به نام خدا

بهینه‌سازی محدّب ۱ (۱-۲۵۷۵۶)

تمرین کامپیوتری سری دوم

ترم بهار ۰۲-۱۴۰۱

دانشکده‌ی مهندسی برق

دانشگاه صنعتی شریف

استاد: دکتر محمّدحسین یاسائی میبدی

مهلت تحویل: جمعه ۲۹ اردیبهشت ساعت ۲۳:۵۹

(*) مسائلی که با ستاره مشخّص شده‌اند امتیازی هستند و حل‌کردن آن‌ها نمره‌ی امتیازی خواهد داشت!

۱ برآورد درست‌نمایی بیشینه

تعداد دفعاتی که نوعی رویداد در هر ساعت از روز رخ می‌دهد یک متغیر تصادفی است و ما آن را با متغیرهای مستقل پواسون مدل می‌کنیم. که در آن احتمال k بار رخ دادن یک رویداد به صورت زیر مشخص می‌شود:

$$\mathbb{P}[N_t = k] = e^{-\lambda_t}\frac{\lambda_t^k}{k!}, \quad k = 0, 1, \ldots .$$

که در آن پارامتر $\lambda_t \geq 0$ و $t = 1, \ldots, 24$ (اگر $\lambda_t = 0$ باشد، به این معناست که $k = 0$ رویداد با احتمال ۱ رخ داده‌است.) در اینجا t نشان‌دهنده‌ی ساعت است. در واقع $t = 1$ نشان‌دهنده‌ی زمان‌های بین نیمه‌شب و ۱ بامداد، و $t = 24$ نشان‌دهنده‌ی زمان‌های بین ۱۱ شب و نیمه‌شب است (توزیع پواسون متناوب). پارامتر λ_t نیز امید ریاضی تعداد رخدادهای ساعت t است؛ به عبارتی دیگر می‌توان آن را به عنوان نرخ وقوع رخدادها در ساعت t در نظر گرفت. در طول یک روز، رویدادهای $N_1, \ldots, N_{24}$ را مشاهده می‌کنیم.

۱.۱ بخش اول

تخمین بیشترین درست‌نمایی (maximum likelihood estimate) برای پارامترهای $\lambda_1, \ldots, \lambda_{24}$ چیست؟
(راهنمایی: یک راه‌حل تحلیلی ساده وجود دارد. شما باید موارد $N_t \geq 0$ و $N_t = 0$ را جداگانه درنظر بگیرید.)

۲.۱ بخش دوم

در بسیاری از کاربردها، منطقی است که فرض کنیم λ_t به آرامی در طول روز تغییر می کند. برای مثال، میزان وقوع رویدادها بین ساعت ۱۵ و ۱۶ خیلی متفاوت از میزان وقوع بین ساعت ۱۶ تا ۱۷ نیست. برای به دست آوردن یک تخمین هموار از λ_t، لگاریتم درست‌نمایی منهای جمله‌ی هموارسازی (regularization) زیر را بیشینه می‌کنیم:

$$\rho\left(\sum_{t=1}^{23}(\lambda_{t+1} - \lambda_t)^2 + (\lambda_1 - \lambda_{24})^2\right)$$

که در آن $\rho \geq 0$. توضیح دهید که چگونه می‌توان مقادیر $\lambda_1, \ldots, \lambda_{24}$ را به کمک بهینه‌سازی محدب یافت.

۳.۱ بخش سوم

اگر $\rho \to \infty$ چه اتفاقی می‌افتد؟ جواب خود را به صورت تحلیلی و کوتاه بیان کنید.

۴.۱ بخش چهارم

در طول یک روز، وقوع رخدادها را به صورت زیر مشاهده می‌کنیم:

$$N = (0, 4, 2, 2, 3, 0, 4, 5, 6, 6, 4, 1, 4, 4, 0, 1, 3, 4, 2, 0, 3, 2, 0, 1)$$

پارامترهای بهینه مسئله‌ی بهینه‌سازی که در بخش دوم نوشته‌اید را به کمک CVXPY برای $\rho \in \{0/1, 1, 10, 100\}$ بیابید. برای هر ρ نمودار λ_t بر حسب t را نیز رسم کنید. (برای آشنایی با CVXPY و حل مسائل بهینه‌سازی به کمک آن، می‌توانید مثال‌های این صفحه را مطالعه کنید)

۵.۱ بخش پنجم (*)

یک راه برای انتخاب مقدار پارامتر ρ ، بررسی مدل‌های بخش چهارم است. در واقع باید دید که کدام‌یک از آن مدل‌ها، لگاریتم درست‌نمایی بیشتری روی داده‌ی تست خواهد داشت (داده‌ی تست، می‌تواند داده‌ی یک روز دیگر باشد که در ساخت مدل استفاده نشده است). برای هر کدام از مقادیر ρ بخش چهارم لگاریتم بیشینه‌نمایی را برای داده‌ی زیر به دست بیاورید. کدام پارامتر ρ بهتر است؟

$$N^{\text{test}} = (0, 1, 3, 2, 3, 1, 4, 5, 3, 1, 4, 3, 5, 5, 2, 1, 1, 1, 2, 0, 1, 2, 1, 0)$$

۲ سطوح فعالیت بهینه

انتخاب n سطح فعالیت غیرمنفی را با $x_1, \ldots, x_n$ نشان می‌دهیم. این فعالیت‌ها از m منبع استفاده می‌کنند که این منابع محدود هستند. فعالیت j میزان $A_{ij}x_j$ از منبع i استفاده می‌کند (A_{ij} از داده‌های مسئله است). مصرف منابع نیز به صورت افزایشی است. در نتیجه مصرف کل از منبع i به صورت $c_i = \sum_{j=1}^{n} A_{ij}x_j$ محاسبه می‌شود (معمولاً A_{ij} مقداری نامنفی است و به عبارتی، فعالیت j منبع i را مصرف می‌کند. اما این امکان را نیز می‌دهیم که A_{ij} منفی باشد و در واقع فعالیت j، منبع i را به عنوان محصول فرعی تولید کند). مصرف هر منبع محدود است و داریم $c_i \le c_i^{\max}$ که مقدار $c_i^{\max}$ داده شده است. هر فعالیت یک درآمد ایجاد می‌کند، که تابعی تکه‌ای و مقعر از سطح فعالیت است:

$$r_j(x_j) = \begin{cases} p_j x_j & 0 \le x_j \le q_j \\ p_j q_j + p_j^{\text{disc}}(x_j - q_j) & x_j \ge q_j \end{cases}$$

در اینجا $p_j > 0$ قیمت پایه، $q_j > 0$ سطح تخفیف و p_j^{disc} قیمت تخفیف‌خورده برای محصول فعالیت j هستند. (داریم $0 < p_j^{\text{disc}} < p_j$). درآمد کل، مجموع درآمدهای مرتبط با هر فعالیت است یا به عبارتی $\sum_{j=1}^{n} r_j(x_j)$. هدف این است که سطوح فعالیت را به گونه‌ای انتخاب کنیم تا با رعایت محدودیت‌های منابع، درآمد کل را بیشینه کنیم.

۱.۲ بخش اول

نشان دهید که چگونه می‌توان این مسئله را به صورت یک مسئله‌ی LP فرمول‌بندی کرد.

۲.۲ بخش دوم

این مسئله را به کمک CVXPY برای داده‌های زیر حل کنید:

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 3 & 1 \\ 0 & 3 & 1 & 1 \\ 2 & 1 & 2 & 5 \\ 1 & 0 & 3 & 2 \end{bmatrix}, \quad c_{\max} = \begin{bmatrix} 100 \\ 100 \\ 100 \\ 100 \\ 100 \end{bmatrix}, \quad p = \begin{bmatrix} 3 \\ 2 \\ 7 \\ 6 \end{bmatrix}, \quad q = \begin{bmatrix} 4 \\ 10 \\ 5 \\ 10 \end{bmatrix}, \quad p^{\text{disc}} = \begin{bmatrix} 2 \\ 1 \\ 4 \\ 2 \end{bmatrix}$$

سطوح فعالیت بهینه، درآمد تولیدشده از هر کدام از آنها، و درآمد کل پاسخ بهینه را به دست بیاورید. همچنین میانگین قیمت در یک واحد فعالیت را برای هر کدام از سطوح به دست بیاورید (این مقدار در واقع نسبت درآمد مربوط به هر فعالیت، به سطح آن فعالیت است. این مقادیر برای هر فعالیت باید بین قیمت پایه، و قیمت تخفیف‌خورده باشد)

۳ برنامه ریزی بهینه سرعت وسیله نقلیه

یک وسیله نقلیه (برای مثال یک هواپیما) در امتداد یک مسیر مشخص متشکل از n بخش، بین $n+۱$ نقطه با برچسب‌های $۰,\dots,n$ حرکت می‌کند. بخش iام از نقطه $i-۱$ شروع می‌شود و در نقطه‌ی i پایان می‌یابد. این وسیله نقلیه در زمان $t=۰$ از نقطه‌ی صفر شروع به حرکت می‌کند و هر بخش را با یک سرعت ثابت و غیرمنفی طی می‌کند. s_i سرعت در بخش iام است. همجنین در هر بخش، یک حد بالا و حد پایین روی سرعت‌ها داریم و به عبارتی $s_i^{\min} \le s_i \le s_i^{\max}$. وسیله در نقاط بین بخش‌ها توقف نمی‌کند و به راحتی به بخش بعد می‌رود. مسافت طی‌شده در بخش iام برابر d_i و یک مقدار مثبت است. در نتیجه زمان طی‌شده در بخش iام برابر با d_i/s_i است. فرض کنید $i=۱,\dots,n$ و همچنین τ_i زمانی باشد که وسیله به نقطه i می‌رسد. وسیله باید بین زمان‌های داده‌شده $\tau_i^{\min}$ و $\tau_i^{\max}$ به نقطه‌ی i برسد. وسیله در بخش iام با نرخی که به سرعت آن بستگی دارد، مقدار $\Phi(s_i)$ سوخت مصرف می‌کند که تابع Φ یک تابع مثبت، صعودی و محدب است و واحد آن نیز $\mathrm{kg/s}$ می‌باشد.

داده‌های d (مسافت طی کردن هر بخش)، $s_i^{\min}$ و $s_i^{\max}$ (محدوده سرعت در هر بخش)، $\tau_i^{\min}$ و $\tau_i^{\max}$ (محدوده زمانی رسیدن به نقاط بین بخش‌ها) و تابع مصرف سوخت $\Phi:\mathbb{R}\to\mathbb{R}$ در اختیار شما قرار داده می‌شود. از شما می‌خواهیم تا سرعت‌های $s_۱,\dots,s_n$ را به گونه‌ای انتخاب کنید تا مجموع سوخت مصرفی (بر حسب kg) را کمینه کند.

۱.۳ بخش اول

نشان دهید که چگونه می‌توان مسئله‌ی گفته‌شده را به صورت یک مسئله‌ی بهینه‌سازی محدب مطرح کرد. اگر متغیرهای جدیدی معرفی می‌کنید یا متغیرها را تغییر می‌دهید، باید توضیح دهید که چگونه سرعت‌های بهینه را از حل مسئله خود بازیابی می‌کنید. اگر تحدب تابع هدف و یا هر کدام از قیود در فرمول‌بندی شما واضح نیست، دلیل محدب بودن آن را بیان کنید.

۲.۳ بخش دوم

روش گفته‌شده در بخش قبل را روی یک نمونه از مسئله با داده‌های موجود در `veh_speed_sched_data.py` اجرا کنید. از تابع مصرف سوخت $\Phi(si)=as_i^۲+bs_i+c$ (پارامترهای a، b و c در فایل داده‌های گفته‌شده تعریف شده است) استفاده کنید. مصرف بهینه‌ی سوخت چقدر است؟ نمودار سرعت بهینه بر حسب بخش‌های موجود در مسیر را رسم کنید. می‌توانید از تابع `step` در `matplotlib` برای بهتر نشان دادن مقدار ثابت سرعت در هر بخش مسیر استفاده کنید.

۴ بازنویسی قیود

در انتهای این بخش، چندین قید محدب را برای متغیرهای اسکالر x، y و z داده شده است. هر یک را به صورت مجموعه‌ای از قیود معتبر در CVX بیان کنید. (بیان مستقیم آنها در CVX منجر به خطا خواهد شد) همچنین می توانید در صورت نیاز متغیرهای افزوده‌ای را معرفی کنید. فرمول‌بندی جدید خود را با ایجاد یک مسئله‌ی جدید با قیود گفته‌شده در CVX و حل آن بررسی کنید. این مسئله لزوما نباید feasible باشد و تنها کافی است بررسی کنید که CVX قیود شما را بدون خطا پردازش می‌کند.

- $1/x+1/y \le 1, x \ge 0, y \ge 0$
- $xy \ge 1, x \ge 0, y \ge 0$
- $(x+y)^2/\sqrt{y} \le x-y+5, y \ge 0$
- $x+z \le 1+\sqrt{xy-z^2}, x \ge 0, y \ge 0$