

動的潜在グラフ構造の動的・非動的成分への分解

原 聡*

Satoshi Hara

鷲尾 隆*

Takashi Washio

Abstract: 共分散選択は確率変数間の条件付き独立性に基づいて変数間の本質的な依存関係（潜在グラフ構造）を解析する統計的手法である．本発表ではこの潜在グラフ構造が動的に変化する問題，特に変化が構造の一部にのみ現れる場合を扱う．これはシステムの部分的な故障を発見する異常検知や，依存関係の短い時間間隔での変化を解析する場合の自然な仮定であると考えられる．従来の共分散選択の枠組みを拡張した，変数間の依存関係を動的・非動的な成分へと分解する手法について述べ，その異常検知への応用について紹介する．

Keywords: 共分散選択，グラフィカル・ガウシアン・モデル，共通部分構造，ブロック座標降下法，異常検知

1 まえがき

実世界の多変量データ，例えば為替 [1]，遺伝子ネットワーク [2] や生体データ [3] は各観測変数の背後に複雑な変数間依存関係（潜在グラフ構造）を内包している．このような潜在グラフ構造はデータを生成する機構と密接な関係を持っている．例えば脳の各部位間の相互作用は fMRI 信号間の依存関係として表現される [3] ．

データの生成機構が周辺環境の変化を受けたり，時間に対する非定常性を有している場合，その変化に対応して潜在グラフ構造もまた動的に変化する．本研究では，このような構造変化が特に潜在グラフ構造の一部にのみ起こる場合を扱う．例えば工学システムではシステム全体の異常よりもその部分システムの故障の方がより起きやすい．この時，システム各部から得られるセンサデータには故障に起因する異常が見受けられる．しかし，そのような場合においても故障の影響を受けていない正常な部分システムの挙動は依然として不変な依存関係を示す．そのため，依存関係の変化を解析してその動的・非動的な成分を抽出することで故障部位を特定する手掛かりを得られることが期待される．この部分変動の仮定は工学システム以外でも，例えば比較的短い時間単位における変化を見る場合に，潜在グラフ構造の変化が急激な部分（動的な変化）と比較的ゆっくりとした部分（非動的な変化）の 2 種類を有している場合の近似的なモデルとしても解釈できる．

本研究ではこのように潜在グラフ構造が動的・非動的

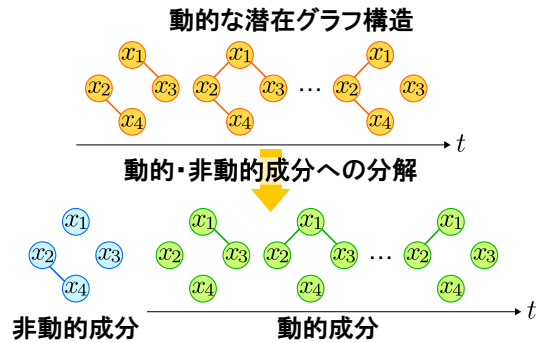


図 1: 潜在グラフ構造の動的・非動的成分への分解

な成分を持つ場合に両者を識別・分解する手法を提案する．提案法のイメージを図 1 に示す．我々は潜在グラフ構造のモデルとしてグラフィカルガウシアンモデル (GGM) を導入し，既存の GGM 構造推定問題 [4, 5, 6] の拡張として共通部分構造推定問題を定式化した．また，それに対してブロック座標降下法 [7] によるアルゴリズムを導出した．本稿の後半では，さらに共通部分構造推定の相関異常検知 [8] への応用について紹介する．

2 共分散選択

多変量解析において，変数間の関係性を計る指標として共分散がしばしば用いられる．しかし，一般に 2 変数 x_j と $x_{j'}$ の間の共分散は第 3 変数 x_k の影響を受ける．そのため，変数間の本質的な依存関係を調べるためにはこのような第 3 変数の影響を排除した，変数間の条件付き独立・従属性を解析する必要がある．確率変数 $x = (x_1, x_2, \dots, x_d)^T \in \mathbb{R}^d$ が正規分布に従う時，この変数間の条件付き独立・従属性は共分散行列 Σ の逆行

*大阪大学産業科学研究所，〒 567-0047 大阪府茨木市美穂ヶ丘 8-1, e-mail: {hara, washio}@ar.sanken.osaka-u.ac.jp, The Institute of Scientific and Industrial Research (ISIR), Osaka University, 8-1, Mihogaoka, Ibarakishi, Osaka, 5670047, Japan

列, 精度行列 Λ により表現される. 正規分布の下では, 変数 x_j と $x_{j'}$ が条件付き独立であることと, Λ の (j, j') 成分が 0 であることは同値である. また, このとき各変数 x_j を頂点とし, 隣接行列が Λ で与えられるグラフをグラフィカル・ガウシアン・モデル (GGM) と呼ぶ. 特に, 一部の变数間のみが依存関係を持つ場合は Λ の多くの非対角要素は 0 となり, GGM は疎なグラフとなる. 疎な GGM は変数間の本質的な依存構造を表現し, 背後のデータ生成機構に関する重要な知見を提供する. このような疎な潜在グラフ構造をデータから推定する問題は共分散選択 [9] と呼ばれている.

2.1 ℓ_1 正則化最尤推定

一般に真の精度行列 Λ が疎であっても, 標本から最尤推定等により得られた共分散行列 $\hat{\Sigma}$ の逆行列 $\hat{\Lambda} = \hat{\Sigma}^{-1}$ は密であり, 対応する GGM は完全グラフとなる. このような事態を避けて標本から疎な精度行列を学習する方法として近年 ℓ_1 正則化を用いる方法 [10] が提案された. これはさらに正規分布の最尤推定の枠組みへと拡張され, 以下の最適化問題として定式化された [4, 5]:

$$\max_{\Lambda \in \mathbb{R}^{d \times d}} \ell(\Lambda; \hat{\Sigma}) - \rho \|\Lambda\|_1 \quad \text{s.t. } \Lambda \succ 0. \quad (1)$$

ここで ρ は非負定数, $\ell(\Lambda; \hat{\Sigma})$ は次式で与えられる正規分布の対数尤度関数である:

$$\ell(\Lambda; \hat{\Sigma}) = \log \det \Lambda - \text{tr}(\hat{\Sigma} \Lambda). \quad (2)$$

また, $\Lambda \succ 0$ は Λ が有効な精度行列となるために正定値行列であることを要請している. 式 (1) を解くことにより得られる精度行列は第 2 項の ℓ_1 正則化の影響により疎となる. 式 (1) は凸制約条件下での凹関数最大化問題であり, ブロック座標降下法を用いたアルゴリズム (GLasso) [6] により効率的に解を得ることができる.

2.2 同時構造推定

ℓ_1 正則化最尤推定 (1) は 1 つの共分散行列 $\hat{\Sigma}$ から疎な精度行列 Λ を推定する問題であった. これをマルチタスク学習 [11] の枠組みを用いて N 個の共分散行列 $\hat{\Sigma}_1, \hat{\Sigma}_2, \dots, \hat{\Sigma}_N$ から N 個の疎な精度行列 $\Lambda_1, \Lambda_2, \dots, \Lambda_N$ を推定する問題へと拡張したものが GGM の同時構造推定問題である. Honorio [12] らは N 個の GGM が共通のグラフ構造を持つと仮定し, group Lasso [13] の方法を用いて以下の同時構造推定問題を定式化した:

$$\begin{aligned} \max_{\{\Lambda_i\}_{i=1}^N} & \sum_{i=1}^N t_i \ell(\Lambda_i; \hat{\Sigma}_i) - \rho \sum_{j \neq j'} \max_i |\Lambda_{i,jj'}| \\ \text{s.t. } & \Lambda_1, \Lambda_2, \dots, \Lambda_N \succ 0. \end{aligned} \quad (3)$$

ここで ρ は非負の正則化パラメータ, t_1, t_2, \dots, t_N は非負定数である. この問題では (1) とは異なり同時構造 $\hat{\Lambda}_{jj'} = \max_i |\Lambda_{i,jj'}|$ に制約を加えて疎にすることで, 行列の一部の成分について共通して $\Lambda_{1,jj'} = \Lambda_{2,jj'} = \dots = \Lambda_{N,jj'} = 0$ となるように構造の推定を行う. マルチタスク学習ではこのように N 個の推定問題間で情報を共有することにより結果の精度を向上させることができる.

3 共通部分構造推定問題

3.1 問題設定

GGM の共通構造推定問題の目的は N 個の共分散行列 $\hat{\Sigma}_1, \hat{\Sigma}_2, \dots, \hat{\Sigma}_N$ から, 精度行列 $\Lambda_1, \Lambda_2, \dots, \Lambda_N$ に共通する要素 (非動的な成分) を推定することである. ただし, 共分散行列は全て等しく d 次元で, 確率変数 x_1, \dots, x_d は同質であるとする (たとえば, どの共分散行列についても x_1 は "センサ A の値"). この仮定の下で複数 GGM の共通部分構造を以下により定義する:

定義 1 (複数 GGM の共通部分構造) 各 GGM に対応する精度行列を $\Lambda_1, \Lambda_2, \dots, \Lambda_N$ とする. この時, これらの共通部分構造を以下の隣接行列 Θ により定義する:

$$\Theta_{jj'} = \begin{cases} \Lambda_{1,jj'}, & \text{if } \Lambda_{1,jj'} = \Lambda_{2,jj'} = \dots = \Lambda_{N,jj'} \\ 0, & \text{otherwise} \end{cases}.$$

この定義は時系列データの弱定常性の定義を偏共分散へと拡張したものとして解釈することができる. ここで定義される共通部分構造は各変数間の依存構造のうち N 個の GGM 全てにおいて共通した重みを持つ辺のみを有するグラフである. 上記の共通部分構造の条件は N 個の精度行列間での最大変動幅 $\max_{i,i'} |\Lambda_{i,jj'} - \Lambda_{i',jj'}|$ が 0 であることと等価である. そこで, 我々はこれを新たな正則化項として加えた以下の正則化最尤推定問題を定式化した:

$$\begin{aligned} \max_{\{\Lambda_i\}_{i=1}^N} & \sum_{i=1}^N t_i \ell(\Lambda_i; \hat{\Sigma}_i) \\ & - \sum_{j \neq j'} \left(\rho \max_i |\Lambda_{i,jj'}| + \gamma \max_{i,i'} |\Lambda_{i,jj'} - \Lambda_{i',jj'}| \right) \\ \text{s.t. } & \Lambda_1, \Lambda_2, \dots, \Lambda_N \succ 0. \end{aligned} \quad (4)$$

ここで ρ, γ は非負の正則化パラメータ, t_1, t_2, \dots, t_N は $\sum_{i=1}^N t_i = 1$ を満たす非負定数であり, 通常は i 番目のデータセットのサイズを n_i とした時に $t_i = \frac{n_i}{\sum_{i=1}^N n_i}$ により与えられる. 1 つ目の正則化項は (3) で用いられている同時正則化項であり疎な推定結果を与える. 他方, 今回加えた 2 つ目の正則化項は N 個の推定結果のうち最大の変動幅に制約を与えるものであり, これに

より一部の成分が N 個の GGM について共通, つまり $\Lambda_{1,jj'} = \Lambda_{2,jj'} = \dots = \Lambda_{N,jj'}$ とすることができる. 最終的な共通部分構造は推定された精度行列に定義 1 を適用することにより得ることができる.

3.2 既存手法との関連

2つのデータセット間での依存構造の変化を検出する方法として Zhang[14] らの研究がある. これは従来の ℓ_1 正則化の枠組み [10] に Fused Lasso[15] 型の正則化を加えたものである. 今回定義した共通部分構造推定問題 (4) はこれを正規分布の正則化最尤法として書き直し, さらに $N \geq 3$ の場合へと拡張したより一般的な問題とみなすことができる. また, Zhang らの提案したアルゴリズムは $N = 2$ の場合に限定されたものであり, 本問題 (4) へと適用することはできない. 次章で我々の提案アルゴリズムを紹介する.

4 アルゴリズム

共通部分構造推定問題 (4) は凸問題であり, 解を効率的に得ることができる. 本章ではブロック座標降下法を用いた最適化アルゴリズムの概略¹を紹介する. なお, ブロック座標降下法の最適化への収束は [7] の定理 4.1 により保証される.

ブロック座標降下法では各精度行列 Λ_i の中の 1 つの成分を選び, それ以外の値を固定する. その上で対象とする成分に関する部分最適化問題を解く. これを対象とする成分を変えながら逐次的に解を更新していく. 各ステップにおける部分最適化問題は行列の対角成分と非対角成分についてそれぞれ異なった問題となる.

まず行列の (m, m) 成分 (対角成分) を最適化する場合を考える. この時, 精度行列, 共分散行列の中の変数 x_m に関する成分を行列の最終行・列へと並び替えて以下のように分割する:

$$\Lambda_i = \begin{bmatrix} Z_i & z_i \\ z_i^\top & \omega_i \end{bmatrix}, \quad \hat{\Sigma}_i = \begin{bmatrix} P_i & p_i \\ p_i^\top & q_i \end{bmatrix}. \quad (5)$$

今, 行列の (m, m) 成分 ω_i 以外は全て定数として扱い, 最適な ω_i の値は以下により与えられる:

$$\omega_i = z_i^\top Z_i^{-1} z_i + q_i^{-1}. \quad (6)$$

この時 $Z_i \succ 0$ であれば常に $\Lambda_i \succ 0$ となる. そのため, Λ_i の初期値を正定になるように選ぶことで, ブロック座標降下法の各ステップで常に行列の正定値性を保証することができる.

¹詳細については [16] を参照. また問題は多少異なるが [17] においても同様のアルゴリズムを採用している.

次に各行列 Λ_i の (m, m') 成分 v_i を最適化する問題を考える. これは N 個の行列の (m, m') 成分を並べたベクトル $\mathbf{v} = (v_1, v_2, \dots, v_N)^\top$ に関する以下の最適化問題として与えられる:

$$\min_{\mathbf{v}} \frac{1}{2} \mathbf{v}^\top \text{diag}(\mathbf{a}) \mathbf{v} - \mathbf{b}^\top \mathbf{v} + \rho \|\mathbf{v}\|_\infty + \gamma \max_{i,i'} |v_i - v_{i'}|. \quad (7)$$

ここで $\mathbf{a}, \mathbf{b} \in \mathbb{R}^N$ は各行列 Λ_i の (m, m') 成分以外の値及び $\hat{\Sigma}_i$ から定義されるベクトルである [16]. さらにこの問題の双対問題は以下で与えられる:

$$\min_{\xi} \frac{1}{2} (\mathbf{b} - \xi)^\top \text{diag}(\mathbf{a})^{-1} (\mathbf{b} - \xi) \quad \text{s.t.} \quad |\mathbf{1}_N^\top \xi| \leq \rho, \quad \|\xi\|_1 \leq \rho + 2\gamma. \quad (8)$$

ここで双対変数 ξ は $\xi = \mathbf{b} - \text{diag}(\mathbf{a})\mathbf{w}$ により定義される. この問題は解を実行可能領域上の位置に応じて 3 通りに場合分けすることができ, それぞれ Lasso, 連続 2 次ナップザック問題 [12] を解くことで解が得られる [16].

5 相関異常検知への応用

本章では共通部分構造推定の異常検知問題への応用について紹介する. 異常検知の目的は各変数が 2 つのデータセット間の差異にどれだけ寄与しているかを定量的に求めることである. 相関異常検知は各変数間の相関の変化に基づいて異常を判定する問題である. Idé[8] らはこれに対し疎な精度行列, 特に GLasso (2.1 節) [6] による推定結果が比較的ノイズにロバストであることを実験的に示し, 相関異常検知への応用を行った. 共通部分構造推定 (4) ではデータセット間での精度行列の変動幅に制約を加えているため, よりばらつきの少ない推定結果を与える. 特に, 異常に寄与していない変数間の構造は共通部分構造として推定されうる. このようなよりノイズにロバストで安定した推定結果を用いることで, 正常な変数と異常な変数をより明確に識別できるようになることが期待される.

5.1 相関異常スコア

Idé[8] らは精度行列に基づいて各変数の相関異常度をスコアリングする方法を提案した. 今, 2 つのデータセットがそれぞれ正規分布 $p_1(\mathbf{x}) = \mathcal{N}(\mathbf{0}_d, \Lambda_1^{-1})$, $p_2(\mathbf{x}) = \mathcal{N}(\mathbf{0}_d, \Lambda_2^{-1})$ から生成されているとする. この時, \mathbf{x}_j を \mathbf{x} から j 番目の変数を除いたものとする, それぞれのデータセットにおける変数 x_j の条件付き分布 $p_1(x_j | \mathbf{x}_{\setminus j})$ と $p_2(x_j | \mathbf{x}_{\setminus j})$ の間の Kullback-Leibler (KL) ダイバージェンス $D_{\text{KL}}[p_1(x_j | \mathbf{x}_{\setminus j}) || p_2(x_j | \mathbf{x}_{\setminus j})]$ を $p_1(\mathbf{x}_{\setminus j})$ について期

待値を取った量 d_j^{12} は以下で与えられる [8] :

$$d_j^{12} = \mathbf{p}_1^\top (\mathbf{z}_2 - \mathbf{z}_1) + \frac{1}{2} \left\{ \frac{\mathbf{z}_2^\top \mathbf{P}_2^{-1} \mathbf{z}_2}{\omega_2} - \frac{\mathbf{z}_1^\top \mathbf{P}_1^{-1} \mathbf{z}_1}{\omega_1} \right\} + \frac{1}{2} \left\{ \log \frac{\omega_1}{\omega_2} + q_1(\omega_2 - \omega_1) \right\}. \quad (9)$$

ここで変数 j に対応する成分を Λ_i の最終行・列へと並び替え、行列を以下のように分割した :

$$\Lambda_i = \begin{bmatrix} Z_i & \mathbf{z}_i \\ \mathbf{z}_i^\top & \omega_i \end{bmatrix}, \quad \Lambda_i^{-1} = \begin{bmatrix} P_i & \mathbf{p}_i \\ \mathbf{p}_i^\top & q_i \end{bmatrix}. \quad (10)$$

KL ダイバージェンスは非対称なため、各変数 x_j についての相関異常スコア a_j は以下により定義される :

$$a_j = \max(d_j^{12}, d_j^{21}). \quad (11)$$

5.2 実データへの適用実験

本節では提案法と前節の相関異常スコアを用いて実データの相関異常検知を行う。対象とするデータは Idé[8] により用いられた自動車センサデータである。これは 42 個のセンサ (変数) からなるデータセットであり、正常状態での 79 試行及びセンサ異常を含む状態での 20 試行からなる。ここでセンサ異常は 24 番目と 25 番目のセンサが逆に接続されていることに起因する。各試行につき 1 つの標本共分散行列が得られ、これらから異常センサを検出することが目的である。Idé[8] らは GLasso を用いて各標本共分散行列に対応する精度行列を求め、そこから前節の相関異常スコアにより異常検知を行った。しかし、今のように複数の標本共分散行列が得られている状況下ではそれらを合わせて用いることで検出精度がより向上することが期待できる。そこで我々は GLasso により精度行列を個別に推定するのではなく、共通部分構造推定により同時に全ての精度行列を推定し、それにより異常検知性能がどのように変化するかを観察する。本実験では比較対象として GLasso と従来の同時構造推定手法 (2.2 節, MSL) [12] を導入した。

以下が本実験の流れである。まず初めに対象データとして正常状態の 79 個の共分散行列から 20 個、異常状態の 20 個の共分散行列から 5 個をそれぞれランダムに選び出す。次に、この 25 個の共分散行列に対し、GLasso, MSL 及び提案法により精度行列の推定を行う。ただし、ここで MSL 及び提案法の重み t_i は 2 つの状態間での影響のバランスを取るために正常状態については $t_i = \frac{1}{40}$ 、異常状態については $t_i = \frac{1}{10}$ とした。最後に、得られた 25 個の精度行列について正常状態と異常状態から 1 つずつ選び出した各ペア 20×5 通りについて相関異常スコアを計算する。

5.3 実験結果

以上の実験を選び出す共分散行列を変えて 100 回繰り返した時の平均の異常検知性能の ROC 曲線を図 2 に示す。ここで正則化パラメータ ρ は結果の AUC が最大になるように選び、また提案法の γ についてはヒューリスティック [16] により計算した。ここではまず提案法及び MSL が GLasso よりも良い異常検知性能を与えていることを見ることができる。これは提案法、MSL が共に精度行列の同時推定を行う手法であり、GLasso により個別に推定を行った場合よりも推定結果のばらつきが抑えられ安定するためと考えられる。

提案法と GLasso, MSL の差異をより詳細に見るために、100 回の繰り返し実験により得られた相関異常スコアの中央値及び 25%, 75% 点を図 3 に示す。ここではそれぞれ正常-異常状態間及び正常状態間、異常状態間におけるスコアを各手法について描画した。まず正常-異常状態間のスコア中央値を各手法について比較すると、提案法が 16 ~ 21, 33 ~ 42 番のセンサを中心に正常なセンサについてより低いスコアを与えていることを見ることができる。また 25%, 75% 点からは提案法が他 2 手法に比べてばらつきが少ない、安定したスコアを与えていることがわかる。この傾向は正常状態間、異常状態間ではより顕著になっている。正常状態間、異常状態間ではどちらの場合についても全ての変数についてスコアが 0 になることが理想的な結果である。しかし、提案法の結果からは 3 番や 7 番, 11 番や 26 番は各状態においても高いスコアを与えていることを見ることができる。これらの結果を正常-異常状態のものに対比することで相関異常を与えているセンサの候補を 22 番, 24 番, 25 番, 28 番, 32 番へとさらに絞り込むことができる。GLasso や MSL は各状態での相関異常スコアも比較的高い値を有しており、またそのばらつきも大きいため提案法に比べてこのような対比が困難となっている。今回の実験においてはこのような提案法の優位性は異常検出性能そのものには直接的に影響を及ぼさず MSL と同等の結果であったが、より複雑な相関異常を有するデータへと適応する際にはその差異が明確になることが期待される。

6 おわりに

本研究ではデータの潜在グラフ構造を動的・非動的成分へと分解する手法の開発を行った。我々はこの問題を複数の GGM に対する共通部分構造推定として定式化し、それに対しブロック座標降下法による最適化アルゴリズムを与えた。また共通部分構造推定の相関異常検知への応用について実データに基づいた検討を行った。実

データ実験において提案法は従来法と同等の異常検知性能を与え、また得られた異常スコアは提案法の有用性を示すものであった。

今後の課題としては共通部分構造推定の漸近的挙動の解析や adaptive Lasso 型 [18] の定式化への拡張などが挙げられる。また、より異常検知に適した構造推定手法の開発も興味深い問題である。

謝辞

本研究は日本学術振興会科学研究費補助金、基盤研究(B)#22300054の補助を受けて行われた。また、自動車センサデータをご提供くださったIBM東京基礎研究所の井手剛氏と本研究を進める上で有用なアドバイスを下さった清水昌平氏に感謝申し上げる。

参考文献

- [1] R.T. Baillie, and T. Bollerslev, “Common stochastic trends in a system of exchange rates,” *The Journal of Finance*, vol.44, no.1, pp.167–181, 1989.
- [2] B. Zhang, H. Li, R.B. Riggins, M. Zhan, J. Xuan, Z. Zhang, E.P. Hoffman, R. Clarke, and Y. Wang, “Differential dependency network analysis to identify condition-specific topological changes in biological networks,” *Bioinformatics*, vol.25, no.4, pp.526–532, 2009.
- [3] G. Varoquaux, A. Gramfort, J.B. Poline, and B. Thirion, “Brain covariance selection: better individual functional connectivity models using population prior,” *Arxiv preprint arXiv:1008.5071*, 2010.
- [4] M. Yuan, and Y. Lin, “Model selection and estimation in the gaussian graphical model,” *Biometrika*, vol.94, pp.19–35, 2007.
- [5] O. Banerjee, L. El Ghaoui, and A. d’Aspremont, “Model selection through sparse maximum likelihood estimation for multivariate gaussian or binary data,” *The Journal of Machine Learning Research*, vol.9, pp.485–516, 2008.
- [6] J. Friedman, T. Hastie, and R. Tibshirani, “Sparse inverse covariance estimation with the graphical lasso,” *Biostatistics*, vol.9, no.3, p.432, 2008.
- [7] P. Tseng, “Convergence of a block coordinate descent method for nondifferentiable minimization,” *Journal of Optimization Theory and Applications*, vol.109, no.3, pp.475–494, 2001.
- [8] T. Idé, A.C. Lozano, N. Abe, and Y. Liu, “Proximity-based anomaly detection using sparse structure learning,” *Proceedings of the 2009 SIAM International Conference on Data Mining*, 2009.
- [9] A. Dempster, “Covariance selection,” *Biometrics*, vol.28, no.1, pp.157–175, 1972.
- [10] N. Meinshausen, and P. Bühlmann, “High-dimensional graphs and variable selection with the lasso,” *The Annals of Statistics*, vol.34, no.3, pp.1436–1462, 2006.
- [11] R. Caruana, “Multitask learning,” *Machine Learning*, vol.28, no.1, pp.41–75, 1997.
- [12] J. Honorio, and D. Samaras, “Multi-task learning of gaussian graphical models,” *In Proc. 27th Conf. on Machine Learning*, 2010.
- [13] F.R. Bach, “Consistency of the group lasso and multiple kernel learning,” *The Journal of Machine Learning Research*, vol.9, pp.1179–1225, 2008.
- [14] B. Zhang, and Y. Wang, “Learning structural changes of gaussian graphical models in controlled experiments,” *In Proc. 26th Conf. on Uncertainty in Artificial Intelligence*, 2010.
- [15] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, and K. Knight, “Sparsity and smoothness via the fused lasso,” *Journal of the Royal Statistical Society: Series B*, vol.67, no.1, pp.91–108, 2005.
- [16] S. Hara, and T. Washio, “Common substructure learning of multiple graphical gaussian models,” *In Proc. ECML PKDD 2011, 2011 (to appear)*.
- [17] S. Hara, and T. Washio, “Simultaneous learning of graphical structures,” *The 25th Annual Conference of the Japanese Society for Artificial Intelligence*, 2011 (in Japanese).
- [18] H. Zou, “The adaptive lasso and its oracle properties,” *Journal of the American Statistical Association*, vol.101, no.476, pp.1418–1429, 2006.

	best AUC	ρ
Proposed	0.97	0.05
GLasso	0.96	0.20
MSL	0.97	0.05

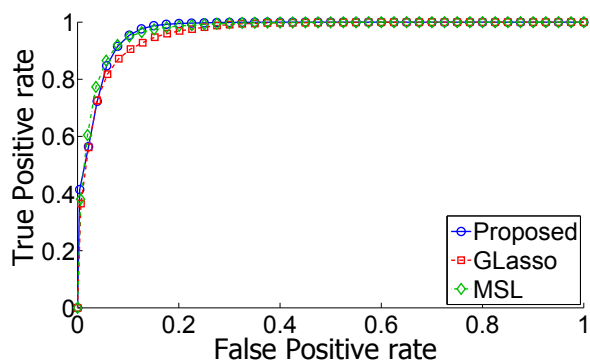


図 2: 相関異常検知性能: 最大 AUC と ROC 曲線

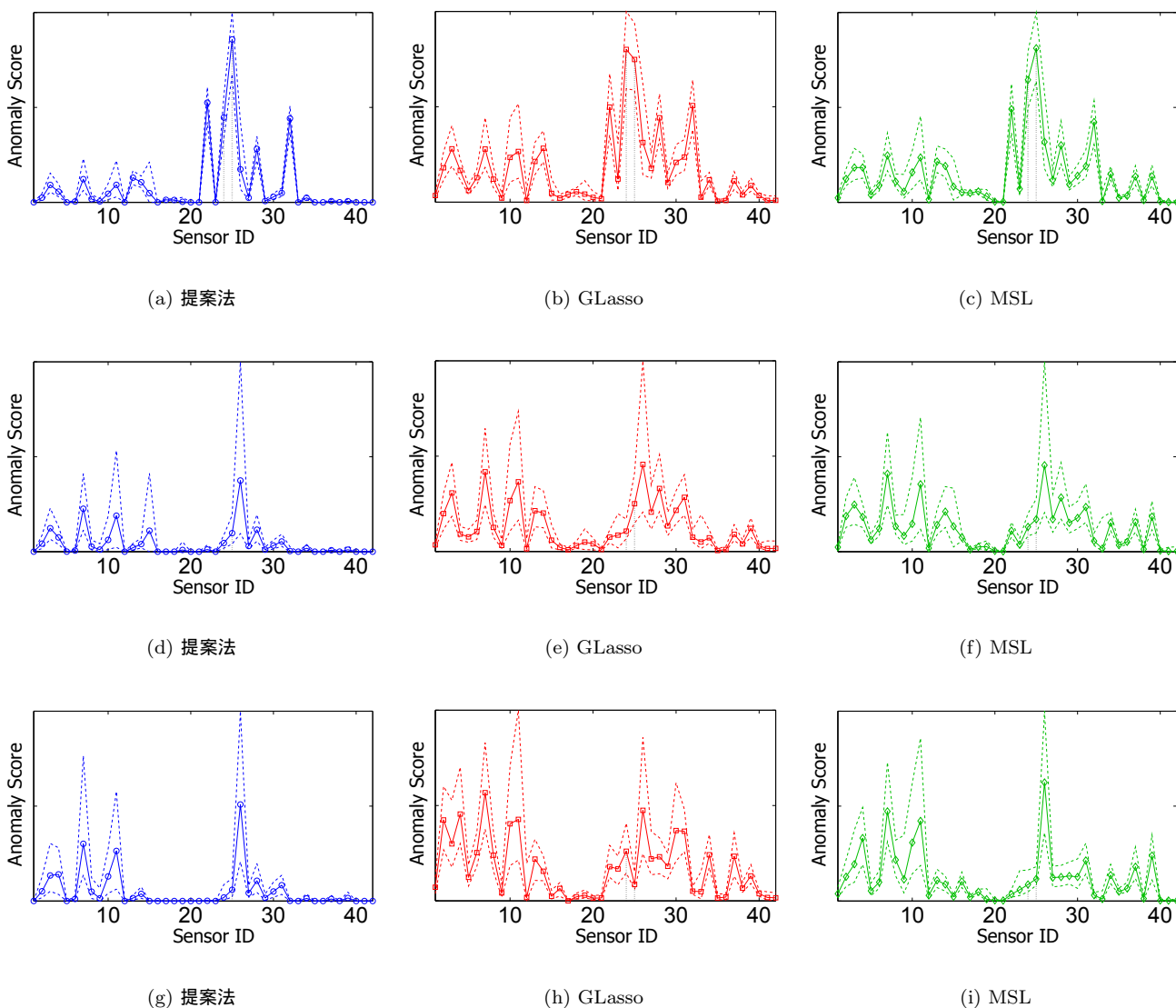


図 3: 相関異常スコア: 各プロットは全て最大値が等しくなるように正規化してある。各グラフ中の実線は各センサについて 100 回の実験により得られたスコアの中央値, 破線はそれぞれ 25%点, 75%点を表している。また, 縦の点線は真の異常センサを表している。上段 (a), (b), (c) は正常-異常状態間の相関異常スコア, 中段 (d), (e), (f) は正常状態間の相関異常スコア, そして下段 (g), (h), (i) は異常状態間の相関異常スコアである。