دانشگاه تربیت مدرس

دانشکده علوم ریاضی

پایان نامه دوره کارشناسی ارشد آمار

تحلیل مدل‌های ... با روش تقریب لاپلاس آشیانی ترکیبی

توسط

کبری قلی‌زاده
زهرا قیومی

استاد راهنما

دکتر محسن محمدزاده

تیر ماه ۱۳۹۱

فهرست مندرجات

۱ مفاهیم مقدماتی ۱

۱.۱ مقدمه ۱

۲.۱ تعاریف و مفاهیم اولیه ۱

۱.۲.۱ استقلال شرطی ۳

۲.۲.۱ گراف ۳

۳.۲.۱ ماتریس همیشه مثبت متقارن ۵

۳.۱ میدان تصادفی مارکوفی گاوسی ۷

۱.۳.۱ خواص مارکوفی GMRF ۹

۲.۳.۱ ویژگی‌های شرطی میدان تصادفی مارکوفی گاوسی ۱۱

۳.۳.۱ مشخص‌سازی GMRF با شرطی‌های کامل ۱۲

۴.۳.۱ میدان تصادفی مارکوفی گاوسی چندمتغیره ۱۴

۴.۱ الگوریتم‌های محاسباتی و شبیه‌سازی GMRF ۱۵
۱.۴.۱ شبیه‌سازی غیرشرطی میدان تصادفی مارکوفی گاوسی ۱۷
۲.۴.۱ شبیه‌سازی شرطی میدان تصادفی مارکوفی گاوسی ۲۰

۵.۱ روش‌های عددی برای ماتریس‌های تُنک ۲۳
۱.۵.۱ تجزیهٔ یک ماتریس تُنک ۲۳

۲ میدان تصادفی مارکوفی گاوسی ذاتی ۳۰

۱.۲ تعاریف و مفاهیم اولیه ۳۱
۱.۱.۲ تفاضل پیشرو ۳۱
۲.۱.۲ چندجمله‌ای‌ها ۳۲

۲.۲ میدان تصادفی مارکوفی گاوسی تحت قیدهای خطی ۳۴

۳.۲ میدان تصادفی مارکوفی گاوسی ذاتی مرتبهٔ اول ۳۹
۱.۳.۲ میدان تصادفی مارکوفی گاوسی ذاتی مرتبهٔ اول روی خط ۴۰
۲.۳.۲ میدان تصادفی مارکوفی گاوسی ذاتی مرتبهٔ اول روی مشبکه ۴۴

۴.۲ میدان‌های تصادفی مارکوفی گاوسی ذاتی مراتب بالا ۴۶
۱.۴.۲ میدان تصادفی گاوسی مارکوفی ذاتی مراتب بالاتر روی خط ۴۷

۵.۲ مدل قدم زدن تصادفی زمان پیوسته ۵۲

۳ تقریب لاپلاس آشیانی ترکیبی ۵۶

۱.۳ تقریب گاوسی ۵۸

۲.۳ تقریب لاپلاس ۶۱
۱.۲.۳ میانگین و واریانس پسینی ۶۲
۲.۲.۳ چگالی پسینی کناری ۶۴

۳.۳ تقریب لاپلاس آشیانی ترکیبی ۶۴
۱.۳.۳ تقریب پسینی توام بردار ابرپارامترها ۶۵
۲.۳.۳ طرح مرکب مرکزی ۶۹
۳.۳.۳ طرح ۷۰
۴.۳.۳ محاسبه تقریب کناری عناصر بردار ابرپارامتر ۷۴
۵.۳.۳ محاسبهٔ تقریب کناری پسینی عناصر میدان پنهان ۷۴

فصل ۱

مفاهیم مقدماتی

۱.۱ مقدمه

۲.۱ تعاریف و مفاهیم اولیه

در این پایان‌نامه بردارها و ماتریس‌ها به صورت برجسته، مانند $\boldsymbol{x}$ و $\boldsymbol{A}$ نشان داده می‌شوند. ترانهادهٔ $\boldsymbol{A}$ با $\boldsymbol{A}^T$ مشخص می‌شود. نماد $\boldsymbol{A} = (A_{ij})$ مشخص می‌کند که مولفهٔ سطر iام و ستون jام ماتریس $\boldsymbol{A}$، A_{ij} است. همچنین مشابه این نماد نیز برای بردار استفاده می‌شود، به عبارتی $\boldsymbol{x} = (x_i)$. $\boldsymbol{x}_{i:j}$ کوتاه شدهٔ بردار $\boldsymbol{x}$ به‌صورت $(x_i, x_{i+1}, \ldots, x_j)^T$ است. اگر $\boldsymbol{A}_{n\times m}$ ماتریسی با m ستون $\boldsymbol{A}_m, \ldots, \boldsymbol{A}_1$ باشد، $Vec(A) = (\boldsymbol{A}_m, \ldots, \boldsymbol{A}_1)$ برداری ستونی است که از زیر هم قرار دادن ستون‌های ماتریس $\boldsymbol{A}$ حاصل می‌شود. ماتریسی که از حذف بعضی از سطرها و ستون‌های $\boldsymbol{A}$ به‌دست می‌آید زیرماتریس $\boldsymbol{A}$ نامیده می‌شود. زیر ماتریسی را که از حذف سطر و ستون با اندیس یکسان به‌دست آید زیرماتریس اصلی گویند. به‌عنوان مثال، ماتریس $\boldsymbol{B} = \begin{pmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{pmatrix}$ که از حذف سطر و ستون سوم ماتریس $\boldsymbol{A}$

حاصل شده است یک زیرماتریس اصلی ماتریس $\boldsymbol{A}$ است. زیر ماتریس $r \times r$، که از حذف $n-r$ سطر و ستون آخر ماتریس $\boldsymbol{A}$ به‌دست می‌آید زیر ماتریس اصلی مقدم [۱] $\boldsymbol{A}$ نامیده می‌شود.

اگر $\boldsymbol{A}$ ماتریس $n \times n$ و $\boldsymbol{a}$ برداری با $n \times ۱$ باشد، آن‌گاه $diag(\boldsymbol{A})$ و $diag(\boldsymbol{a})$ به ترتیب به منظور نمایش ماتریس‌های قطری

$$diag(\boldsymbol{A}) = \begin{pmatrix} A_{۱۱} & & \\ & \ddots & \\ & & A_{nn} \end{pmatrix}, \qquad diag(\boldsymbol{a}) = \begin{pmatrix} a_{۱} & & \\ & \ddots & \\ & & a_{n} \end{pmatrix}$$

استفاده می‌شوند. پهنای‌باند ماتریس $\boldsymbol{A}$ به‌صورت $\max\{|i-j| : A_{ij} \neq ۰\}$ تعریف می‌شود. به‌طور مشابه پهنای‌باند پایینی $\boldsymbol{A}$ برابر با $\max\{|i-j| : A_{ij} \neq ۰, \quad i > j\}$ و پهنای باند بالایی برابر $\max\{|i-j| : A_{ij} \neq ۰, \quad i < j\}$ است. $\mathcal{I}_{\boldsymbol{n}} = \{(i,j) : i = ۱, \ldots, n_{۱}, \quad j = ۱, \ldots, n_{۲}\}$ نشانگر مشبکه [۲] (توری [۳]) از اندازهٔ $\boldsymbol{n} = (n_{۱}, n_{۲})$ است. برای $C \subset \mathcal{I} = \{۱, \ldots, n\}$، $\boldsymbol{x}_C = \{x_i : i \in C\}$ تعریف می‌شود. اگر مجموعهٔ $\mathcal{I} - C$ با $-C$ نشان داده شود، آن‌گاه $\boldsymbol{x}_{-C} = \{x_i : i \in -C\}$. برای دو زیرمجموعهٔ $C_{۱}$ و $C_{۲}$ از $\mathcal{I}$، $C_{۱} \backslash C_{۲} = \{i \in \mathcal{I} : \rangle \in \mathcal{C}_{\infty}, \quad \rangle \notin \mathcal{C}_{\in}\}$ است.

ل‌اشه ن‌اوند ه‌ب .دوشی‌م ه‌دیمان [۴] flop)(روانش زیمم م‌قر ت‌رابع ک‌ی ه‌بساحم ی‌اهرگلمع دادعت

.ت‌سا روانش زیمم م‌قر ود ی‌اراد ت‌سا ع‌مج رگلمع ک‌ی و ب‌رض رگلمع ک‌ی ل‌ماش ه‌ک $x + a * b$ ت‌رابع

---

[۱] Leading Principal Submatrix

[۲] Lattice

[۳] Grid

[۴] Floating Point Operation

۱.۲.۱ استقلال شرطی

برای $C \subset \mathcal{I}$ چگالی شرطی $\boldsymbol{x}_C$ به شرط $\boldsymbol{x}_{-C}$ عبارت است از

$$\pi(\boldsymbol{x}_C|\boldsymbol{x}_{-C}) = \frac{\pi(\boldsymbol{x}_C, \boldsymbol{x}_{-C})}{\pi(\boldsymbol{x}_{-C})} \propto \pi(\boldsymbol{x}). \qquad (۱.۲.۱)$$

دو متغیر تصادفی X و Y مستقل هستند، اگر و تنها اگر $\pi(x,y) = \pi(x)\pi(y)$ و با نماد $X \perp Y$ نمایش داده می‌شود. دو متغیر تصادفی X و Y برای Z داده شده مستقل شرطی نامیده می‌شوند، اگر و تنها اگر $\pi(x,y|z) = \pi(x|z)\pi(y|z)$ و با نماد $X \perp Y|Z$ نمایش داده می‌شود.

قضیه ۱.۲.۱ (رو و هلد، ۲۰۰۵):

$$x \perp y|z \quad \Leftrightarrow \quad \pi(x,y,z) = f(x,z)g(y,z) \qquad (۲.۲.۱)$$

برای بعضی توابع f و g، و همهٔ zهایی که $\pi(z) > 0$ است.

۲.۲.۱ گراف

گراف مجموعه‌ای از رأس‌ها است که تعدادی یال به هم وصل شده‌اند. گراف[۵] $\mathcal{G}$ به صورت زوج مرتب $(\mathcal{V}, \mathcal{E})$ نشان داده می‌شود، که در آن $\mathcal{V}$ مجموعه‌ای متناهی و غیر تهی از راس‌ها و $\mathcal{E}$ زیر مجموعهٔ تمام زوج‌های نامرتب از عناصر نامساوی $\mathcal{V}$ به صورت

$$\mathcal{E} \subseteq \{\{i,j\} : i,j \in \mathcal{V}, i \neq j\}$$

است. اعضای $\mathcal{V}$ را رأس‌های گراف $\mathcal{G}$ و اعضای $\mathcal{E}$ را یال‌های آن می‌نامند. بین دو رأس یک گراف ساده[۶] حداکثر یک یال وجود دارد.

---

[۵] Graph

[۶] Simple Graph

اگر $\mathcal{E}$ زیر مجموعه‌ای از تمام زوج‌های مرتب عناصر $\mathcal{V}$ باشد، یعنی

$$\mathcal{E} \subseteq \{(i,j) : i,j \in \mathcal{V}, i \neq j\}$$

آن‌گاه $\mathcal{G}$ گراف سودار [۷] نامیده می‌شود.

گراف $\mathcal{G}$ کامل نامیده می‌شود هرگاه بین هر دو رأس آن دقیقاً یک یال وجود داشته باشد. در ادامه فرض می‌شود که $\mathcal{V} = \{۱, \ldots, n\}$ است، در این موارد گراف را نشان‌دار[۸] گویند. شکل ۱.۲.۱ یک گراف ساده شامل سه رأس و دو یال را نشان می‌دهد.

شکل ۱.۲.۱: نمایی از یک گراف ساده

همسایه‌های[۹] رأس i، رأس‌هایی در گراف $\mathcal{G}$ هستند که یالی مشترک با آن دارند، به بیان دیگر

$$ne(i) = \{j \in \mathcal{V} : \{i,j\} \in \mathcal{E}\}.$$

به‌طور مشابه برای مجموعه $A \subset \mathcal{V}$، همسایه‌های A به صورت

$$ne(A) = \bigcup_{i \in A} ne(i) \setminus A$$

تعریف می‌شود و شامل همه رأس‌هایی است که در A نیستند اما همسایگی با رأس های موجود در A دارند. یک مسیر از رأس $i_۱$ به رأس i_m دنباله‌ای مجزا از رئوس داخل $\mathcal{V}$ به‌صورت $i_۱ \ldots i_m$ است که

[۷] directed graph

[۸] Labelled

[۹] Neighbors

به ازای هر $j = ۱, \dots m-۱$، زوج مرتب $(i_j, i_{j+۱})$ در $\mathcal{E}$ باشد.

زیر مجموعه $C \subset \mathcal{V}$ جدا کنندهٔ دو رأس $i \notin C$ و $j \notin C$ است، اگر هر مسیر از i به j شامل حداقل یک رأس از C باشد. دو مجموعهٔ مجزای $A \subset \mathcal{V} \setminus C$ و $B \subset \mathcal{V} \setminus C$ به‌وسیله C جدا شده هستند، اگر هر $i \in A$ و $j \in B$ به‌وسیله C جدا شده باشند، یعنی نمی‌توان روی گراف از A به سمت B قدم زد بدون این که از C عبور کرد.

وقتی دو رأس i و j همسایه‌هایی در گراف $\mathcal{G}$ باشند، با نماد $i \overset{\mathcal{G}}{\sim} j$ نشان داده می‌شود. بدیهی است که $j \overset{\mathcal{G}}{\sim} i$ اگر و تنها اگر $i \overset{\mathcal{G}}{\sim} j$.

اگر A زیر مجموعه‌ای از $\mathcal{V}$ باشد، گراف حاصل بعد از حذف همه رأس‌هایی که در A نیستند و همهٔ یال‌هایی که حداقل یک رأس آن‌ها در A قرار دارند را گراف محدود شده به A نامند و یا ابا $\mathcal{G}^A = \{\mathcal{V}^A, \mathcal{E}^A\}$ نمایش داده می‌شود، که در آن

$$\mathcal{V}^A = A, \qquad \mathcal{E}^A = \{\{i,j\} \in \mathcal{E}, \{i,j\} \in A \times A\}.$$

است.

مثال ۱.۲.۱ اگر $\mathcal{G}$ گراف شکل ۱.۲.۱ و $A = \{۱, ۲\}$ باشد، آن‌گاه گراف محدود به A عبارت است از $\mathcal{G}^A = \{\mathcal{V}^A, \mathcal{E}^A\}$ که در آن $\mathcal{V}^A = \{۱, ۲\}$ و $\mathcal{E}^A = \{۱, ۲\}$ است.

۳.۲.۱ ماتریس همیشه مثبت متقارن

ماتریس $\boldsymbol{A}_{n\times n}$ همیشه مثبت است اگر و تنها اگر برای هر بردار $\boldsymbol{x} \neq \mathbf{0}$، $\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} > ۰$ باشد. اگر $\boldsymbol{A}$ متقارن نیز باشد به آن ماتریس همیشه مثبت متقارن[۱۰] (SPD) گویند. ماتریس $n \times n$، $\boldsymbol{A}$ نیمه

[۱۰] Symmetric Positive Definite

مثبت[۱۱] نامیده می‌شود، اگر و تنها اگر

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} \geq ۰, \quad \forall \boldsymbol{x} \neq \mathbf{0}.$$

و با نماد $\boldsymbol{A} \geq ۰$ نشان داده می‌شود. همچنین اگر $\boldsymbol{A}$ متقارن باشد به آن ماتریس نیمه‌مثبت متقارن (SPSD) گویند.

برخی از خواص ماتریس‌های SPD عبارتند از

۱) $rank(\boldsymbol{A}) = n$

۲) $|\boldsymbol{A}| > ۰$

۳) $A_{ii} > ۰$

۴) به ازای $i \neq j$، $A_{ii} - A_{jj} - A_{ij}^۲ > ۰$

۵) به ازای $i \neq j$، $A_{ii} - A_{jj} - ۲|A_{ij}| > ۰$

٦) $\max A_{ii} > \max_{i \neq j} |A_{ij}|$

۷) $\boldsymbol{A}^{-۱}$ ماتریس SPD است.

۸) همه زیرماتریس‌های اصلی $\boldsymbol{A}$ نیز SPD هستند.

۹) اگر $\boldsymbol{A}$ و $\boldsymbol{B}$ هر دو ماتریس SPD باشند، $\boldsymbol{A} + \boldsymbol{B}$ نیز SPD خواهد بود، اما عکس آن لزوماً برقرار نیست.

۱۰) اگر $\boldsymbol{A}$ و $\boldsymbol{B}$ هر دو ماتریس SPD و $\boldsymbol{AB} = \boldsymbol{BA}$ باشد، آن گاه $\boldsymbol{AB}$ نیز یک ماتریس SPD است.

[۱۱] Positive Semidefinite

۱۱) شرایط لازم و کافی برای این که ماتریس $\boldsymbol{A}$، SPD باشد عبارتند از:

الف) همه مقادیر ویژهٔ $\boldsymbol{A}$ اکیداً مثبت هستند.

ب) ماتریسی مانند $\boldsymbol{C}$ وجود دارد که $\boldsymbol{A} = \boldsymbol{C}\boldsymbol{C}^T$ است.

ج) همهٔ زیرماتریس‌های اصلی مقدم، دارای دترمینان اکیداً مثبت هستند.

۳.۱ میدان تصادفی مارکوفی گاوسی

فرض کنید بردار تصادفی $\boldsymbol{x} = (x_1, \ldots, x_n)^T$ دارای توزیع چندمتغیرهٔ نرمال با میانگین $\boldsymbol{\mu}$ و ماتریس کواریانس $\boldsymbol{\Sigma}$ باشد. به علاوه $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ گرافی نشان‌دار باشد، که در آن $\mathcal{E}$ شامل همهٔ زوج‌های $\{i, j\}$ است که راس‌های i و j هیچ یال مشترکی نداشته باشند، اگر و تنها اگر $x_i \perp x_j | x_{-ij}$. در این صورت $\boldsymbol{x}$ یک میدان تصادفی مارکوفی گاوسی[۱۲] (GMRF) نسبت به گراف $\mathcal{G}$ نامیده می‌شود.

قبل از بیان یک تعریف رسمی برای GMRF، ارتباط بین گراف G و پارامترهای توزیع نرمال بیان می‌شود. از آن‌جا که $\boldsymbol{\mu}$ تاثیری بر استقلال شرطی دوبه‌دوی درایه‌های $\boldsymbol{x}$ ندارد، می‌توان استنباط کرد که این اطلاعات تنها در ماتریس کواریانس $\boldsymbol{\Sigma}$ پنهان شده است، بنابراین ماتریس دقت، که به‌صورت $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1}$ تعریف می‌شود، نقش مهمی را در تحلیل GMRF ایفا می‌کند.

قضیه ۱.۳.۱ (رو و هلد، ۲۰۰۵): اگر $\boldsymbol{x}$ دارای توزیع نرمال چندمتغیره با میانگین $\boldsymbol{\mu}$ و ماتریس دقت $\boldsymbol{Q} > \mathbf{0}$ باشد، آن‌گاه درایهٔ ij ام ماتریس $\boldsymbol{Q}$، یعنی Q_{ij} برابر صفر است۷ اگر و تنها اگر $x_i \perp x_j \mid \boldsymbol{x}_{-ij}$.

[۱۲] Gaussian Markov Random Fields

با توجه به قضیهٔ ۱.۳.۱ درایه‌های غیر صفر ماتریس $\boldsymbol{Q}$ گراف $\mathcal{G}$ را مشخص می‌کند و براساس آن‌ها می‌توان استقلال شرطی x_i و x_j را بررسی نمود.

تعریف ۱.۳.۱ **میدان تصادفی مارکوفی گاوسی**: بردار تصادفی $\boldsymbol{x} = (x_1, \ldots, x_n)^T \in R^n$ یک GMRF، تحت گراف نشان‌دار $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ با میانگین $\boldsymbol{\mu}$ و ماتریس دقت $\boldsymbol{Q} > \mathbf{0}$ است، اگر و تنها اگر تابع چگالی آن به‌صورت

$$\pi(\boldsymbol{x}) = (2\pi)^{-\frac{n}{2}} |\boldsymbol{Q}|^{\frac{1}{2}} \exp\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{Q}(\boldsymbol{x}-\boldsymbol{\mu})\} \qquad (۳.۳.۱)$$

باشد و برای هر $i \neq j \in \mathcal{E}$

$$Q_{ij} \neq 0 \iff \{i, j\} \in \mathcal{E}.$$

اگر $\boldsymbol{Q}$ یک ماتریس چگال باشد، یعنی تُنک نباشد، گراف $\mathcal{G}$ همبند کامل[۱۳] نامیده می‌شود و نشان دهندهٔ آن است که هر توزیع نرمال با ماتریس دقت SPD یک GMRF است و هر GMRF دارای توزیع نرمال با ماتریس دقت SPD است.

قضیه ۲.۳.۱ (رو و هلد، ۲۰۰۵): فرض کنید $\boldsymbol{x}$ یک GMRF تحت گراف $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ با میانگین $\boldsymbol{\mu}$ و ماتریس دقت $\boldsymbol{Q} > \mathbf{0}$ باشد، آن‌گاه میانگین، دقت و همبستگی شرطی به ترتیب عبارتند از

$$E(x_i|\boldsymbol{x}_{-i}) = \mu_i - \frac{1}{Q_{ii}} \sum_{j:j\sim i} Q_{ij}(x_j - \mu_j) \qquad (۴.۳.۱)$$

$$Prec(x_i|\boldsymbol{x}_{-i}) = Q_{ii} \qquad (۵.۳.۱)$$

$$Corr(x_i, x_j|\boldsymbol{x}_{-ij}) = -\frac{Q_{ij}}{\sqrt{Q_{ii}Q_{jj}}}, \qquad i \neq j. \qquad (۶.۳.۱)$$

[۱۳] Full Connected

درایه‌های قطری ماتریس $\boldsymbol{Q}$ دقت‌های شرطی x_i به شرط $\boldsymbol{x}_{-i}$ و مولفه‌های خارج از قطر اطلاعاتی در مورد همبستگی شرطی بین x_i و x_j به شرط $\boldsymbol{x}_{-ij}$ ارائه می‌کنند. اما از آن‌جا که درایه‌های ماتریس کواریانس $\boldsymbol{\Sigma} = (\Sigma_{ij})$، واریانس $Var(x_i) = \Sigma_{ii}$ و همبستگی $Corr(x_i, x_j) = \Sigma_{ij}/\sqrt{(\Sigma_{ii}\Sigma_{jj})}$ را نتیجه می‌دهند، ماتریس کواریانس اطلاعاتی در مورد واریانس کناری x_i و همبستگی کناری بین x_i و x_j دراختیار قرار می‌دهد. تفسیر کناری با Σ مستقیماً حاوی اطلاعات مفیدی است و می‌توان از آن برای بیان توزیع‌های کناری یک یا دو بعدی متغیره استفاده کرد، در حالی که در مورد $\boldsymbol{Q}$ این کار دشوار است و باید از توزیع توأم روی $\boldsymbol{x}_{-i}$ یا $\boldsymbol{x}_{-ij}$ انتگرال گرفت، زیرا با توجه به تعریف $\boldsymbol{Q}^{-1} = \boldsymbol{\Sigma}$[۱]، در حالت کلی Σ_{ii} به همهٔ مولفه‌های $\boldsymbol{Q}$ مرتبط است.

۱.۳.۱ خواص مارکوفی GMRF

گراف $\mathcal{G}$ با بررسی وجود یا عدم وجود استقلال شرطی $x_i \perp x_j | \boldsymbol{x}_{-ij}$ تعریف شد، که با توجه به قضیهٔ ۱.۳.۱ معادل با صفر یا غیر صفر بودن درایه های غیرقطری Q_{ij} است. از این رو گراف $\mathcal{G}$ از مولفه‌های غیر صفر Q ساخته می‌شود. بنابراین از ویژگی‌های قابل توجه و مفید GMRF این است که اطلاعات زیادی در مورد استقلال شرطی از گراف $\mathcal{G}$ قابل استنتاج است. ویژگی‌های مارکوفی موضعی[۱۴]، مارکوفی فراموضعی[۱۵] و مارکوفی دوبه‌دو[۱۶] که برای تعریف گراف به کار می‌رود، برای GMRFها هم ارزند.

---

[۱۴] Local

[۱۵] Global

[۱۶] Pairwise

قضیه ۳.۳.۱ فرض کنید $\boldsymbol{x}$ یک GMRF تحت گراف $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ باشد. آن‌گاه جملات زیر هم‌ارزند.

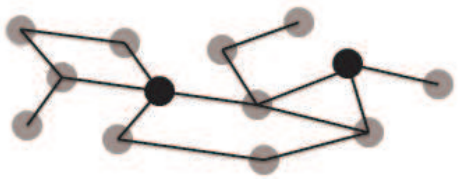
الف) ویژگی مارکوفی دوبه‌دو: اگر $\{i,j\} \notin \mathcal{E}$ و $i \neq j$ آن‌گاه $x_i \perp x_j | \boldsymbol{x}_{-ij}$.

ب) ویژگی مارکوفی موضعی: برای هر $i \in \mathcal{V}$، $x_i \perp \boldsymbol{x}_{-\{i,ne(i)\}} | \boldsymbol{x}_{ne(i)}$.

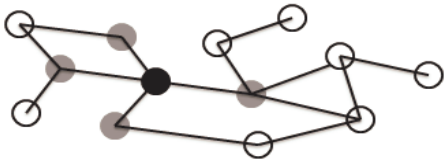
ج) ویژگی مارکوفی فراموضعی: اگر A، B و C مجموعه‌های مجزا و C مجموعه‌ای ناتهی باشد که A و B را از هم جدا می‌کند، آن‌گاه

$$\boldsymbol{x}_A \perp \boldsymbol{x}_B | \boldsymbol{x}_C. \qquad (۷.۳.۱)$$

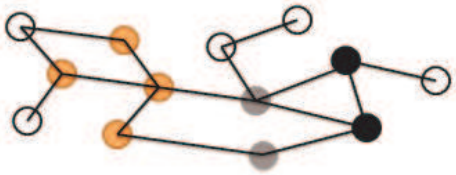
در شکل ۲.۳.۱ دو رأس مشکی به شرط معلوم بودن رأس‌های توسی از هم مستقل هستند. در شکل ۳.۳.۱ رأس مشکی از رأس‌های سفید به شرط معلوم بودن رأس‌های توسی مستقل است. در شکل ۴.۳.۱ رأس‌های مشکی و توسی به شرط معلوم بودن رأس‌های هاشور از هم مستقل هستند. شکل‌های ۲.۳.۱، ۲.۳.۱ و ۳.۳.۱ قضیهٔ ؟؟ را شرح می‌دهد.



شکل ۲.۳.۱: ویژگی مارکوفی دوبه‌دو



شکل ۳.۳.۱: ویژگی مارکوفی موضعی



شکل ۴.۳.۱: ویژگی مارکوفی فراموضعی

۲.۳.۱ ویژگی‌های شرطی میدان تصادفی مارکوفی گاوسی

بای بررسی ویژگی‌های شرطی GMRF، پارامتریدن کانونی که روشی دیگر برای نمایش توزیع نرمال است مفید خواهد بود. توزیع شرطی $\boldsymbol{x}_A$ به شرط $\boldsymbol{x}_B$، که در آن A و $B = -A$ دو مجموعه ناتهی هستند را درنظر بگیرید. در این صورت میانگین و ماتریس دقت بردار $\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_A \\ \boldsymbol{x}_B \end{pmatrix}$ به‌صورت زیر افراز می‌شوند:

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_A \\ \boldsymbol{\mu}_B \end{pmatrix}, \quad \text{و} \quad \boldsymbol{Q} = \begin{pmatrix} \boldsymbol{Q}_{AA} & \boldsymbol{Q}_{AB} \\ \boldsymbol{Q}_{BA} & \boldsymbol{Q}_{BB} \end{pmatrix}. \tag{۸.۳.۱}$$

قضیه ۴.۳.۱ (رو و هلد، ۲۰۰۵): فرض کنید $\boldsymbol{x}$ یک GMRF تحت گراف $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ با میانگین $\boldsymbol{\mu}$ و ماتریس دقت $\boldsymbol{Q} > \mathbf{0}$ باشد. برای دو مجموعهٔ غیرتهی $A \subset \mathcal{V}$ و $B = \mathcal{V} \setminus A$، توزیع شرطی $\boldsymbol{x}_A | \boldsymbol{x}_B$

یک GMRF تحت گراف $\mathcal{G}^A$ بامیانگین $\boldsymbol{\mu}_{A|B}$ و ماتریس دقت $\boldsymbol{Q}_{A|B} > \mathbf{0}$ است، که در آن

$$\begin{aligned} \boldsymbol{\mu}_{A|B} &= \boldsymbol{\mu}_A - \boldsymbol{Q}_{AA}^{-1}\boldsymbol{Q}_{AB}(\boldsymbol{x}_B - \boldsymbol{\mu}_B) \\ \boldsymbol{Q}_{AB} &= \boldsymbol{Q}_{AA}. \end{aligned} \qquad (۹.۳.۱)$$

قضیه۴.۳.۱ که تعمیم قضیهٔ ۲.۳.۱ است، نتایج مفیدی را به دست می‌دهد. اول این‌که $\boldsymbol{Q}_{A|B}$ و در واقع $\boldsymbol{Q}_{AA}$ در دسترس است و نیاز به محاسبه ماتریس دقت شرطی نیست و گراف $\mathcal{G}^A$ تنها با حذف راس‌هایی که در A نیستند و یال‌های متناظر آن‌ها به‌دست می‌آید. دوم این‌که، چون Q_{ij} صفر است مگر این‌که $j \in ne(i)$ باشد، میانگین شرطی تنها به مقدار $\boldsymbol{\mu}$ و $\boldsymbol{Q}$ در $A \cup ne(A)$ وابسته است. بنابراین برای محاسبهٔ میانگین شرطی $\boldsymbol{\mu}_{A|B}$ فقط نیاز به حل دستگاه خطی

$$\boldsymbol{Q}_{AA}(\boldsymbol{\mu}_{A|B} - \boldsymbol{\mu}_A) = -\boldsymbol{Q}_{AB}(\boldsymbol{x}_B - \boldsymbol{\mu}_B)$$

است، و لزوماً نیاز به معکوس کردن $\boldsymbol{Q}_{AA}$ نیست.

تعریف ۲.۳.۱ (پارامتریدن کانونی): یک GMRF تحت گراف $\mathcal{G}$ با تابع چگالی

$$\pi(\boldsymbol{x}) \propto \exp(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x}),$$

فرم کانونی با پارامترهای کانونی $\boldsymbol{b}$ و $\boldsymbol{Q} > \mathbf{0}$ نامیده می‌شود. این فرم دارای ماتریس دقت $\boldsymbol{Q}$ و میانگین $\boldsymbol{\mu} = \boldsymbol{Q}^{-1}\boldsymbol{b}$ است و با نماد $\boldsymbol{x} \sim N_C(\boldsymbol{b}, \boldsymbol{Q})$ نشان داده می‌شود.

رابطه پارامتریدن کانونی با توزیع نرمال به صورت $N(\boldsymbol{\mu}, \boldsymbol{Q}^{-1}) = N_C(\boldsymbol{Q}\boldsymbol{\mu}, \boldsymbol{Q})$ است.

۳.۳.۱ مشخص‌سازی GMRF با شرطی‌های کامل

شیوهٔ دیگر مشخص کردن یک GMRF با میانگین و ماتریس دقت، استفاده از شرطی‌های کامل $\{\pi(x_i|\boldsymbol{x}_{-i})\}$ است. بسیج (۱۹۷۵) در این شیوه پیشگام است و این مدل‌ها را با نام اتو رگرسیو

شرطی[۱۷](CAR) معرفی کرد. در ادامه خصوصیاتی که باید روی شرطی‌های کامل تحمیل شود تا یک GMRF معتبر باشد، بیان می‌شود.

فرض کنید شرطی‌های کامل، توزیع نرمال با میانگین‌ها و دقت‌های شرطی به‌ترتیب به‌صورت

$$E(x_i|\boldsymbol{x}_{-i}) = \mu_i - \Sigma_{j:j\sim i}\beta_{ij}(x_j-\mu_j) \quad i=1,\ldots,n \qquad (10.3.1)$$

$$Prec(x_i|\boldsymbol{x}_{-i}) = \kappa_i > 0 \qquad (11.3.1)$$

برای بعضی $\{\beta_{ij}, i\neq j\}$ و بردارهای $\boldsymbol{\mu}$ و $\boldsymbol{\kappa}$ باشند. با مقایسهٔ عبارات (۴.۳.۱) و (۱۰.۳.۱)، اگر درایه‌های ماتریس $\boldsymbol{Q}$ به صورت

$$Q_{ii}=\kappa_i, \quad \text{و} \quad Q_{ij}=\kappa_i\beta_{ij}$$

انتخاب شوند و ماتریس $\boldsymbol{Q}$ متقارن باشد، یعنی $\kappa_i\beta_{ij}=\kappa_j\beta_{ji}$، آن‌گاه چگالی $\boldsymbol{x}$ موجود و بنابر قضیه بعد منحصربه‌فرد است.

قضیه ۵.۳.۱ (رو و هلد، ۲۰۰۵): فرض کنید n توزیع شرطی کامل نرمال با میانگین‌ها و دقت‌های شرطی (۱۰.۳.۱) و (۱۱.۳.۱) داده شده باشند. آن‌گاه $\boldsymbol{x}$ یک GMRF تحت گراف نشاندار $\mathcal{G}=(\mathcal{V},\mathcal{E})$ با میانگین μ و ماتریس دقت $\boldsymbol{Q}=(Q_{ij})$ با درایه‌های

$$Q_{ij}=\begin{cases}\kappa_i\beta_{ij} & i\neq j\\ \kappa_i & i=j\end{cases}$$

است، به طوری که $\kappa_i\beta_{ij}=\kappa_j\beta_{ji}, i\neq j$ و $\boldsymbol{Q}>\mathbf{0}$.

[۱۷] Conditional Autoregression

۴.۳.۱ میدان تصادفی مارکوفی گاوسی چندمتغیره

فرض کنید $\boldsymbol{x}$ یک GMRF نسبت به گراف $\mathcal{G}$ است. ویژگی مارکوفی بر

$$\pi(x_i|\boldsymbol{x}_{-i}) = \pi(x_i|\{x_j : j \sim i\})$$

دلالت دارد، که در آن x_i مقدار متناظر با رأس i است. اغلب رأس‌ها دارای تعبیر فیزیکی مانند رأس‌های یک مشبکه یاعنصرهای تصویری [۱۸] یک منطقه بر نقشه یک کشور هستند که ممکن است برای تعریف همسایگی رأس i استفاده شوند. برای تعمیم این مسئله، از به هم پیوستن بردارهای p–بعدی $\boldsymbol{x}_i$، $(i = ۱, \ldots, n)$ برای هر یک از n رأس یک GMRF چندمتغیره با اندازهٔ np به‌صورت $\boldsymbol{x} = (\boldsymbol{x}_۱^T, \ldots, \boldsymbol{x}_n^T)^T$ حاصل می‌شود. ویژگی مارکوفی در هر رأس محفوظ است، یعنی $\pi(\boldsymbol{x}_i|\boldsymbol{x}_{-i}) = \pi(\boldsymbol{x}_i|\{\boldsymbol{x}_j : j \sim i\})$، که در آن $\sim$ نسبت به گراف یکسان $\mathcal{G}$ است. فرض کنید $\boldsymbol{\mu} = (\boldsymbol{\mu}_۱^T, \ldots, \boldsymbol{\mu}_n^T)^T$ میانگین و $\widetilde{\boldsymbol{Q}} = (\widetilde{\boldsymbol{Q}}_{ij})$ ماتریس دقت $\boldsymbol{x}$ باشد. هر مولفه $\widetilde{\boldsymbol{Q}}_{ij}$ یک ماتریس $p \times p$ است. با توجه به قضیه ۱.۳.۱ رابطه

$$\boldsymbol{x}_i \perp \boldsymbol{x}_j \mid \boldsymbol{x}_{-ij} \Longleftrightarrow \widetilde{\boldsymbol{Q}}_{ij} = \mathbf{0}$$

به‌دست می‌آید. یک GMRF P–متغیره را می‌توان با بسط تعریف ۱.۳.۱ برای GMRF تعریف نمود.

تعریف ۳.۳.۱ (GMRF چندمتغیره): بردار تصادفی $\boldsymbol{x} = (\boldsymbol{x}_۱^T, \ldots, \boldsymbol{x}_n^T)^T$، که در آن $\dim(\boldsymbol{x}_i) = p$ است، یک GMRF p–متغیره نسبت به گراف $\mathcal{G} = (\mathcal{V} = \{۱, \ldots, n\}, \mathcal{E})$ با میانگین $\boldsymbol{\mu}$ و ماتریس دقت $\widetilde{\boldsymbol{Q}} > \mathbf{0}$ نامیده می‌شود، اگر و تنها اگر چگالی آن به‌صورت

$$\pi(x) \quad = \quad (\frac{۱}{۲\pi})^{-\frac{np}{۲}} |\widetilde{\boldsymbol{Q}}|^{\frac{۱}{۲}} \exp\left(-\frac{۱}{۲}(\boldsymbol{x}-\boldsymbol{\mu})^T \widetilde{\boldsymbol{Q}}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

[۱۸] pixel

$$= (\frac{۱}{۲\pi})^{-\frac{np}{۲}} |\widetilde{\boldsymbol{Q}}|^{\frac{۱}{۲}} \exp\left(-\frac{۱}{۲}\sum_{ij}(\boldsymbol{x}_i - \boldsymbol{\mu}_i)^T \widetilde{\boldsymbol{Q}}_{ij}(\boldsymbol{x}_j - \boldsymbol{\mu}_j)\right)$$

باشد و برای هر $i \neq j$

$$\widetilde{\boldsymbol{Q}}_{ij} \neq \mathbf{o} \Longleftrightarrow \{i, j\} \in E.$$

همهٔ نتایج به‌دست آمده برای GMRF قابل تعمیم برای p GMRF—متغیره است.

۴.۱ الگوریتم‌های محاسباتی و شبیه‌سازی GMRF

در این بخش الگوریتم‌هایی برای محاسبهٔ لگاریتم چگالی، محاسبهٔ چگالی‌های شرطی، شبیه‌سازی از GMRF و شبیه‌سازی شرطی روی زیرمجموعه‌ای از GMRF، قیدهای سخت یا قیدهای نرم و ارزیابی لگاریتم چگالی‌های شرطی متناظر ارائه می‌شود. همهٔ موارد با عملکردهای سادهٔ ماتریسی روی ماتریس دقت $\boldsymbol{Q}$، که تُنک است فرمول‌بندی خواهند شد. به دلیل تُنک بودن ماتریس دقت، ذخیره‌سازی آسان‌تر و محاسبات سریع‌تر صورت خواهد گرفت. به عنوان مثال، در تجزیهٔ چولسکی $\boldsymbol{Q} = \boldsymbol{L}^T\boldsymbol{L}$، ماتریس $\boldsymbol{L}$ پایین مثلثی است که می‌تواند تُنک باشد و شامل درایه‌های غیر صفر ماتریس $\boldsymbol{Q}$ باشد.

در حالت کلی، مقدار flop برای محاسبهٔ تجزیهٔ چولسکی از مرتبهٔ $\mathcal{O}(n^۳)$ است، که در ماتریس‌های تُنک به $\mathcal{O}(n)$ کاهش می‌یابد، این تعداد در الگوهای فضایی از مرتبهٔ $\mathcal{O}(n^{۳/۲})$ و GMRFهای فضایی زمانی $\mathcal{O}(n^۲)$ است. به‌طور مشابه اگر $\boldsymbol{L}$، ماتریس تُنک باشد، حل معادلهٔ $\boldsymbol{L}\boldsymbol{x} = \boldsymbol{b}$ سریع‌تر انجام می‌شود. قبل از بررسی روش‌های عددی برای ماتریس‌های تُنک ابتدا به نحوهٔ شبیه‌سازی و ارزیابی لگاریتم چگالی برای حالت کلی ماتریس $\boldsymbol{Q}$ ارائه می‌شود. فرض کنید $\boldsymbol{A}_{n\times n}$ یک ماتریس متقارن همیشه مثبت باشد، بنابراین تنها یک مثلث چولسکی $\boldsymbol{L}$ وجود دارد که برای هر i، $\boldsymbol{L}_{ii} > ۰$ و $\boldsymbol{A} = \boldsymbol{L}^T\boldsymbol{L}$ است. تعداد flop برای محاسبهٔ $\boldsymbol{L}$ برابر $n^۳/۳$ است. این تجزیه مبنایی

برای حل دستگاه‌های خطی مانند $\boldsymbol{Ax} = \boldsymbol{b}$ و $\boldsymbol{AX} = \boldsymbol{B}$ یا به طور معادل محاسبهٔ $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$ و $\boldsymbol{X} = \boldsymbol{A}^{-1}\boldsymbol{B}$ است.

الگوریتم ۱.۴.۱ (حل معادلهٔ $\boldsymbol{Ax} = \boldsymbol{b}$):

۱) تجزیهٔ چولسکی $\boldsymbol{A} = \boldsymbol{L}^T\boldsymbol{L}$ محاسبه شود.

۲) معادلهٔ $\boldsymbol{L\nu} = \boldsymbol{b}$ برحسب $\boldsymbol{\nu}$ حل شود.

۳) معادلهٔ $\boldsymbol{L}^T\boldsymbol{x} = \boldsymbol{\nu}$ بر حسب $\boldsymbol{x}$ حل شود.

از آن‌جا که

$$\boldsymbol{x} = (\boldsymbol{L}^{-1})^T\boldsymbol{\nu} = (\boldsymbol{L}^{-1})^T(\boldsymbol{L}^{-1}\boldsymbol{b}) = (\boldsymbol{L}\boldsymbol{L}^T)^{-1}\boldsymbol{b} = \boldsymbol{A}^{-1}\boldsymbol{b}$$

پس x پاسخ معادلهٔ $\boldsymbol{Ax} = \boldsymbol{b}$ است. در گام دوم الگوریتم ۱.۴.۱ برای محاسبه $\boldsymbol{\nu}$ از یک حلقهٔ پیشرو به صورت

$$\nu_i = \frac{1}{L_{ii}}(b_i - \sum_{j=1}^{i-1} L_{ij}\nu_j), \qquad i = 1, \ldots, n$$

استفاده می‌شود، که جانشین‌سازی پیشرو[19] نامیده می‌شود. به طور مشابه گام سوم نیز جانشین‌سازی پسرو[20] نامیده می‌شود که در آن x از یک حلقهٔ پسرو به صورت

$$x_i = \frac{1}{L_{ii}}(\nu_i - \sum_{j=i+1}^{n} L_{ji}x_j), \qquad i = n, \ldots, 1 \qquad (۱۲.۴.۱)$$

محاسبه می‌شود. تعداد کل flopهای الگوریتم ۱.۴.۱ برابر n^2 است.

---

[19] Forward Subsitution

[20] Backward Subsitution