به نام خدا

بهینه‌سازی محدّب ۱ (۱-۲۵۷۵۶)

تمرین کامپیوتری سری سوم

ترم بهار ۰۲-۱۴۰۱

دانشکده‌ی مهندسی برق

دانشگاه صنعتی شریف

استاد: دکتر محمّدحسین یاسائی میبدی

مهلت تحویل: جمعه ۹ تیر ۱۴۰۲ ساعت ۲۳:۵۹

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

۱ برنامه‌ریزی برای تخلیه‌ی بهینه

در مسئله‌ی تخلیه‌ی بهینه‌ی مردم از ناحیه‌ی خطر، هدف کمینه کردن مجموع ریسک کل است. محدوده‌ی موردنظر را به شکل گرافی همبند با n گره و m یال مدل می‌کنیم، به طوری که مردم می‌توانند در گره‌ها بمانند و یا در هر جهتی روی یال‌ها حرکت کنند. بردار $\mathbf{q}_t \in \mathbb{R}_+^n$ را بردار تعداد نفرات در هر راس در لحظه‌ی t در نظر می‌گیریم که $t \in \{1, \ldots, T\}$. همچنین توجه داریم که بردار $\mathbf{q}_t$ متشکل از مقادیر حقیقی است. گره‌های گراف محدودیت ظرفیت دارند، به طوری که قید $\mathbf{q}_t \leq \mathbf{Q}$ برقرار است که $\mathbf{Q} \in \mathbb{R}_+^n$ بردار حداکثر ظرفیت است. برای توصیف اتصالات گراف از ماتریس $\mathbf{A} \in \mathbb{R}^{n\times m}$ استفاده می‌کنیم که اعضای آن به شکل زیر تعریف می‌شوند:

$$A_{ij} = \begin{cases} +1 & \text{if edge } j \text{ enters node } i \\ -1 & \text{if edge } j \text{ exits node } i \\ 0 & \text{otherwise} \end{cases}$$

توزیع جمعیت در هر لحظه مانند t از معادله‌ی $\mathbf{q}_{t+1} = \mathbf{A} \cdot \mathbf{f}_t + \mathbf{q}_t$ به دست می‌آید که بردار $\mathbf{f}_t \in \mathbb{R}^m$ حرکت جمعیت روی یال‌ها را نشان می‌دهد و علامت مقدار مربوط به هر یال نشان‌دهنده‌ی جهت آن است.

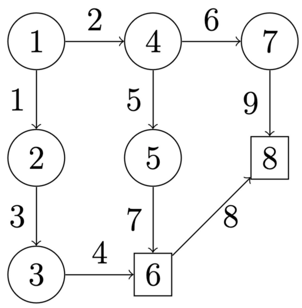
هر یال ظرفیتی محدود دارد به طوری که $|\mathbf{f}_t| \leq \mathbf{F}$ و برقرار است و بردار $\mathbf{F} \in \mathbb{R}_+^m$ بردار حداکثر ظرفیت است.

یک برنامه‌ی تخلیه دنباله‌ای از $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_T$ و $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_{T-1}$ است که از قیود داده شده تبعیت می‌کنند. ریسک کل یک برنامه از رابطه‌ی زیر به دست می‌آید:

$$R_{\text{tot}} = \sum_{t=1}^{T} \left(\mathbf{r}^\top \mathbf{q}_t + \mathbf{q}_t^\top \operatorname{diag}(\mathbf{s})\mathbf{q}_t\right) + \sum_{t=1}^{T-1} \left(\tilde{\mathbf{r}}^\top |\mathbf{f}_t| + \mathbf{f}_t^\top \operatorname{diag}(\tilde{\mathbf{s}})\mathbf{f}_t\right) \tag{1}$$

که در آن $\mathbf{r}, \mathbf{s} \in \mathbb{R}_+^n$ ضرایب ریسک مربوط به گره‌ها و $\tilde{\mathbf{r}}, \tilde{\mathbf{s}} \in \mathbb{R}_+^m$ ضرایب مربوط به ریسک یال‌ها است. در مدل‌سازی ریسک داده شده ضریب بخش خطی $(\mathbf{r}, \tilde{\mathbf{r}})$ را می‌توان ریسک به ازای هر نفر و ضریب بخش متناسب با مربع جمعیت $(\mathbf{s}, \tilde{\mathbf{s}})$ را می‌توان ریسک مربوط به شلوغی ناحیه در نظر گرفت. زیرمجموعه‌ای از گره‌های گراف امن هستند، به این معنی که ضرایب مربوط به ریسک آن‌ها برابر صفر است. همچنین گراف در لحظه‌ی t در صورتی تخلیه شده است که $\mathbf{r}^\top \mathbf{q}_t + \mathbf{q}_t^\top \operatorname{diag}(\mathbf{s})\mathbf{q}_t = 0$ باشد (به عبارت دیگر تمام جمعیت در گره‌های امن باشند).

با فرض این که T به اندازه کافی بزرگ است و مجموع ظرفیت گره‌های امن حداقل به اندازه‌ی کل جمعیت است، برنامه‌ی بهینه (که مجموع ریسک کل را کمینه می‌کند) را برای داده‌های `opt_evac_data.py` به دست آورید (گراف مربوط به ناحیه‌ی داده شده در شکل ۱ رسم شده است). مقدار مجموع ریسک کل برنامه را گزارش کنید. اگر تخلیه‌ی کامل را به طور تقریبی با برقراری معادله‌ی $\mathbf{r}^\top \mathbf{q}_t + \mathbf{q}_t^\top \operatorname{diag}(\mathbf{s})\mathbf{q}_t \leq 10^{-4}$ بررسی کنیم، در چه زمانی در برنامه بهینه تخلیه نهایی صورت می‌گیرد؟ نمودار مقادیر جمعیت در هر گره، جمعیتی که از هر یال عبور می‌کنند و ریسک لحظه‌ای را بر حسب تمام لحظات $t = 1, \ldots, T$ رسم کنید.



شکل ۱: گراف مربوط به منطقه داده شده

۲ طراحی مدار بهینه

در طراحی مدار، برای قرار دادن عرض مجموعه‌ای از اجزای مدار که با بردار $\mathbf{w} = (w_1, \ldots, w_n)$ نمایش داده می‌شوند، باید آنها را در یک بازه‌ی مشخص قرار دهیم. به عبارت دیگر، داریم:

$$W_{\min} \leq w_i \leq W_{\max}, \qquad i = 1, \ldots, n$$

به طوری که $W_{\min}$ و $W_{\max}$ اعداد مثبت داده شده هستند. هر طراحی با سه معیار توان مدار ($P(\mathbf{w})$)، تاخیر مدار ($D(\mathbf{w})$)، و مساحت مدار ($A(\mathbf{w})$) سنجیده می‌شود که علاقه‌مند هستیم هر سه را کوچک نگه داریم.

این سه معیار توابع چندجمله‌ای پیچیده‌ای بر حسب $\mathbf{w}$ هستند و هیچ اطلاعاتی در مورد ضرایب یا توان جملات این سه تابع نداریم (یک تابع چندجمله‌ای به صورت:

$$f(\mathbf{w}) = \sum_m c_m \cdot w_1^{a_m^{(1)}} \cdot w_2^{a_m^{(2)}} \cdot \ldots \cdot w_n^{a_m^{(n)}} = \sum_m c_m \prod_{i=1}^{n} w_i^{a_m^{(i)}}$$

تعریف می‌شود).

اطلاعات چند طراحی قبلی که در شرط مربوط به عرض‌ها صدق می‌کنند را داریم، یعنی بردارهای $\{\tilde{\mathbf{w}}^{(j)}\}_{j=1}^{K}$ که به ازای هر $j \in [K]$ می‌دانیم $\tilde{\mathbf{w}}^{(j)} \in \mathbb{R}^n$ و مقادیر توان، تأخیر و مساحت متناظر با هر طرّاحی یعنی $\left\{ \left(P(\tilde{\mathbf{w}}^{(j)}), D(\tilde{\mathbf{w}}^{(j)}), A(\tilde{\mathbf{w}}^{(j)}) \right) \right\}_{j=1}^{K}$ را داریم. هدف ما این است که $\mathbf{w}$ را به گونه‌ای طراحی کنیم که علاوه بر شروط مربوط به عرض، در شروط زیر نیز صدق کند:

$$P(\mathbf{w}) \leq P_{\text{spec}}, \quad D(\mathbf{w}) \leq D_{\text{spec}}, \quad A(\mathbf{w}) \leq A_{\text{spec}}$$

که سه مقدار $(P_{\text{spec}}, D_{\text{spec}}, A_{\text{spec}})$ داده شده‌اند. به دادگان موجود در فایل `blend_design_data.py` توجه کنید. بردار $\mathbf{w}$ را به گونه‌ای تعیین کنید که تمام قیود گفته شده را ارضا کند. توضیح دهید چگونه با وجود اینکه اطلاعات خیلی محدودی از توابع $P(\mathbf{w}), D(\mathbf{w}), A(\mathbf{w})$ داریم، مطمئن هستیم که بردار $\mathbf{w}$ در تمام قیود صدق می‌کند. روش خود را پیاده‌سازی کنید.

راهنمایی: از نامساوی جنسن-کوشی استفاده کنید.

۳ $(*)$ پر کردن ماتریس کوواریانس

در این مسئله به دنبال به دست آوردن تخمین ماتریس کواریانس یک سری متغیر تصادفی از روی تخمین‌های دو زیرماتریس متقاطع از آن هستیم که این دو زیرماتریس می‌توانند در درایه‌های مشترک ناسازگار هم باشند. به طور دقیق‌تر، فرض کنید $n = n_1 + n_2 + n_3$ متغیر تصادفی با ماتریس کواریانس ناشناخته‌ی $\mathbf{C}$ داریم که می‌توان به صورت زیر آن را بلوک‌بندی کرد:

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{12}^{\top} & \mathbf{C}_{22} & \mathbf{C}_{23} \\ \mathbf{C}_{13}^{\top} & \mathbf{C}_{23}^{\top} & \mathbf{C}_{33} \end{bmatrix}$$

که در آن $\mathbf{C}_{11}, \mathbf{C}_{22}, \mathbf{C}_{33}$ ماتریس‌های کواریانس مربع با بعد n_1, n_2, n_3 هستند و بعد باقی زیرماتریس‌ها هم متقابلاً از روی بعدهای داده شده قابل محاسبه است.

تخمین‌های $\mathbf{S}, \mathbf{T}$ از ماتریس‌های کواریانس متناظر با $n_1 + n_2$ متغیّر تصادفی اول و $n_2 + n_3$ متغیّر تصادفی آخر در اختیار ما قرار گرفته و هدف ما تعیین C است که با این ماتریس‌ها همخوانی داشته باشد. به زبان ریاضی می‌خواهیم:

$$\mathbf{C}^{(۱)} = \begin{bmatrix} \mathbf{C}_{۱۱} & \mathbf{C}_{۱۲} \\ \mathbf{C}_{۱۲}^{\top} & \mathbf{C}_{۲۲} \end{bmatrix} \approx \mathbf{S} = \begin{bmatrix} \mathbf{S}_{۱۱} & \mathbf{S}_{۱۲} \\ \mathbf{S}_{۱۲}^{\top} & \mathbf{S}_{۲۲} \end{bmatrix},$$

$$\mathbf{C}^{(۲)} = \begin{bmatrix} \mathbf{C}_{۲۲} & \mathbf{C}_{۲۳} \\ \mathbf{C}_{۲۳}^{\top} & \mathbf{C}_{۳۳} \end{bmatrix} \approx \mathbf{T} = \begin{bmatrix} \mathbf{T}_{۲۲} & \mathbf{T}_{۲۳} \\ \mathbf{T}_{۲۳}^{\top} & \mathbf{T}_{۳۳} \end{bmatrix}$$

چون $\mathbf{S}, \mathbf{T}$ دو تخمین همراه با خطا هستند، بر خلاف حالت ایده‌آل لزوماً $\mathbf{S}_{۲۲}, \mathbf{T}_{۲۲}$ یکسان نیستند. ولی فرض می‌کنیم که تخمین‌های $\mathbf{S}, \mathbf{T}$ مثبت معین هستند.

۱. یک راه ساده برای پرکردن ماتریس کواریانس، استفاده از حدس‌های زیر است:

$$\mathbf{C}_{۱۱} = \mathbf{S}_{۱۱}, \quad \mathbf{C}_{۱۲} = \mathbf{S}_{۱۲}, \quad \mathbf{C}_{۱۳} = \mathbf{0},$$
$$\mathbf{C}_{۲۲} = \frac{۱}{۲}(\mathbf{S}_{۲۲} + \mathbf{T}_{۲۲}), \quad \mathbf{C}_{۲۳} = \mathbf{T}_{۲۳}, \quad \mathbf{C}_{۳۳} = \mathbf{T}_{۳۳}$$

با استفاده از یک مثال نقض ساده نشان دهید که ماتریس حاصل از این روش نمی‌تواند یک ماتریس کوواریانس باشد.

۲. با استفاده از بهینه‌سازی محدب، سعی می‌کنیم مسئله قسمت قبل را حل کنیم! حدس ساده‌ی قسمت قبل را با $\mathbf{C}_{\text{sim}}$ نشان می‌دهیم. مسئله‌ی زیر را در نظر بگیرید:

$$\begin{array}{ll} \text{Minimize} & \|\mathbf{C} - \mathbf{C}_{\text{sim}}\|_F^{۲} \\ \text{subject to:} & \mathbf{C} \succeq \mathbf{0}. \end{array}$$

تحت چه شرایطی جواب مسئله فوق برابر با $\mathbf{C}_{\text{sim}}$ است؟

۳. برنامه‌ای بنویسید که این مسئله را حل کند. برای مثال نقض قسمت اول این برنامه را اجرا کنید و نشان دهید پاسخ بدست آمده قابل قبول است.

۴ (∗) کنترل به کمک توابع هدف مختلف

یک مسئله کنترل بهینه استاندارد را با دینامیک $\mathbf{x}_{t+۱} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t$، که $t = ۰, ۱, \ldots, T - ۱$ در نظر بگیرید.
در اینجا $\mathbf{x}_t \in \mathbb{R}^n$ حالت در زمان t، $\mathbf{u}_t \in \mathbb{R}^m$ کنترل یا ورودی در زمان t، $\mathbf{A} \in \mathbb{R}^{n\times n}$ ماتریس دینامیک و $\mathbf{B} \in \mathbb{R}^{n\times m}$ ماتیس ورودی است. حالت اولیه به عنوان $\mathbf{x}_۰ = \mathbf{x}^{\text{init}}$ به ما داده شده و نیاز است که حالت نهایی صفر باشد، یعنی $\mathbf{x}_T = \mathbf{0}$ (در کاربرد، حالت $\mathbf{0}$ مربوط به برخی حالت‌های مطلوب است). حال باید دنباله‌ای از ورودی‌ها مانند $\mathbf{u}_۰, \ldots, \mathbf{u}_{T-۱}$ را به گونه‌ای انتخاب کنید که تابع هدف را کمینه کند.
مقادیر $\mathbf{x}^{\text{init}}, \mathbf{A}, \mathbf{B}$ و T در فایل `various_obj_regulator_data.py` داده شده‌اند.
برای این امر توابع هدف مختلفی را در نظر می‌گیریم، که همه آن‌ها اندازه‌ی ورودی‌ها (یا به عبارتی، اثر کنترل) را اندازه‌گیری می‌کنند.

1. $\sum_{t=0}^{T-1} \|\mathbf{u}_t\|_2^2$.
2. $\sum_{t=0}^{T-1} \|\mathbf{u}_t\|_2$.
3. $\max_{t=0,\ldots,T-1} \|\mathbf{u}_t\|_2$.
4. $\sum_{t=0}^{T-1} \|\mathbf{u}_t\|_1$.

برای هر کدام از تابع هدف‌ها، نمودار مؤلّفه‌های ورودی بهینه و همچنین $\|\mathbf{u}_t\|_۲$ بر حسب t را رسم کنید و رفتار نمودارها را به صورت مختصر تحلیل کنید.

۵ بهینه‌سازی پورتفولیو

فرض کنید $\mathbf{w} \in \mathbb{R}^n$ بردار وزن‌های یک پورتفولیو باشد، به طوری که مقادیر منفی مربوط به پوزیشن‌های کوتاه[1] هستند و وزن‌ها به گونه‌ای نرمالیزه شده‌اند که $\mathbf{1}^\top \mathbf{w} = ۱$. بازدهی مورد انتظار پورتفولیو برابر $\boldsymbol{\mu}^\top \mathbf{w}$ است، که $\boldsymbol{\mu} \in \mathbb{R}^n$ بردار بازدهی انتظاری[2] دارایی‌هاست. ما ریسک پورتفولیو را به کمک واریانس بازدهی آن اندازه‌گیری می‌کنیم. در این مسئله ما ماتریس کوواریانس $\boldsymbol{\Sigma}$ بازدهی دارایی‌ها را نمی‌دانیم؛ در عوض فرض می‌کنیم که $\boldsymbol{\Sigma}$ یکی از M ماتریس کوواریانس $\{\boldsymbol{\Sigma}^{(k)}\}_{k=۱}^{M}$، است که به ازای هر $k \in [M]$ داریم $\boldsymbol{\Sigma}^{(k)} \in \mathbb{S}_n^{++}$.

ما می‌توانیم $\boldsymbol{\Sigma}^{(k)}$ها را به عنوان M مدل ریسک مختلف در نظر بگیریم که مرتبط با M رژیم مختلف بازار هستند. برای یک بردار وزن $\mathbf{w}$، M مقدار مختلف ممکن برای ریسک وجود دارد:

$$\mathbf{w}^\top \boldsymbol{\Sigma}^{(k)} \mathbf{w}, \qquad k = ۱, \ldots, M$$

بدترین حالت ریسک، در میان مدل‌های مختلف،

$$\max_{k \in [M]} \mathbf{w}^\top \boldsymbol{\Sigma}^{(k)} \mathbf{w}$$

است (که معادل با بدترین حالت ریسک در میان تمام ماتریس‌های کوواریانس در پوش محدب $\{\boldsymbol{\Sigma}^{(k)}\}_{k=۱}^{M}$ خواهد بود.)

ما وزن‌های پورتفولیو را به گونه‌ای انتخاب خواهیم کرد که بازدهی انتظاری بیشنه شود. همچنین بدترین حالت ریسک را در نظر خواهیم گرفت. در واقع مسئله‌ی بهینه‌سازی ما بدین صورت خواهد بود:

$$\begin{aligned} \text{Maximize} \quad & \boldsymbol{\mu}^\top \mathbf{w} - \gamma \max_{k \in [M]} \mathbf{w}^\top \boldsymbol{\Sigma}^{(k)} \mathbf{w} \\ \text{subject to:} \quad & \mathbf{1}^\top \mathbf{w} = ۱ \end{aligned} \tag{۲}$$

که متغیر مسئله $\mathbf{w}$ است و $\gamma > ۰$ پارامتر ریسک‌گریزی است که داده‌شده است.

۱.۵ بخش اول

ثابت کنید مقادیر نامنفی $\gamma_۱, \ldots, \gamma_M$ وجود دارند به طوری که $\sum_{k=۱}^{M} \gamma_k = \gamma$ و جواب $\mathbf{w}^*$ مسئله‌ی (۲) با جواب مسئله‌ی زیر یکسان است:

$$\begin{aligned} \text{Maximize} \quad & \boldsymbol{\mu}^\top \mathbf{w} - \sum_{k=۱}^{T} \gamma_k \mathbf{w}^\top \boldsymbol{\Sigma}^{(k)} \mathbf{w} \\ \text{subject to:} \quad & \mathbf{1}^\top \mathbf{w} = ۱ \end{aligned} \tag{۳}$$

نتیجه بالا تفسیر جالبی دارد: می‌توانیم γ_k را عنوان تخصیص دهنده‌ی ریسک کلی خودمان (γ) به مسئله‌ی (۲) در میان M رژیم مختلف در نظر بگیریم.

راهنمایی: مقادیر γ_k به راحتی یافت نمی‌شوند و برای پیدا کردن آن‌ها باید ابتدا مسئله‌ی (۲) را حل کنید. بنابراین، این نتیجه به ما در حل مسئله‌ی (۲) کمک نمی‌کند؛ بلکه تفسیر جالبی از راه‌حل آن ارائه می‌دهد.

۲.۵ بخش دوم

وزن‌های بهینه‌ی پورتفولیو را برای نمونه‌ی مسئله با داده‌های داده شده در فایل `multi_risk_portfolio_data.py` پیدا کنید. وزن‌ها و مقادیر $\{\gamma_k\}_{k=۱}^{M}$ را به همراه M مقدار ممکن ریسک مرتبط به وزن‌ها، و بدترین حالت ریسک را گزارش کنید.

[1] short
[2] expected return