به نام خدا

تئوری تخمین (۱-۲۵۱۶۳)

تمرین سری پنجم

بهار ۱۴۰۴-۱۴۰۳

دانشکده‌ی مهندسی برق

دانشگاه صنعتی شریف

اساتید: روح الله امیری و محمّدمهدی مجاهدیان

موعد تحویل: شنبه ۲۴ خرداد

۱ فرآیند MA

فرایند گسسته $x(n)$ یک فرایند $MA(2)$ است که به صورت زیر تعریف می‌شود:

$$x(n) = w(n) + 0.8w(n-1) + 1.5w(n-2)$$

که در آن $w(n)$ یک نویز سفید ایستان با میانگین صفر است و تابع خودهمبستگی آن $R_w(m) = \delta(m)$ می‌باشد. ما به اندازه‌گیری‌های نویزی این فرایند دسترسی داریم: $y(n) = x(n) + v(n)$ که در آن $v(n)$ یک فرایند سفید ایستان با میانگین صفر است و تابع خودهمبستگی آن $R_v(m) = 0.1\delta(m)$ می‌باشد. $v(n)$ و $x(n)$ ناهمبسته هستند، یعنی $R_{xv}(m) = 0$.

۱.۱

یک فیلتر بهینه برای تخمین خطی حداقل مربعات خطای میانگین (MMSE) $x(n)$ بر اساس مشاهدات $y(k)$ برای $-\infty < k < \infty$ به دست آورید. خطای مربعات میانگین تخمین را محاسبه کنید.

۲.۱

یک فیلتر بهینه برای تخمین خطی MMSE $x(n)$ بر اساس مشاهدات $y(n-k)$ برای $k \geq 1$ به دست آورید. خطای مربعات میانگین تخمین را محاسبه کنید.

۳.۱

یک فیلتر بهینه برای تخمین خطی MMSE $x(n)$ بر اساس مشاهدات $y(n-k)$ برای $k \geq -1$ به دست آورید. خطای مربعات میانگین تخمین را محاسبه کنید.

۴.۱

معادلات فیلتر کالمن را برای تخمین خطی MMSE $x(n)$ بر اساس مشاهدات $y(n-k)$ برای $1 \leq k \leq n$ تعیین کنید. خطای مربعات میانگین تخمین را به عنوان تابعی از n بیابید. خطای مربعات میانگین را وقتی $n \to \infty$ محاسبه کنید.

۵.۱

معادلات فیلتر کالمن را برای تخمین خطی MMSE $x(n)$ بر اساس مشاهدات $y(n-k)$ برای $-1 \leq k \leq n$ تعیین کنید. خطای مربعات میانگین تخمین را به عنوان تابعی از n بیابید. خطای مربعات میانگین را وقتی $n \to \infty$ محاسبه کنید.

۲ فرآیند AR (اختیاری*)

فرض کنید $x(n)$ یک فرایند $AR(2)$ است که به صورت زیر تعریف می‌شود:

$$x(n) = 0/3x(n-1) + 0/18x(n-2) + w(n)$$

که در آن $w(n)$ یک نویز سفید ایستان با میانگین صفر است و تابع خودهمبستگی آن $R_w(m) = 1/44\delta(m)$ می‌باشد. ما به اندازه‌گیری‌های نویزی این فرایند دسترسی داریم: $y(n) = x(n) + v(n)$ که در آن $v(n)$ یک فرایند سفید ایستان با میانگین صفر است و تابع خودهمبستگی آن $R_v(m) = 0/1\delta(m)$ می‌باشد. $v(n)$ و $x(n)$ ناهمبسته هستند، یعنی $R_{xv}(m) = 0$.

۱.۲

یک فیلتر بهینه برای تخمین خطی MMSE $x(n)$ بر اساس مشاهدات $y(k)$ برای $-\infty < k < \infty$ به دست آورید. خطای مربعات میانگین تخمین را محاسبه کنید.

۲.۲

یک فیلتر بهینه برای تخمین خطی MMSE $x(n)$ بر اساس مشاهدات $y(n-k)$ برای $k \geq 0$ به دست آورید. خطای مربعات میانگین تخمین را محاسبه کنید.

۳.۲

یک فیلتر بهینه برای تخمین خطی MMSE $x(n)$ بر اساس مشاهدات $y(n-k)$ برای $k \geq -2$ به دست آورید. خطای مربعات میانگین تخمین را محاسبه کنید.

۴.۲

معادلات فیلتر کالمن را برای تخمین خطی MMSE $x(n)$ بر اساس مشاهدات $y(n-k)$ برای $0 \leq k \leq n$ تعیین کنید. خطای مربعات میانگین تخمین را به عنوان تابعی از n بیابید. خطای مربعات میانگین را وقتی $n \to \infty$ محاسبه کنید.

۵.۲

معادلات فیلتر کالمن را برای تخمین خطی MMSE $x(n)$ بر اساس مشاهدات $y(n-k)$ برای $-2 \leq k \leq n$ تعیین کنید. خطای مربعات میانگین تخمین را به عنوان تابعی از n بیابید. خطای مربعات میانگین را وقتی $n \to \infty$ محاسبه کنید.

۳ تخمین فرکانس با فیلتر کالمن تعمیم‌یافته

در این مسأله، یک فیلتر کالمن تعمیم‌یافته برای کاربرد ردیابی فرکانس در نظر می‌گیریم. به طور خاص، هدف ما ردیابی فرکانس یک سیگنال سینوسی در حضور نویز است. فرکانس به صورت زیر مدل شده است:

$$f_0[n] = af_0[n-1] + u[n] \quad n \geq 0$$

که در آن $u[n]$ یک نویز گوسی سفید با واریانس σ_u^2 است و $f_0[-1]$ دارای توزیع نرمال $\mathcal{N}(\mu_{f_0}, \sigma_{f_0}^2)$ بوده و مستقل از $u[n]$ است. داده‌های مشاهده‌شده به صورت زیر هستند:

$$x[n] = \cos(2\pi f_0[n]) + w[n] \quad n \geq 0$$

که در آن $w[n]$ یک نویز گوسی سفید با واریانس σ^2 است و مستقل از $u[n]$ و $f_0[-1]$ می‌باشد. معادلات فیلتر کالمن توسعه‌یافته برای این مسأله را بنویسید.

۴ آماره کافی ترتیبی

در کلاس دیدیم که در فیلترهای ترتیبی مانند فیلتر کالمن، تخمین پارامتر با آمدن داده جدید در طول زمان به روز رسانی می‌شود. در این تمرین، قصد داریم رابطه‌ای ترتیبی برای آماره کافی ارائه دهیم.

فرض کنید اندازه‌گیری‌های اسکالر به صورت زیر تولید می‌شوند:

$$x_t = \mathbf{c}_t^\mathsf{T}\boldsymbol{\theta} + n_t \quad (t = 0, 1, \ldots)$$

$\boldsymbol{\theta}$ پارامتری ثابت با ابعاد $p \times 1$ است، اما بردار $\mathbf{c}_t$ با زمان تغییر می‌کند. دنباله‌ای از اندازه‌گیری‌ها را می‌توان در یک مدل آماری خطی بیان کرد:

$$\mathbf{x}_t = \mathbf{H}_t\boldsymbol{\theta} + \mathbf{n}_t, \quad \mathbf{H}_t = \begin{bmatrix} \mathbf{H}_{t-1} \\ \mathbf{c}_t^\mathsf{T} \end{bmatrix}^\mathsf{T}, \mathbf{x}_t = [x_0, x_1, \ldots, x_t]^\mathsf{T}, \mathbf{n}_t = [n_0, n_1, \ldots, n_t]^\mathsf{T}$$

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

$$\mathbf{n}_t \sim \mathcal{N}(0, \mathbf{R}_t), \mathbf{R}_t = \mathbb{E}\{\mathbf{n}_t \mathbf{n}_t^\mathsf{T}\} = \begin{bmatrix} \mathbf{R}_{t-1} & \mathbf{r}_t \\ \mathbf{r}_t^\mathsf{T} & r_{tt} \end{bmatrix}, \mathbf{r}_t = \mathbb{E}\{\mathbf{n}_{t-1} n_t\}, r_{tt} = \mathbb{E}\{n_t^2\}$$

۱.۴

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

$$\mathbf{T}(\mathbf{x}_t) = \mathbf{H}_t^\mathsf{T} \mathbf{R}_t^{-1} \mathbf{x}_t$$

۲.۴

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

$$\begin{aligned} \mathbf{T}(\mathbf{x}_t) &= \mathbf{T}(\mathbf{x}_{t-1}) + \gamma_t^{-1} [\mathbf{c}_t + \mathbf{H}_{t-1}^\mathsf{T} \mathbf{b}_t](x_t + \mathbf{b}_t^\mathsf{T} \mathbf{x}_{t-1}) \\ \mathbf{b}_t &= -\mathbf{R}_{t-1}^{-1} \mathbf{r}_t, \quad \gamma_t = r_{tt} + \mathbf{r}_t^\mathsf{T} \mathbf{b}_t \end{aligned}$$

۳.۴

به عنوان حالت خاص، اگر نویزهای اندازه‌گیری مستقل باشند، آنگاه ماتریس کوواریانس $\mathbf{R}_t$ قطری خواهد بود، به این معنی که $\mathbf{r}_t = 0$ و $\mathbf{b}_t = 0$. در این صورت آماره کافی را می‌توان به صورت زیر به‌روزرسانی کرد:

$$\mathbf{T}(\mathbf{x}_t) = \mathbf{T}(\mathbf{x}_{t-1}) + r_{tt}^{-1} \mathbf{c}_t x_t.$$

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

۵ کران کرامر-رائوی بیزی ترتیبی

در طول درس، با تخمینگرهای ترتیبی مختلفی (مانند فیلتر کالمن) آشنا شدیم. برای ارزیابی عملکرد آن‌ها نیاز است که فرمی ترتیبی برای کران‌های عملکردی پیدا کنیم. در این سوال، این کار را برای کران کرامر-رائوی بیزی انجام خواهیم داد.
بردار حالت $\mathbf{s}_n$ با ابعاد $p \times 1$ را در نظر بگیرید و فرض کنید که n مشاهده در طول زمان داشته باشیم $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$. بنابراین، ماتریس اطلاعات فیشر کل بردارهای حالت $\mathbf{S}_n$ یک ماتریس $np \times np$ خواهد بود. با تجزیه $\mathbf{S}_n$ به صورت $\mathbf{S}_n = \begin{bmatrix} \mathbf{S}_{n-1} \\ \mathbf{s}_n \end{bmatrix}$ که در آن $\mathbf{S}_{n-1} = \mathrm{vec}[\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_{n-1}]$ است، آنگاه فقط به ماتریس $p \times p$ در گوشه پایین-راست، یعنی BCRB_{nn} علاقه‌مندیم. توجه کنید که کران کرامر-رائو بیزی (BCRB) یک ماتریس $np \times np$ با گرفتن امیدریاضی روی تمام دنباله بردارهای حالت $\mathbf{s}_1, \ldots, \mathbf{s}_n$ خواهد بود. تعریف می‌کنیم:

$$\mathbf{J}_n \triangleq \mathrm{BCRB}_{nn}^{-1}$$

۱.۵

نشان دهید که

$$\mathbf{J}_{n+1} = \mathbf{D}_n^{22} - \mathbf{D}_n^{21} \left(\mathbf{J}_n + \mathbf{D}_n^{11}\right)^{-1} \mathbf{D}_n^{12}$$

که در آن:

$$\begin{aligned} \mathbf{D}_n^{11} &= \mathbb{E}_{\mathbf{s}_n, \mathbf{s}_{n+1}} \left[-\Delta_{\mathbf{s}_n}^{\mathbf{s}_n} \ln p(\mathbf{s}_{n+1}|\mathbf{s}_n)\right] \\ \mathbf{D}_n^{12} &= \mathbb{E}_{\mathbf{s}_n, \mathbf{s}_{n+1}} \left[-\Delta_{\mathbf{s}_n}^{\mathbf{s}_{n+1}} \ln p(\mathbf{s}_{n+1}|\mathbf{s}_n)\right], \\ \mathbf{D}_n^{22} &= \mathbb{E}_{\mathbf{s}_n, \mathbf{s}_{n+1}} \left[-\Delta_{\mathbf{s}_{n+1}}^{\mathbf{s}_{n+1}} \ln p(\mathbf{s}_{n+1}|\mathbf{s}_n)\right] + \mathbb{E}_{\mathbf{s}_{n+1}, \mathbf{x}_{n+1}} \left[-\Delta_{\mathbf{s}_{n+1}}^{\mathbf{s}_{n+1}} \ln p(\mathbf{x}_{n+1}|\mathbf{s}_{n+1})\right]. \end{aligned}$$

در رابطه فوق، $\Delta_\beta^\alpha = \nabla_\beta \nabla_\alpha^T$ است که در آن عملگرهای گرادیان به صورت $\nabla_\alpha = \begin{bmatrix} \frac{\partial}{\partial \alpha_1} & \frac{\partial}{\partial \alpha_2} & \cdots & \frac{\partial}{\partial \alpha_p} \end{bmatrix}^T$ تعریف می‌شوند.

۲.۵

هر کدام از جملات $\mathbf{D}_n^{12}$، $\mathbf{D}_n^{11}$ و $\mathbf{D}_n^{22}$ را تعبیر کنید. همچنین، عبارت $\mathbf{J}_{n+1}$ محاسبه‌شده در قسمت قبل را به صورت مجموع جملات مربوط به پیش‌بینی و به‌روزرسانی مرتب کنید.

۳.۵

اکنون یک مدل خطی گاوسی برای حالت و مشاهدات مطابق زیر در نظر بگیرید:

$$\begin{aligned}\mathbf{s}_{n+1} &= \mathbf{F}_n\mathbf{s}_n + \mathbf{v}_n,\\ \mathbf{x}_n &= \mathbf{C}_n\mathbf{s}_n + \mathbf{w}_n,\end{aligned}$$

که در آن $\mathbf{v}_n \sim \mathcal{N}(0, \mathbf{Q}_n)$ و $\mathbf{w}_n \sim \mathcal{N}(0, \mathbf{R}_n)$ است. بر اساس، توزیع‌های ذکرشده و استفاده از لم معکوس ماتریس، نشان دهید که رابطه ترتیبی ماتریس اطلاعات فیشر به صورت زیر قابل محاسبه است:

$$\begin{aligned}\mathbf{J}_{n+1} &= \mathbf{Q}_n^{-1} - \mathbf{Q}_n^{-1}\mathbf{F}_n\left[\mathbf{J}_n + \mathbf{F}_n^T\mathbf{Q}_n^{-1}\mathbf{F}_n\right]^{-1}\mathbf{F}_n^T\mathbf{Q}_n^{-1} + \mathbf{C}_{n+1}^{-1}\mathbf{R}_{n+1}^{-1}\mathbf{C}_{n+1}\\ &= \left[\mathbf{Q}_n + \mathbf{F}_n\mathbf{J}_n^{-1}\mathbf{F}_n^T\right]^{-1} + \mathbf{C}_{n+1}^{-1}\mathbf{R}_{n+1}^{-1}\mathbf{C}_{n+1}\end{aligned}$$

که در آن عبارت اول پیش‌بینی فرآیند و عبارت دوم به‌روزرسانی اندازه‌گیری است. این رابطه را با MSE فیلتر کالمن مقایسه کنید. چه نتیجه‌ای می‌گیرید.

۶ نمونه‌برداری

متغیر تصادفی $X \sim \mathcal{N}(0, 1)$ را در نظر بگیرید. می‌خواهیم احتمال رخداد نادر زیر را تخمین بزنیم:

$$P(X > 4) = \mathbb{E}[\mathbb{I}_{X>4}]$$

که در آن $\mathbb{I}_{X>4}$ تابع نشانگر است.

۱.۶ روش مونت‌کارلوی ساده

با تولید $N = 10^6$ نمونه از توزیع $\mathcal{N}(0, 1)$، مقدار $P(X > 4)$ را تخمین زده و میانگین تجربی و واریانس تخمین را محاسبه کنید. چرا این روش برای تخمین این احتمال ناکارآمد است؟

۲.۶ نمونه‌برداری اهمیت

برای بهبود کارایی، از توزیع $\mathcal{N}(\mu, 1)$ استفاده می‌کنیم. تخمین‌گر نمونه‌برداری اهمیت را به صورت زیر بدست آورید:

$$\hat{P} = \frac{1}{N}\sum_{i=1}^{N}\mathbb{I}_{Y_i>4}\cdot\frac{\phi(Y_i)}{\phi(Y_i;\mu,1)}$$

که در آن $\phi(x)$ تابع چگالی احتمال نرمال استاندارد است.

این تخمین‌گر را برای $\mu = 4$ و $N = 10^4$ پیاده‌سازی کرده و واریانس آن را با روش مونت‌کارلوی ساده مقایسه کنید. چرا نمونه‌برداری اهمیت واریانس را کاهش می‌دهد؟ مقدار بهینه‌ی μ که واریانس را کمینه می‌کند محاسبه کنید.

۳.۶ فیلتر ذره‌ای

در فیلتر ذره‌ای برای سیستم‌های دینامیکی، از نمونه‌برداری اهمیت برای تخمین توزیع حالت استفاده می‌شود. اگر مدل حالت به صورت زیر باشد:

$$\begin{aligned}\mathbf{s}_n &= \mathbf{f}(\mathbf{s}_{n-1}) + \mathbf{v}_n\\ \mathbf{x}_n &= \mathbf{h}(\mathbf{s}_n) + \mathbf{w}_n\end{aligned}$$

که در آن $\mathbf{w}_n$ و $\mathbf{v}_n$ نویز هستند، نشان دهید چگونه می‌توان از نمونه‌برداری اهمیت برای تخمین $p(\mathbf{s}_n|\mathbf{x}_{1:n})$ استفاده کرد.