Detection of Gauss-Markov Random Field on Nearest-Neighbor Graph

A. Anandkumar[1] **L. Tong**[1] **A. Swami**[2]

[1] School of Electrical and Computer Engineering
Cornell University, Ithaca, NY 14853

[2] Army Research Laboratory, Adelphi MD 20783

2007 International Conference on Acoustics, Speech and Signal Processing

Supported by the Army Research Laboratory CTA

Introduction: Distributed Detection

Setup

- Sensors: transmit local decisions
- Fusion center: Global Decision
- Classical data model: Conditionally IID

Sensor signal field

- Correlated sensor readings
- Large coverage area
- Large number of sensors
- Arbitrary sensor placement

Influence of correlation structure on detection performance

Detection of Correlation

Binary hypothesis testing

$\mathcal{H}_1$: Correlated data vs. $\mathcal{H}_0$: Independent observations

Questions

- How to model correlation?
- Is there an analytically tractable performance metric?
- How does correlation affect performance?
- How does node density affect performance?

New tradeoffs not encountered in IID scenario

Detection of Correlation

Binary hypothesis testing

$\mathcal{H}_1$: Correlated data vs. $\mathcal{H}_0$: Independent observations

Questions

- How to model correlation?
- Is there an analytically tractable performance metric?
- How does correlation affect performance?
- How does node density affect performance?

New tradeoffs not encountered in IID scenario

Detection of Correlation

Binary hypothesis testing

$\mathcal{H}_1$: Correlated data vs. $\mathcal{H}_0$: Independent observations

Questions

- How to model correlation?
- Is there an analytically tractable performance metric?
- How does correlation affect performance?
- How does node density affect performance?

New tradeoffs not encountered in IID scenario

Summary of Results

Questions Answered

- How to model correlation?
  - Gauss-Markov random field
- Is there an analytically tractable performance metric?
  - Closed-form detection error exponent for Neyman Pearson
- How does correlation affect performance?
  - Depends on variance ratio
    - If signal under $\mathcal{H}_1$ is weak (low variance), correlation helps
    - If signal under $\mathcal{H}_1$ is strong (high variance), correlation hurts
- How does node density affect performance?
  - More node density more correlation as edge length is reduced

Summary of Results

Questions Answered

- How to model correlation?
  - Gauss-Markov random field
- Is there an analytically tractable performance metric?
  - Closed-form detection error exponent for Neyman Pearson
- How does correlation affect performance?
  - Depends on variance ratio
    - If signal under $\mathcal{H}_1$ is weak (low variance), correlation helps
    - If signal under $\mathcal{H}_1$ is strong (high variance), correlation hurts
- How does node density affect performance?
  - More node density more correlation as edge length is reduced

Summary of Results

Questions Answered

- How to model correlation?
  - Gauss-Markov random field
- Is there an analytically tractable performance metric?
  - Closed-form detection error exponent for Neyman Pearson
- How does correlation affect performance?
  - Depends on variance ratio
    - If signal under $\mathcal{H}_1$ is weak (low variance), correlation helps
    - If signal under $\mathcal{H}_1$ is strong (high variance), correlation hurts
- How does node density affect performance?
  - More node density more correlation as edge length is reduced

Summary of Results

Questions Answered

- How to model correlation?
  - Gauss-Markov random field
- Is there an analytically tractable performance metric?
  - Closed-form detection error exponent for Neyman Pearson
- How does correlation affect performance?
  - Depends on variance ratio
    - If signal under $\mathcal{H}_1$ is weak (low variance), correlation helps
    - If signal under $\mathcal{H}_1$ is strong (high variance), correlation hurts
- How does node density affect performance?
  - More node density more correlation as edge length is reduced

Previous Results on Detection Error Exponent

I.I.D case

- Closed-form for optimal detector and threshold
- Error exponent - Stein's lemma

Correlated case

- Stationary Gaussian process (Donsker & Varadhan, 85)
- General formulas for Neyman-Pearson exponent (Chen, 96)
- Closed-form for Gauss-Markov random process (Sung & etal, 06)

Limitations of the closed form

- Requires causality, valid in 1-D case
- Cannot handle random placement of nodes

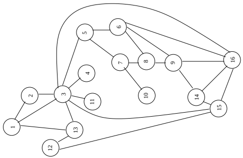
Outline

Outline

Model for Correlated Data : Graphical Model

$X_{i-1} \perp X_{i+1} | X_i$

Linear graph corresponding to autoregressive process of order 1



Graph of German states and states with common borders are neighbors

Temporal signals

- Conditional independence based on ordering
- Fixed number of neighbors
- Causal (random processes)

Spatial signals

- Conditional independence based on (undirected) Dependency Graph
- Variable set of neighbors
- Maybe acausal

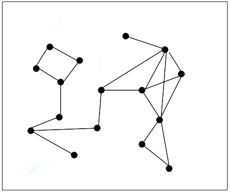
Remark

Dependency graph is NOT related to communication capabilities, but to the correlation structure of data!

Markov Random Field

Definition : MRF with Dependency Graph $\mathcal{G}_d(\mathcal{V},\mathcal{E})$

$\mathbf{Y}(\mathcal{V}) = \{Y_i : i \in \mathcal{V}\}$ is MRF with $\mathcal{G}_d(\mathcal{V},\mathcal{E})$ if $\mathbf{Y}$ is Gaussian random field, PDF satisfies positivity condition and Markov property



Markov Property

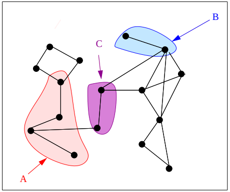
- A, B, C are disjoint
- A, B non-empty
- C separates A, B

$$\mathbf{Y}_A \perp \mathbf{Y}_B | \mathbf{Y}_C$$

Markov Random Field

Definition : MRF with Dependency Graph $\mathcal{G}_d(\mathcal{V}, \mathcal{E})$

$\mathbf{Y}(\mathcal{V}) = \{Y_i : i \in \mathcal{V}\}$ is MRF with $\mathcal{G}_d(\mathcal{V}, \mathcal{E})$ if $\mathbf{Y}$ is Gaussian random field, PDF satisfies positivity condition and Markov property



Markov Property

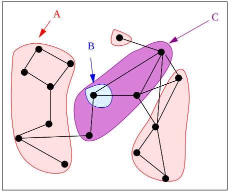
- A, B, C are disjoint
- A, B non-empty
- C separates A, B

$$\mathbf{Y}_A \perp \mathbf{Y}_B | \mathbf{Y}_C$$

Markov Random Field

Definition : MRF with Dependency Graph $\mathcal{G}_d(\mathcal{V}, \mathcal{E})$

$\mathbf{Y}(\mathcal{V}) = \{Y_i : i \in \mathcal{V}\}$ is MRF with $\mathcal{G}_d(\mathcal{V}, \mathcal{E})$ if $\mathbf{Y}$ is Gaussian random field, PDF satisfies positivity condition and Markov property



Markov Property

- A, B, C are disjoint
- A, B non-empty
- C separates A, B

$$\mathbf{Y}_A \perp \mathbf{Y}_B | \mathbf{Y}_C$$

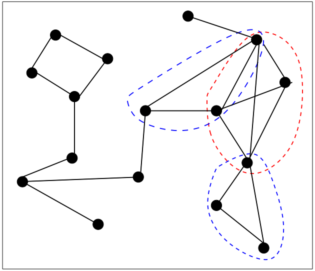
Likelihood Function of MRF

Hammersley-Clifford Theorem (1971)

For a MRF $\mathbf{Y}$ with dependency graph $\mathcal{G}_d(\mathcal{V}, \mathcal{E}_d)$,

$$\log \mathbb{P}(\mathbf{Y}; \mathcal{G}_d) = Z + \sum_{c \in \mathcal{C}} \Psi_c(\mathbf{Y}_c), \;\; Z \stackrel{\Delta}{=} e^{-\int\limits_{\mathbf{Y}} \prod\limits_{c \in \mathcal{C}} \Psi_c(\mathbf{Y}_c)},$$

where $\mathcal{C}$ is the set of all cliques in $\mathcal{G}_d$ and Ψ_C the clique potential

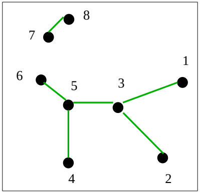


Dependency Graph

Potential Matrix of GMRF

Potential Matrix

- Inverse of covariance matrix of a GMRF
- Non-zero elements of Potential matrix correspond to graph edges



Dependency Graph

$$\begin{bmatrix} \times & & \times & & & & & \\ & \times & \times & & & & & \\ \times & \times & \times & & \times & & & \\ & & & \times & \times & & & \\ & & \times & \times & \times & \times & & \\ & & & & \times & \times & & \\ & & & & & & \times & \times \\ & & & & & & \times & \times \end{bmatrix}$$

$\times$: Non-zero element of Potential Matrix

Form of Log-Likelihood of zero-mean GMRF with potential matrix $\mathbf{A}$

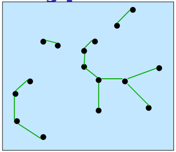
$$-\log P(\mathbf{Y}_n; \mathcal{G}_d, \mathbf{A}) = \frac{1}{2}\Big(-n\log 2\pi + \log|\mathbf{A}| + \sum_{(i,j)\in\mathcal{E}_d} A(i,j)Y_iY_j + \sum_{i\in\mathcal{V}} A(i,i)Y_i^2\Big)$$

Acyclic Dependency Graph

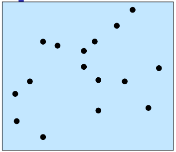
Given Covariance matrix, closed-form expression of likelihood

Outline

Hypothesis Testing for Independence



$\mathcal{H}_1$: GMRF with dependency graph $\mathcal{G}_d$



$\mathcal{H}_0$: Independent observations

Model for Dependency Graph $\mathcal{G}_d$ under $\mathcal{H}_1$

- Dependency graph is a proximity graph (edges between nearby points)
- Simplest proximity graph: nearest-neighbor graph

Definition of Nearest-Neighbor Graph

In NNG, (i, j) is an edge if i is nearest neighbor of j or vice versa

Additional assumptions

- Random placement of nodes (Uniform or Poisson distribution)
- Correlation function g : function of spatial distance

Optimal Detection

Log Likelihood Ratio (LLR) Detector

$$\log \frac{P[\mathbf{Y}_n, \mathcal{V}; \mathcal{H}_1]}{P[\mathbf{Y}_n, \mathcal{V}; \mathcal{H}_0]} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \tau_n$$

Neyman-Pearson Detection

Minimize Miss Probability

$$P_M \stackrel{\Delta}{=} P[\text{Decision} = \mathcal{H}_0 | \mathcal{H}_1]$$

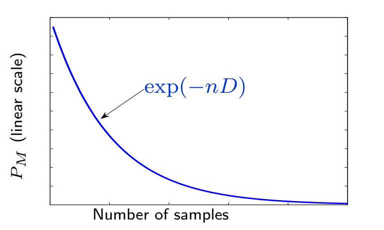
with false alarm constraint

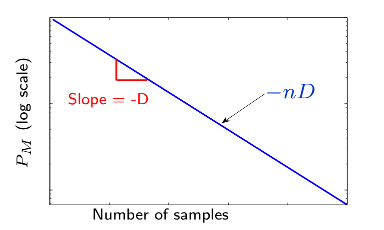
$$P_F = P[\text{Decision} = \mathcal{H}_1 | \mathcal{H}_0] \leq \alpha$$

Outline

Error Exponent D

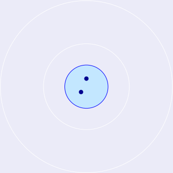
Closed-form of error probability not tractable





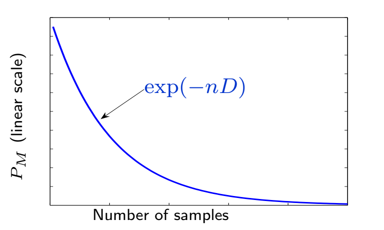
$$P_M \approx e^{-nD} \qquad \log P_M \approx -nD$$

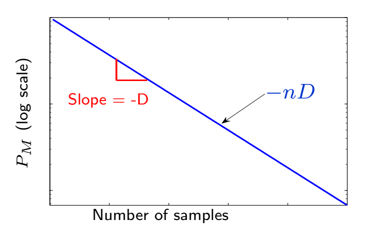
Sensors Placed in region with constant node density λ



Error Exponent D

Closed-form of error probability not tractable

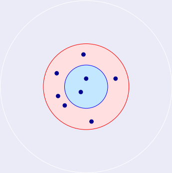




$$P_M \approx e^{-nD}$$

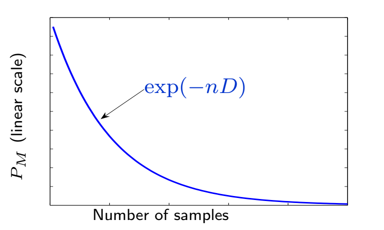
$$\log P_M \approx -nD$$

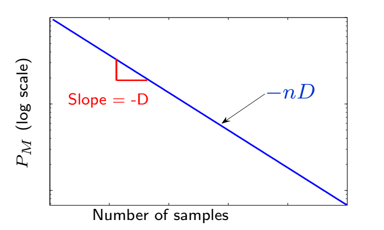
Sensors Placed in region with constant node density λ



Error Exponent D

Closed-form of error probability not tractable

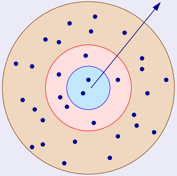




$$P_M \approx e^{-nD}$$

$$\log P_M \approx -nD$$

Sensors Placed in region with constant node density λ



Our Methodology

Approaches

- LLR as sum of node and edge functionals of dependency graph
- Error exponent through limit of LLR
- Evaluate limit using Law of Large Numbers for graph functionals
- Error exponent for performance analysis

Detailed Methodology

LLR as sum of node and edge functionals of dependency graph

$$\mathsf{LLR}(\mathbf{Y}_n, \mathcal{G}_d) = n \log \frac{\sigma_1}{\sigma_0} + \frac{1}{2}\Big[\sum_{i \in \mathcal{V}} \Big(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2}\Big) Y_i^2$$

$$+ \sum_{\substack{(i,j) \in \mathcal{E}_d \\ i<j}} \Big\{ \log[1 - g^2(R_{ij})] + \frac{g^2(R_{ij})}{1 - g^2(R_{ij})} \frac{Y_i^2 + Y_j^2}{\sigma_1^2} - \frac{2g(R_{ij})}{1 - g^2(R_{ij})} \frac{Y_i Y_j}{\sigma_1^2} \Big\} \Big]$$

Error exponent through limit of LLR

$$D = \lim_{n \to \infty} \frac{1}{n} \mathsf{LLR}(\mathbf{Y}_n; \mathcal{G}_d), \quad \mathcal{H}_0$$

LLR is sum of graph functionals of a Marked process

Y_i are independent under $\mathcal{H}_0$

Detailed Methodology

LLR as sum of node and edge functionals of dependency graph

$$\mathsf{LLR}(\mathbf{Y}_n, \mathcal{G}_d) = n \log \frac{\sigma_1}{\sigma_0} + \frac{1}{2}\Bigg[\sum_{i \in \mathcal{V}} \left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2}\right) Y_i^2$$

$$+ \sum_{\substack{(i,j)\in\mathcal{E}_d \\ i<j}} \left\{ \log[1 - g^2(R_{ij})] + \frac{g^2(R_{ij})}{1 - g^2(R_{ij})} \frac{Y_i^2 + Y_j^2}{\sigma_1^2} - \frac{2g(R_{ij})}{1 - g^2(R_{ij})} \frac{Y_i Y_j}{\sigma_1^2} \right\}\Bigg]$$

Error exponent through limit of LLR

$$D = \lim_{n \to \infty} \frac{1}{n} \mathsf{LLR}(\mathbf{Y}_n; \mathcal{G}_d), \quad \mathcal{H}_0$$

LLR is sum of graph functionals of a Marked process

Y_i are independent under $\mathcal{H}_0$

Detailed Methodology

LLR as sum of node and edge functionals of dependency graph

$$\mathsf{LLR}(\mathbf{Y}_n, \mathcal{G}_d) = n \log \frac{\sigma_1}{\sigma_0} + \frac{1}{2}\Bigg[\sum_{i \in \mathcal{V}} \Big(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2}\Big) Y_i^2$$
$$+ \sum_{\substack{(i,j)\in\mathcal{E}_d\\ i<j}} \Big\{ \log[1 - g^2(R_{ij})] + \frac{g^2(R_{ij})}{1-g^2(R_{ij})} \frac{Y_i^2 + Y_j^2}{\sigma_1^2} - \frac{2g(R_{ij})}{1-g^2(R_{ij})} \frac{Y_i Y_j}{\sigma_1^2} \Big\} \Bigg]$$

Error exponent through limit of LLR

$$D = \lim_{n\to\infty} \frac{1}{n} \mathsf{LLR}(\mathbf{Y}_n; \mathcal{G}_d), \quad \mathcal{H}_0$$

LLR is sum of graph functionals of a Marked process

Y_i are independent under $\mathcal{H}_0$

Detailed Methodology

LLR as sum of node and edge functionals of dependency graph

$$\mathsf{LLR}(\mathbf{Y}_n, \mathcal{G}_d) = n \log \frac{\sigma_1}{\sigma_0} + \frac{1}{2}\Bigg[\sum_{i \in \mathcal{V}} \Big(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2}\Big) Y_i^2$$
$$+ \sum_{\substack{(i,j) \in \mathcal{E}_d \\ i<j}} \Big\{ \log[1 - g^2(R_{ij})] + \frac{g^2(R_{ij})}{1 - g^2(R_{ij})} \frac{Y_i^2 + Y_j^2}{\sigma_1^2} - \frac{2g(R_{ij})}{1 - g^2(R_{ij})} \frac{Y_i Y_j}{\sigma_1^2} \Big\} \Bigg]$$

Error exponent through limit of LLR

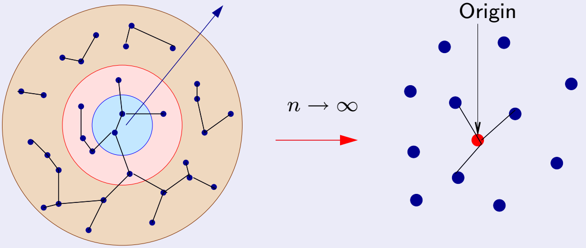
$$D = \lim_{n \to \infty} \frac{1}{n} \mathsf{LLR}(\mathbf{Y}_n; \mathcal{G}_d), \quad \mathcal{H}_0$$

LLR is sum of graph functionals of a Marked process

Y_i are independent under $\mathcal{H}_0$

LLN for graph functionals (Penrose & Yukich, 02)

Pictorial Representation of result



Normalized sum of edge weights

Expectation of edges of origin of Poisson process

$$\frac{\sum_{e\in\mathcal{E}} \Phi(R_e)}{n} \longrightarrow \mathbb{E} \sum_{\mathbf{x}\in\mathcal{P}_\lambda} \phi(R_{\mathbf{0},\mathbf{x}})$$

Remarks

LLN states that limit is a localized effect around origin

Result on Error Exponent D

Applying LLN (Penrose & Yukich, 02)

$$D = \frac{1}{2}\Big[\mathbb{E}\sum_{\mathbf{X}\in\mathcal{P}_\lambda} f(g(R_{\mathbf{0},\mathbf{X}})) + \log K + \frac{1}{K} - 1\Big],$$

$$f(x) \triangleq \log[1-x^2] + \frac{2x^2}{K[1-x^2]}, \qquad K \triangleq \frac{\sigma_1^2}{\sigma_0^2}$$

- $R_{\mathbf{0},\mathbf{X}}$: edge-lengths in a NNG of origin of a homogeneous Poisson process of intensity λ

Closed-form Expression for D

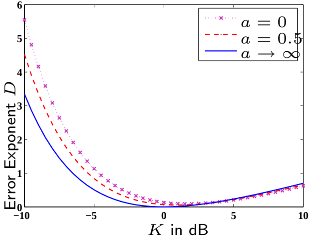
$$D = \frac{1}{2}\Big[\mathbb{E}f(g(Z_1)) - \frac{\pi}{2\omega}\mathbb{E}f(g(Z_2)) + \log K + \frac{1}{K} - 1\Big]$$

- Z_1, Z_2 : Rayleigh distributed with Variances $(2\pi\lambda)^{-1}$, $(2\omega\lambda)^{-1}$
- $\omega \approx 5.06$: area of union of two unit radii circles, with centers unit distant apart

Numerical Analysis

Questions

- How does correlation affect performance?
  - Depends on variance ratio
    - If signal under $\mathcal{H}_1$ is weak (low variance), correlation helps
    - If signal under $\mathcal{H}_1$ is strong (high variance), correlation hurts
- How does node density affect performance?
  - More node density more correlation as edge length is reduced

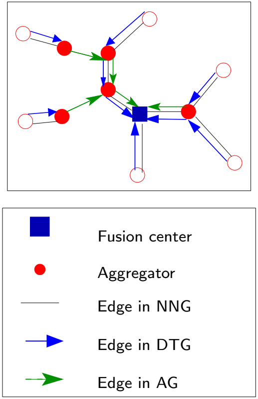


Correlation coefficient a, M=0.5, $\lambda = 1$

Exponential Correlation Function

$$g(r) = Me^{-ar},\ a > 0, 0 < M < 1$$

Minimum Energy Routing for Optimal Inference[1]



Transmission scheme delivering LLR

Minimum Energy Routing for Inference

Minimize total energy of routing such that LLR is delivered to fusion center

Summary of Results

- Concept of dependency graph based routing
  - Exploit correlation to fuse data
- Proposed 2-approximation algorithm

[1] A.Anandkumar, L.Tong, A. Swami, "Energy Efficient Routing for Statistical Inference of Markov Random Fields," *Conference on Information Sciences and Systems*, March 2007

Conclusion

Summary

- Derived a closed-form expression for error exponent of detection a GMRF with nearest-neighbor dependency
- Studied effect of correlation and node density on performance

Outlook

- Relax assumptions
- Extend to other dependency models
- Study Performance-Routing Energy tradeoff

Thank You !