(\Pi(T^*) \subset \Pi(\hat{T})) \rightarrow 1$

# Experimental Results

## Real Climate Data Analysis

held-out risk minimization with a greedy tree learning algorithm

100 locations of U.S.; Monthly observation 1990 ~ 2002 (training: 1996~2002); 15 factors (greenhouse gases, e.g. CO<sub>2</sub>; temperature, climate, solar radiation)



Graphical Lasso



## Observations:

- (1): Graphical lasso, no edge connect CO<sub>2</sub> with solar radiation factors; Co-CART, there is.
- (2): Graphs along the coasts are more sparse than the ones in the mainland.