به نام خدا

تئوری تخمین (۱-۲۵۱۶۳)

تمرین سری اوّل

بهمن ۱۴۰۳

دانشکده‌ی مهندسی برق

دانشگاه صنعتی شریف

اساتید: روح الله امیری و محمّدمهدی مجاهدیان

موعد تحویل: سه‌شنبه ۲۸ اسفند

۱ تخمین‌گر MVU و کران کرامر-رائو

توزیع دوجمله‌ای منفی با پارامتر θ و تابع جرم احتمال زیر را در نظر بگیرید:

$$p\left(x;\theta\right)=\binom{r+x-1}{x}\left(1-\theta\right)^{x}\theta^{r}.$$

آیا میانگین تجربی $\bar{X}_n$ تخمینی MVU از میانگین توزیع هست؟ آیا این تخمین‌گر کارا[1] است؟ کران کرامر-رائو را به دست آورید.

۲ کران کرامر-رائو برای تخمین‌گرهای بایاس‌دار

در کلاس دیدیم که کران کرامر-رائو یک حد پایین برای ماتریس کواریانس تخمین‌گرهای بدون بایاس ارائه می‌دهد. در صورتی‌که بخواهیم این کران را به خانواده تخمین‌گرهای بایاس‌دار با بایاس $\mathbf{b}(\boldsymbol{\theta})$ تعمیم دهیم، خواهیم داشت:

$$\mathbf{C}_{\hat{\boldsymbol{\theta}}}\geq\left(\mathbf{I}+\frac{\partial\mathbf{b}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^{\mathsf{T}}}\right)\mathbf{I}^{-1}(\boldsymbol{\theta})\left(\mathbf{I}+\frac{\partial\mathbf{b}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^{\mathsf{T}}}\right)^{\mathsf{T}}+\mathbf{b}(\boldsymbol{\theta})\mathbf{b}^{\mathsf{T}}(\boldsymbol{\theta}).$$

با روندی مشابه آنچه برای اثبات کران کرامر-رائو در کلاس ارائه شد، این رابطه را اثبات نمایید.

۳ کران کرامر-رائوی مقیّد

با داشتن اطلاعات پیشین، می‌توان روابطی قطعی (و نه احتمالاتی) برای پارامترها در نظر گرفت که به صورت قید در تخمین ظاهر خواهند شد. در کلاس اثبات کردیم که اگر قیود از جنس تساوی باشند، یعنی $\mathbf{f}(\boldsymbol{\theta})=0$، کران کرامر-رائوی مقید از رابطه‌ی زیر به دست می‌آید:

$$\mathrm{CCRLB}(\boldsymbol{\theta})=\mathbf{U}(\mathbf{U}^{\mathsf{T}}\mathbf{J}\mathbf{U})^{-1}\mathbf{U}^{\mathsf{T}}$$

که در آن $\mathbf{J}$ ماتریس اطلاعات فیشر در حالت نامقید است و $\mathbf{U}$ به صورت زیر تعریف می‌شود:

$$\mathbf{F}(\boldsymbol{\theta})=\frac{\partial\mathbf{f}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^{\mathsf{T}}},\ \mathbf{F}(\boldsymbol{\theta})\mathbf{U}=0,\ \mathbf{U}^{\mathsf{T}}\mathbf{U}=\mathbf{I}$$

می‌توان ماتریس $\mathbf{U}$ را با هر ماتریسی مانند $\mathbf{V}$ که شرایط زیر را برآورده کند جایگزین کرد:

$$\mathbf{F}(\boldsymbol{\theta})\mathbf{V}=0,\ \mathrm{rank}(\mathbf{V})=\mathrm{rank}(\mathbf{U})$$

همچنین، دیدیم که اگر $\mathbf{J}$ ماتریسی مثبت معین باشد، می‌توان کران فوق را به صورت زیر ساده کرد:

$$\mathrm{CCRLB}(\boldsymbol{\theta})=\mathbf{J}^{-1}-\mathbf{J}^{-1}\mathbf{F}^{\mathsf{T}}(\mathbf{F}\mathbf{J}^{-1}\mathbf{J}^{\mathsf{T}})^{-1}\mathbf{F}\mathbf{J}^{-1}$$

[1] Efficient

شرط حصول این کران نیز مشابه حالت نامقید است:

$$\frac{\partial \ln f(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathrm{CCRLB}(\boldsymbol{\theta})(\mathbf{g}(\mathbf{x}) - \boldsymbol{\theta})$$

۱.۳

روابط فوق را اثبات کنید. آیا شهودی از تعمیم کران کرامر-رائو به حالت وجود قید نابرابری دارید؟.

۲.۳

حال این مسئله را در نظر بگیرید که N ماهواره در موقعیت‌های $\mathbf{s}_n \in \mathbb{R}^3$ در اختیار داریم که هر کدام به یک سنسور دقیق اندازه‌گیری توان مجهز است.[۲] از طرفی یک کوهنورد دستگاهی در اختیار دارد که با توان مشخص سیگنالی را پخش می‌کند. قصد داریم با استفاده از توان اندازه‌گیری‌شده در سنسورهای ماهواره‌ای موقعیت کوهنورد (که آن را با $\mathbf{x} \in \mathbb{R}^3$ نمایش می‌دهیم) را در هر لحظه از زمان تخمین بزنیم. با فرض مدل انتشاری فضای آزاد، مشاهدات توان قابل تبدیل به مشاهدات فاصله خواهد بود. در صورتی‌که مشاهدات را به صورت زیر مدل کنیم، کران کرامر-رائوی $\mathrm{CRLB}(\mathbf{x})$ را به دست بیاورید:

$$r_n = \|\mathbf{s}_n - \mathbf{x}\|_2 + \omega_n, \;\; n = 1, \ldots, N$$

که ω_nها دارای توزیع گاوسی مستقل با میانگین صفر و واریانس σ^2 هستند.

۳.۳

علاوه بر مشاهدات انجام‌شده، به عنوان اطلاعات پیشین می‌دانیم کوهنورد روی زمین واقع شده است. پس مسئله تخمین ما، به صورت مقید خواهد بود. اگر زمین را به صورت یک کره در نظر بگیریم، موقعیت کوهنورد باید در قید $\|\mathbf{x}\|_2^2 = R_e^2$ صدق کند که R_e شعاع مؤثر زمین است. کران کرامر-رائوی مقید را برای این مسئله را برای $N \geq 3$ و $N = 2$ محاسبه کنید. توجه کنید که مسئله نامقید برای $N = 2$ مشاهده‌پذیر نیست. با توجه به شرط حصول باند کرامر-رائوی مقید، آیا تخمین‌گر کارا وجود دارد؟

۴.۳

قسمت ۳.۳ را برای حالتی‌که توان ارسالی مجهول باشد، تکرار کرده و با نتیجه قبلی مقایسه کنید. در این حالت، مشاهدات به صورت زیر مدل می‌شود:

$$r_n = \alpha\|\mathbf{s}_n - \mathbf{x}\|_2 + \omega_n, \;\; n = 1, \ldots, N$$

که α پارامتری متناسب با جذر توان ارسالی است.

۴ کران کرامر-رائوی بهبودیافته

در برخی مسائل، شرایط حصول باند کرامر-رائو برآورده نمی‌شود و لذا، این باند به اندازه کافی محکم[۳] نخواهد بود (به عنوان مثال در تخمین پارامتر از روی داده‌های نویزی در شرایط SNR پایین). کران‌های متعددی در این شرایط مطرح شده‌اند. در این تمرین، یک کران مفهومی ساده برای تخمین‌های بدون بایاس توسعه می‌دهیم که به کران باتاکاریا[۴] موسوم است.

۱.۴

مسئله تخمین پارامتر اسکالر θ از روی مشاهدات $\mathbf{x}$ را در نظر بگیرید. بردار $\mathbf{z}$ را به صورت زیر تعریف می‌کنیم:

$$\mathbf{z} \triangleq \begin{bmatrix} \hat{\theta} - \theta \\ \frac{1}{f(\mathbf{x};\theta)} \frac{\partial f(\mathbf{x};\theta)}{\partial \theta} \\ \frac{1}{f(\mathbf{x};\theta)} \frac{\partial^2 f(\mathbf{x};\theta)}{\partial \theta^2} \\ \vdots \\ \frac{1}{f(\mathbf{x};\theta)} \frac{\partial^N f(\mathbf{x};\theta)}{\partial \theta^N} \end{bmatrix}$$

نشان دهید که ماتریس کواریانس $\mathbf{z}$ به صورت زیر قابل بیان است:

$$\boldsymbol{\Lambda} \triangleq \mathbb{E}\left(\mathbf{z}\mathbf{z}^{\mathsf{T}}\right) = \begin{bmatrix} \sigma^2_{\hat{\theta}} & 1 & 0 \\ 1 & & \\ & \mathbf{J} & \\ 0 & & \end{bmatrix}$$

و عناصر $\mathbf{J}$ را مشخص کنید.

می‌دانیم ماتریس کواریانس همواره مثبت نیمه‌معین است. با فرض مثبت معین بودن $\mathbf{J}$، آیا $\boldsymbol{\Lambda}$ همواره مثبت معین است؟ (اگر پاسخ منفی است، چه زمانی مثبت معین نخواهد بود؟)

---

[۲] کلیه موقعیت‌ها در یک دستگاه مختصات اینرسی با مرکز زمین بیان شده‌اند.

[۳] Tight

[۴] Bhattacharyya

۲.۴

با استفاده از نتایج قسمت ۱.۴ و اعمال مکمل شور[5] روی $\mathbf{\Lambda}$، یک کران پایین برای $\sigma^2_{\hat{\theta}}$ ارائه دهید. تحت چه شرایطی حالت تساوی برقرار است؟

۳.۴

نشان دهید که برای حالت خاص $N=1$، کران باتاکاریا به کران کرامر-رائو تبدیل می‌شود.

۴.۴

برای حالت $N=2$، نشان دهید:

$$\sigma^2_{\hat{\theta}} \geqslant \frac{1}{J_{11}} + \frac{J_{12}^2}{J_{11}(J_{11}J_{22}-J_{12}^2)}$$

که عبارت دوم نشان‌دهنده میزان بهبود در کران است.

۵.۴

حالتی را در نظر بگیرید که $\mathbf{x}$ شامل M مشاهده مستقل با چگالی‌های یکسان و میانگین‌ها و واریانس‌های شرطی محدود باشد. اگر عناصر $\mathbf{J}$ ناشی از M مشاهده را به صورت $J_{ij}(M)$ نمایش دهیم، نشان دهید که $J_{11}(M)=MJ_{11}(1)$ است و روابط مشابهی برای $J_{12}(M)$ و $J_{22}(M)$ استخراج کنید. نشان دهید که

$$\sigma^2_{\hat{\theta}} \geqslant \frac{1}{MJ_{11}(1)} + \frac{J_{12}^2(1)}{2M^2J_{11}^4(1)} + \mathcal{O}\left(\frac{1}{M^3}\right).$$

۵ کاربرد: تخمین موقعیت کاربر در سیستم‌های مخابراتی

۱.۵

در کلاس دیدیم که ماتریس اطلاعات فیشر برای توزیع برداری گاوسی $\mathbf{x}\sim\mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}),\mathbf{C}(\boldsymbol{\theta}))$ به صورت زیر محاسبه می‌شود:

$$[\mathbf{I}(\boldsymbol{\theta})]_{ij} = \left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right]^{\mathsf{T}}\mathbf{C}^{-1}(\boldsymbol{\theta})\left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right] + \frac{1}{2}\mathrm{tr}\left[\mathbf{C}^{-1}(\boldsymbol{\theta})\frac{\partial\mathbf{C}(\boldsymbol{\theta})}{\partial\theta_i}\mathbf{C}^{-1}(\boldsymbol{\theta})\frac{\partial\mathbf{C}(\boldsymbol{\theta})}{\partial\theta_j}\right]$$

این رابطه را اثبات کرده و آن را به حالت گاوسی مختلط $\mathbf{x}\sim\mathcal{CN}(\boldsymbol{\mu}(\boldsymbol{\theta}),\mathbf{C}(\boldsymbol{\theta}))$ تعمیم دهید.

۲.۵

در این تمرین، قصد داریم مسئله تخمین موقعیت کاربر را در یک سیستم مخابراتی یک ورودی-یک خروجی با سیگنالینگ OFDM بررسی کنیم. در صورتی‌که کاربر ساکن فرض شود، رابطه سیگنال دریافتی در زیرحامل n-ام به‌صورت زیر مدل خواهد شد:

$$r[n] = \underbrace{s[n]\sum_{k=0}^{K}\alpha_k e^{-j2\pi n\Delta_f\tau_k}}_{\boldsymbol{\mu}[n]} + w[n],\quad n\in\{-\frac{N}{2},\ldots,\frac{N}{2}\}$$

که در آن $\Delta_f=\frac{W}{N+1}$ بوده، W پهنای باند OFDM و $s[n]$ پایلوت ارسالی $(\mathbb{E}\{|s[n]|^2\}=E_s=\frac{P}{W})$ می‌باشد. در این رابطه، $K+1$ تعداد مسیرهای دریافت سیگنال (۱ مسیر مستقیم و K مسیر غیرمستقیم)، α_k ضریب مختلط افت کانال و $w[n]$ جمله مربوط به نویز است که در فرم برداری برای کل زیرحامل‌ها به صورت $\mathbf{w}\sim\mathcal{CN}(0,\sigma^2\mathbf{I})$ در نظر گرفته می‌شود. همچنین، تاخیر مسیر مستقیم سیگنال و مسیرهای غیرمستقیم به ترتیب با $\tau_0=\frac{\|\mathbf{x}-\mathbf{x}_0\|_2}{c}$ و $\tau_k=\frac{\|\mathbf{x}_k-\mathbf{x}_0\|_2+\|\mathbf{x}_k-\mathbf{x}\|_2}{c}, k=1,\ldots,K$ بیان می‌شود. ($\mathbf{x}_0$ موقعیت فرستنده، $\mathbf{x}_k$ موقعیت انعکاس‌دهنده k-ام و $\mathbf{x}$ موقعیت کاربر در فضای دوبعدی است.)

با تعریف بردار ضرایب کانال به‌صورت $\boldsymbol{\alpha}=[\alpha_0,\ldots,\alpha_K]^{\mathsf{T}}$، کران کرامر-رائو را با استفاده از روابط قسمت ۱.۵ برای پارامترهای مجهول $\boldsymbol{\theta}=[\mathbf{x}^{\mathsf{T}},\boldsymbol{\alpha}^{\mathsf{T}}]^{\mathsf{T}}$ محاسبه کرده و با استفاده از رابطه معکوس ماتریس‌های بلوکی، کران مربوط به تخمین موقعیت کاربر را به صورت مجزا گزارش نمایید. با توجه به کران حاصل‌شده، اثر پهنای باند و SNR را بر دقت تخمین موقعیت تحلیل نمایید.

توجه: ابتدا کران را برای بردار پارامتر $\boldsymbol{\eta}=[\boldsymbol{\tau}^{\mathsf{T}},\boldsymbol{\alpha}^{\mathsf{T}}]^{\mathsf{T}}$ از رابطه زیر محاسبه کرده و سپس با تعریف ژاکوبین، کران کرامر-رائو را برای $\boldsymbol{\theta}$ به دست آورید.

$$\mathbf{J}(\boldsymbol{\eta}) = \frac{1}{\sigma^2}\sum_{n=-N/2}^{N/2}\Re\left\{\frac{\partial\boldsymbol{\mu}[n]}{\partial\boldsymbol{\eta}^{\mathsf{H}}}\frac{\partial\boldsymbol{\mu}[n]}{\partial\boldsymbol{\eta}}\right\}$$

[5] Schur complement

۳.۵

در صورتی‌که کاربر با فرستنده از نظر زمانی سنکرون نباشد، جملات تاخیر در عبارت سیگنال دریافتی به $\tau_k + b$ تغییر می‌کند که b آفست کلاک بین فرستنده و کاربر است. بر این اساس، پارامترهای مجهول $\boldsymbol{\theta} = [\mathbf{x}^{\mathsf{T}}, b, \boldsymbol{\alpha}^{\mathsf{T}}]^{\mathsf{T}}$ خواهد شد. کران کرامر–رائوی موقعیت کاربر را در این حالت به دست آورده و با نتیجه قسمت ۲.۵ مقایسه کنید.

۴.۵

حال اگر از یک آرایه خطی یکنواخت با M المان و فواصل d در گیرنده (کاربر) استفاده کنیم، رابطه سیگنال دریافتی را به صورت تابعی از تاخیرها (τ_k)، افت مسیرها (α_k) و زوایای ورود سیگنال (φ_k) بازنویسی کرده و کران کرامر–رائوی موقعیت کاربر را به دست آورده و با نتیجه قسمت ۲.۵. با توجه به کران حاصل‌شده، اثر طول آرایه را بر دقت تخمین موقعیت تحلیل نمایید.

توجه: ابتدا کران را برای بردار پارامتر $\boldsymbol{\eta} = [\boldsymbol{\tau}^{\mathsf{T}}, \boldsymbol{\varphi}^{\mathsf{T}}, \boldsymbol{\alpha}^{\mathsf{T}}]^{\mathsf{T}}$ محاسبه کرده و سپس با تعریف ژاکوبین، کران کرامر–رائو را برای $\boldsymbol{\theta}$ به دست آورید.

۵.۵

در مسائل بهینه‌سازی با معیارهای مرتبط به تخمین پارامتر، کران کرامر–رائو می‌تواند یک معیار مناسب باشد. به عنوان نمونه، در تمرین فوق تعیین چیدمان بهینه انعکاس‌دهنده‌های محیطی به منظور حصول بهترین دقت تخمین می‌تواند به به صورت یک مسئله بهینه‌سازی مطرح شود. نمونه‌های دیگری این مسائل شامل تخصیص منابع، تعیین شکل موج ارسالی، شکل‌دهی پرتو می‌شود.

در تخمین پارامترهای برداری، کران کرامر–رائو فرم ماتریسی دارد و برای استفاده در بهینه‌سازی باید از یک معیار اسکالر استفاده کنیم. توابع متداول برای اسکالرسازی، اثر، دترمینان و بزرگترین مقدار ویژه ماتریس کرامر–رائو است. بیان کنید که هر کدام از توابع ذکرشده چه مفهومی دارند؟

۶ کاربرد: تخمین ضرایب فیلتر AR در پردازش صحبت

۱.۶

برای یک فرآیند گاوسی WSS نشان دهید که عناصر ماتریس اطلاعات فیشر در حالت حدی $N \to \infty$ به صورت زیر قابل بیان است:

$$[\mathbf{I}(\boldsymbol{\theta})]_{i,j} = \frac{N}{۲} \int_{-\frac{۱}{۲}}^{\frac{۱}{۲}} \frac{\partial \ln P_{xx}(f;\boldsymbol{\theta})}{\partial \theta_i} \frac{\partial \ln P_{xx}(f;\boldsymbol{\theta})}{\partial \theta_j} df$$

که در آن $P_{xx}(f;\boldsymbol{\theta})$ چگالی طیف توان فرآیند (PSD) با وابستگی صریح به $\boldsymbol{\theta}$ است. فرض بر این است که میانگین $x[n]$ صفر است.

۲.۶

در پردازش گفتار، یک مدل مهم برای تولید گفتار، فرآیند AR است که داده‌ها به عنوان خروجی یک فیلتر گسسته‌ی تمام قطب (all–pole) که در ورودی توسط نویز سفید گاوسی (WGN) تحریک شده است، مدل‌سازی می‌شوند. فیلتر تمام قطب برای مدل‌سازی مجرای صوتی استفاده می‌شود و نویز تحریک، مدل‌سازی فشار هوا از طریق تنگی در گلو را انجام می‌دهد که برای تولید صداهای بی‌واک[۶] لازم است. با استفاده از تکنیک‌های تخمین طیف می‌توان نشان داد که چگالی طیف توان سیگنال از روی مشاهدات به صورت زیر قابل بیان است:

$$P_{xx}(f;\boldsymbol{\theta}) = \frac{\sigma_u^۲}{|A(f)|^۲}$$

که در آن $\boldsymbol{\theta} = [a[۱], a[۲], \dots a[p], \sigma_u^۲]^T$ و $A(f) = ۱ + \sum_{m=۱}^{p} a[m] \exp(-j۲\pi f m)$ است. در این رابطه، $\{a[۱], a[۲], \dots, a[p]\}$ پارامترهای فیلتر AR و $\sigma_u^۲$، واریانس نویز سفید تحریک است.

با استفاده از رابطه چگالی طیف توان و نتیجه ارائه شده در قسمت ۱.۶، باند کرامر–رائو را برای تخمین پارامتر $\boldsymbol{\theta}$ محاسبه نمایید.

[۶] صامت