به نام خدا

تئوری تخمین (۱-۲۵۱۶۳)

تمرین سری دوم

بهار ۱۴۰۵

دانشکده‌ی مهندسی برق

دانشگاه صنعتی شریف

اساتید: روح الله امیری و محمّدمهدی مجاهدیان

موعد تحویل: جمعه ۲۵ اردیبهشت

۱ آمارگان کافی و کامل

با فرض در اختیار داشتن N نمونه مستقل و هم‌توزیع $X_1, \ldots, X_N$ از هر کدام از توزیع‌های زیر، یک آمارگان کافی مینیمال برای θ پیدا کنید. در هر مورد، بررسی کنید که آیا آمارگان حاصل کامل است؟

آ)

$$f(x;\theta) = \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty$$

ب)

$$f(x;\theta) = e^{-(x-\theta)}, \quad \theta < x < \infty, \quad -\infty < \theta < \infty$$

ج)

$$f(x;\theta) = e^{-(x-\theta)} \exp\{-e^{-(x-\theta)}\}, \quad \theta < x < \infty, \quad -\infty < \theta < \infty$$

د)

$$f(x;\theta) = \frac{e^{-(x-\theta)}}{(1+e^{-(x-\theta)})^2}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty$$

ه)

$$f(x;\theta) = \frac{1}{\pi[1+(x-\theta)^2]}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty$$

و)

$$f(x;\theta) = \frac{1}{2} e^{-|x-\theta|}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty$$

ز)

$$f(x;\theta) = 1, \quad ,\theta < x < \theta + 1, \quad 0 < \theta < \infty$$

ح)

$$P(X = x;\theta) = e^{-\theta} \frac{\theta^x}{x!}, \quad ,x = 0, 1, 2, \ldots, \quad -\infty < \theta < \infty$$

۲ توزیع نرمال چندمتغیره

فرض کنید $\mathbf{x}$ یک بردار تصادفی نرمال N-بعدی با میانگین $\theta\mathbf{1}$ و ماتریس کوواریانس $\mathbf{I}$ باشد؛ یعنی: $\mathbf{x} \sim \mathcal{N}(\theta\mathbf{1}, \mathbf{I})$
بردار $\mathbf{y}$ را به صورت $\mathbf{y} = \mathbf{A}\mathbf{x}$ تعریف می‌کنیم که در آن:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ -1 & & & \\ \vdots & & \mathbf{I} & \\ -1 & & & \end{bmatrix}.$$

آ) نشان دهید که $T(x)$ (اولین مولفه $\mathbf{y}$) دارای توزیع $\mathcal{N}(N\theta, N)$ است.

ب) نشان دهید که $\mathbf{z}$ ($(N-1)$ مولفه آخر $\mathbf{y}$) دارای توزیع $\mathcal{N}(0, \mathbf{Q})$ است. ماتریس $\mathbf{Q}$ را نیز بیابید.

ج) نشان دهید که توزیع شرطی $\mathbf{z}$ به شرط T برابر $\mathcal{N}(0, \mathbf{Q})$ بوده و مستقل از θ است. همچنین، نشان دهید که توزیع شرطی $\mathbf{x}$ به شرط T مستقل از θ است. با توجه به نتایج فوق، در مورد آمارگان T چه می‌توان گفت؟

د) در این قسمت مثالی را می‌بررسی می‌کنیم که در آن تعداد آمارگان‌های کافی کمتر از تعداد پارامترهای تخمین است. فرض کنید $\mathbf{x}$ یک بردار تصادفی نرمال N-بعدی $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$ باشد که در آن

$$\boldsymbol{\mu} = \begin{bmatrix} N\mu \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} N-1+\sigma^2 & -\mathbf{1}^T \\ -\mathbf{1} & \mathbf{I} \end{bmatrix}.$$

به کمک روابط معکوس و دترمینان ماتریس بلوکی، نشان دهید که برای تخمین بردار پارامتر $\boldsymbol{\theta} = [\mu\ \sigma^2]^T$، تابع درستنمایی به صورت زیر قابل ساده‌شدن است:

$$f(\mathbf{x}; \boldsymbol{\theta}) = \frac{1}{(2\pi)^{\frac{N}{2}}\sigma} \exp\left\{-\frac{1}{2}\left[\frac{N^2}{\sigma^2}(\bar{x}-\mu)^2 + \sum_{n=1}^{N-1} x^2[n]\right]\right\}.$$

بر این اساس، آمارگان کافی برای $\boldsymbol{\theta}$ به چه صورت است؟

۳ مدل خطی و تخمین‌گر MVU

آ) در کلاس دیدیم که برای تخمین بردار $\boldsymbol{\theta}$ طبق مدل مشاهدات خطی با نویز سفید گوسی به صورت $\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w}$، $\mathbf{w} \sim \mathcal{N}(0, \sigma^2\mathbf{I})$ کران کرامر-رائو با تخمین‌گر MVU زیر قابل حصول است:

$$\hat{\boldsymbol{\theta}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x}$$

اکنون از طریق پیدا کردن یک آمارگان کافی کامل و اعمال قضیه RBLS سعی کنید به این تخمین‌گر برسید.

ب) در یک مدل خطی، ماتریس مشاهده زیر را در نظر بگیرید:

$$\mathbf{H} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1+\epsilon \end{bmatrix}$$

که در آن ϵ مقدار کوچکی است. $(\mathbf{H}^\mathsf{T}\mathbf{H})^{-1}$ را محاسبه کنید و بررسی کنید که وقتی $\epsilon \to 0$ چه اتفاقی می‌افتد. اگر $\mathbf{x} = [2\ 2\ 2]^\mathsf{T}$ باشد، تخمین‌گر MVU را پیدا کنید و توضیح دهید که وقتی $\epsilon \to 0$ چه اتفاقی می‌افتد؟

ج) در این قسمت، تخمین یک تابع خطی از $\boldsymbol{\theta}$ را بررسی می‌کنیم. فرض کنید پارامتر جدید $\boldsymbol{\alpha} = \mathbf{A}\boldsymbol{\theta}$ باشد، که در آن $\mathbf{A}$ یک ماتریس معلوم $r \times p$ ($r < p$) با رتبه r است. نشان دهید که تخمین‌گر MVU به صورت زیر است:

$$\hat{\boldsymbol{\alpha}} = \mathbf{A}\hat{\boldsymbol{\theta}}$$

که در آن $\hat{\boldsymbol{\theta}}$ تخمین‌گر MVU برای $\boldsymbol{\theta}$ است. همچنین، ماتریس کوواریانس را پیدا کنید.
راهنمایی: $\mathbf{x}$ را با

$$\mathbf{x}' = \mathbf{A}(\mathbf{H}^\mathsf{T}\mathbf{H})^{-1}\mathbf{H}^\mathsf{T}\mathbf{x}$$

جایگزین کنید که در آن $\mathbf{x}'$ دارای ابعاد $r \times 1$ است. می‌توان نشان داد $\mathbf{x}'$ تمام اطلاعات $\boldsymbol{\theta}$ را شامل می‌شود و لذا می‌توانیم تخمین‌گر را بر اساس یک پایگاه داده با ابعاد پایین‌تر به دست آوریم (مدل خطی کاهش یافته).

۴ سیگنال سینوسی

یک سیگنال سینوسی با فرکانس مشخص که با نویز سفید گاوسی ترکیب شده را در نظر بگیرید:

$$x[n] = A\cos(۲\pi f_\circ n) + w[n] \quad n = \circ, ۱, \ldots, N-۱$$

که در آن $w[n]$ نویز سفید گاوسی با واریانس $\sigma^۲$ است. تخمین‌گر MVU را برای موارد زیر پیدا کنید. می‌توانید فرض کنید که آمارگان‌های کافی، شرط کامل بودن را دارند.

آ) دامنه A، با فرض اینکه $\sigma^۲$ معلوم است.

ب) دامنه A و واریانس نویز $\sigma^۲$.

۵ خانواده نمایی

خانواده توزیع نمایی را با تابع چگالی احتمال زیر در نظر بگیرید:

$$f(\mathbf{x};\boldsymbol{\theta}) = c(\boldsymbol{\theta})h(\mathbf{x})\exp\left[\sum_{i=۱}^{m} w_i(\boldsymbol{\theta})T_i(\mathbf{x})\right]$$

آ) بر اساس قضیه تجزیه نیمن فیشر، نشان دهید که $\mathbf{T} = [T_۱(\mathbf{x}), \ldots, T_m(\mathbf{x})]^\mathsf{T}$ یک آمارگان کافی برای $\boldsymbol{\theta}$ است.

ب) مطابق قضیه ارائه‌شده در کلاس، نشان دهید که یک شرط کافی برای اینکه آمارگان فوق کامل باشد، این است که فضای $\{w_۱(\boldsymbol{\theta}), \ldots, w_m(\boldsymbol{\theta})\}$ شامل یک مستطیل mـبعدی باشد.

ج) برای خانواده نمایی نشان دهید که اگر یک آمارگان، کافی و کامل باشد، لزوما یک آمارگان مینیمال است.

د) در کلاس دیدیم که برای تخمین پارامترهای اسکالر، شرط حصول کران کرامرـرائو معادل این است که توزیع مشاهدات از خانواده نمایی باشد و در این صورت تخمین‌گر کارا همان آماره کافی خانواده نمایی است. این گزاره را اثبات نمایید. آیا این نتیجه در حالت تخمین پارامتر برداری نیز برقرار است؟

ه) فرض کنید θ یک پارامتر مثبت و غیرتصادفی باشد. همچنین فرض کنید مشاهدات $X_۱, X_۲, \ldots, X_N$ مستقل و هم‌توزیع با تابع چگالی احتمال زیر هستند:

$$f(x;\theta) = \begin{cases} \frac{x^M e^{-x/۲\theta}}{(۲\theta)^{M+۱}M!}, & x \geq \circ \\ \circ, & x < \circ \end{cases}$$

که در آن M یک عدد صحیح مثبت معلوم است. نشان دهید که این توزیع در خانواده نمایی است. سپس، با محاسبه آمارگان کافی و کامل، تخمین‌گر MVU برای θ را پیدا کنید و واریانس تخمین را محاسبه کنید.

۶ تعمیم قضیه RBLS

در کلاس، قضیه RBLS را برای تخمین پارامتر اسکالر اثبات کردیم. در این تمرین، قصد داریم این قضیه را برای حالت برداری تعمیم دهیم.

آ) اگر $\mathbf{X}$ و $\mathbf{Z}$ دو بردار تصادفی باشند و $\mathbf{g}(\mathbf{z})$ را به صورت $\mathbf{g}(\mathbf{z}) = \mathbb{E}\{\mathbf{Y}|\mathbf{Z}=\mathbf{z}\}$ تعریف کنیم، نشان دهید:

$$\mathrm{Cov}\{\mathbf{g}(\mathbf{Z})\} \preceq \mathrm{Cov}\{\mathbf{Y}\}$$

ب) با استفاده از نتیجه فوق، نشان دهید اگر $\check{\boldsymbol{\theta}}$ یک تخمین‌گر بدون بایاس دلخواه باشد و تعریف کنیم

$$\hat{\boldsymbol{\theta}} = \mathbb{E}\{\check{\boldsymbol{\theta}}|\mathbf{T}(\mathbf{x})\}$$

آنگاه $\hat{\boldsymbol{\theta}}$ تخمین‌گری بدون بایاس است و به ازای هر بردار دلخواه $\mathbf{a}$ خواهیم داشت:

$$\mathbb{E}\{[\mathbf{a}^T(\check{\boldsymbol{\theta}} - \boldsymbol{\theta})]^۲\} \geq \mathbb{E}\{[\mathbf{a}^T(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})]^۲\}$$

تفسیر خود را از این نتیجه بیان کنید.

ج) اکنون نشان دهید اگر $\mathbf{T}(\mathbf{x})$ یک آمارگان کافی و کامل باشد، آنگاه $\hat{\boldsymbol{\theta}}$ تخمین‌گر MVU است.

۷ آمارگان کافی ترتیبی

در کلاس دیدیم که در تخمین‌گرهای ترتیبی، تخمین پارامتر با آمدن داده جدید در طول زمان به روز رسانی می‌شود. در این تمرین، قصد داریم رابطه‌ای ترتیبی برای آمارگان کافی ارائه دهیم.
فرض کنید اندازه‌گیری‌های اسکالر به صورت زیر تولید می‌شوند:

$$x_t = \mathbf{c}_t^\mathsf{T}\boldsymbol{\theta} + n_t \quad (t = ۰, ۱, \ldots)$$

$\boldsymbol{\theta}$ پارامتری ثابت با ابعاد $p \times ۱$ است، اما بردار $\mathbf{c}_t$ با زمان تغییر می‌کند. دنباله‌ای از اندازه‌گیری‌ها را می‌توان در یک مدل آماری خطی بیان کرد:

$$\mathbf{x}_t = \mathbf{H}_t\boldsymbol{\theta} + \mathbf{n}_t, \quad \mathbf{H}_t = \begin{bmatrix} \mathbf{H}_{t-۱} \\ \mathbf{c}_t^\mathsf{T} \end{bmatrix}^\mathsf{T}, \mathbf{x}_t = [x_۰, x_۱, \ldots, x_t]^\mathsf{T}, \mathbf{n}_t = [n_۰, n_۱, \ldots, n_t]^\mathsf{T}$$

فرض کنید بردار نویز $\mathbf{n}_t$ دارای توزیع نرمال چندمتغیره با میانگین صفر و ماتریس کوواریانس $\mathbf{R}_t$ باشد:

$$\mathbf{n}_t \sim \mathcal{N}(۰, \mathbf{R}_t), \mathbf{R}_t = \mathbb{E}\{\mathbf{n}_t\mathbf{n}_t^\mathsf{T}\} = \begin{bmatrix} \mathbf{R}_{t-۱} & \mathbf{r}_t \\ \mathbf{r}_t^\mathsf{T} & r_{tt} \end{bmatrix}, \mathbf{r}_t = \mathbb{E}\{\mathbf{n}_{t-۱}n_t\}, r_{tt} = \mathbb{E}\{n_t^۲\}$$

آ) برای مقدار ثابت t، آمارگان کافی برای $\boldsymbol{\theta}$ برابر است با:

$$\mathbf{T}(\mathbf{x}_t) = \mathbf{H}_t^\mathsf{T}\mathbf{R}_t^{-۱}\mathbf{x}_t$$

ب) سوالی که مطرح می‌شود این است: آیا می‌توان این آمارگان کافی را به صورت ترتیبی محاسبه کرد؟ با نوشتن معکوس $\mathbf{R}_t$ و استفاده از فرمول معکوس ماتریس بلوکی، رابطه‌ای ترتیبی برای آمارگان کافی به شکل زیر به دست آورید.

$$\begin{aligned} \mathbf{T}(\mathbf{x}_t) &= \mathbf{T}(\mathbf{x}_{t-۱}) + \gamma_t^{-۱}[\mathbf{c}_t + \mathbf{H}_{t-۱}^\mathsf{T}\mathbf{b}_t](x_t + \mathbf{b}_t^\mathsf{T}\mathbf{x}_{t-۱}) \\ \mathbf{b}_t &= -\mathbf{R}_{t-۱}^{-۱}\mathbf{r}_t, \quad \gamma_t = r_{tt} + \mathbf{r}_t^\mathsf{T}\mathbf{b}_t \end{aligned}$$

ج) به عنوان حالت خاص، اگر نویزهای اندازه‌گیری مستقل باشند، آنگاه ماتریس کوواریانس $\mathbf{R}_t$ قطری خواهد بود، به این معنی که $\mathbf{r}_t = ۰$ و $\mathbf{b}_t = ۰$.
در این صورت آمارگان کافی را می‌توان به صورت زیر به‌روزرسانی کرد:

$$\mathbf{T}(\mathbf{x}_t) = \mathbf{T}(\mathbf{x}_{t-۱}) + r_{tt}^{-۱}\mathbf{c}_t x_t.$$

در این رابطه ساده، آمارگان کافی تنها با استفاده از اندازه‌گیری اسکالر جدید x_t به‌روزرسانی می‌شود. این رابطه را تعبیر کنید.

۸ سوال مطالعاتی (امتیازی)

در این سوال، قصد داریم با جنبه‌های دیگری از تخمین کلاسیک آشنا شویم. لطفا مقاله زیر را مطالعه کرده و نتایج و قضایای اصلی مطرح‌شده در آن را در حداکثر یک صفحه بیان کنید.

- در این مقاله، نظریه ریاضی فضاهای هیلبرت با هسته بازتولیدشونده (RKHS) برای مسائل تخمین کلاسیک با کمترین واریانس (MVE) به کار رفته است. کران‌های تخمین، آمارگان کافی و تخمین‌گر MVE در این مقاله بر مبنای نظریه فوق بازنویسی شده‌اند.

A. Jung, S. Schmutzhard, and F. Hlawatsch, "The RKHS Approach to Minimum Variance Estimation Revisited: Variance Bounds, Sufficient Statistics, and Exponential Families," IEEE Transactions on Information Theory, 2014.