In the Name of God

Convex Optimization I (25756-1)

CHW 1

Fall Semester 1402-03

Department of Electrical Engineering

Sharif University of Technology

Instructor: Dr. R. Amiri

Due on Azar 3, 1402 at 23:59

1 DCP Representation with CVXPY

The function $f(x, y) = 1/(xy)$ with $x, y \in \mathbb{R}$, $\mathbf{dom} f = \mathbb{R}^2_{++}$, is convex. How do we represent it using disciplined convex programming (DCP), and the functions $1/u$, $\sqrt{uv}$, $\sqrt{u}$, u^2, u^2/v, addition, subtraction, and scalar multiplication? (These functions have the obvious domains, and you can assume a sign-sensitive version of DCP, e.g., u^2/v increasing in u for $u \geq 0$.)
To learn about CVXPY and how to solve optimization problems with it, you can study the examples on this page.

2 Optimizing a set of disks

A disk $D \subset \mathbb{R}^2$ is parametrized by its center $c \in \mathbb{R}^2$ and its radius $r \geq 0$, with the form $D = \{x \mid \|x - c\|_2 \leq r\}$. (We allow $r = 0$, in which case the disk reduces to a single point $\{c\}$.) The goal is to choose a set of n disks $D_1, \ldots, D_n$ (i.e., specify their centers and radii) to minimize an objective subject to some constraints. One constraint is that the first k disks are fixed, i.e.,

$$c_i = c_i^{\text{fix}}, \quad r_i = r_i^{\text{fix}}, \quad i = 1, \ldots, k,$$

where c_i^{fix} and r_i^{fix} are given. The second constraint is an overlap or intersection constraint, which requires some pairs of disks to intersect:

$$D_i \cap D_j \neq \emptyset \ , \ (i, j) \in I,$$

where $I \subset \{1, \ldots, n\}^2$ is given. You can assume that for each $(i, j) \in I$, $i < j$. We consider two objectives: The sum of the disk areas, and the sum of the disk perimeters. These two objectives result in two different problems.

1. Explain how to solve these two problems using convex optimization.
2. Solve both problems for the problem data given in **disks_data.py**. Give the optimal total area and the optimal total perimeter. Plot the two optimal disk arrangements, using the code included in the data file. Give a very brief comment on the results, especially the distribution of disk radii each problem obtains.

3 Bandlimited signal recovery from zero-crossings

Let $y \in \mathbb{R}^n$ denote a bandlimited signal, which means that it can be expressed as a linear combination of sinusoids with frequencies in a band:

$$y_t = \sum_{j=1}^{B} a_j \cos\left(\frac{2\pi}{n}(f_{\min} + j - 1)t\right) + b_j \sin\left(\frac{2\pi}{n}(f_{\min} + j - 1)t\right), \quad t = 1, \ldots, n$$

where $f_{\min}$ is the lowest frequency in the band, B is the bandwidth, and $a, b \in \mathbb{R}^B$ are the cosine and sine coefficients, respectively. We are given $f_{\min}$ and B, but not the coefficients a, b or the signal y. We do not know y, but we are given its sign $s = \mathrm{sign}(y)$, where $s_t = 1$ if $y_t \geq 0$ and $s_t = -1$ if $y_t < 0$. (Up to a change of the overall sign, this is the same as knowing the zero-crossings of the signal, i.e., when it changes sign. Hence the name of this problem.) We seek an estimate $\hat{y}$ of y that is consistent with the bandlimited assumption and the given signs. Of course, we cannot distinguish y and αy, where $\alpha > 0$, since both of these signals have the same sign pattern. Thus, we can only estimate y up to a positive scale factor. To normalize $\hat{y}$, we will require that $\|\hat{y}\|_1 = n$, i.e., the average value of $|y_i|$ is one. Among all $\hat{y}$ that are consistent with the bandlimited assumption, the given signs, and the normalization, we choose the one that minimizes $\|\hat{y}\|_2$.

1. Show how to find $\hat{y}$ using convex optimization.

2. Apply your method to the problem instance with data given in **zero_crossings_data.py**. The data files also include the true signal y (which, of course, you cannot use to find $\hat{y}$). Plot $\hat{y}$ and y, and report the relative recovery error, $\frac{\|y-\hat{y}\|_2}{\|y\|_2}$. Give one short sentence commenting on the quality of the recovery.

4 Fitting a periodic Poisson distribution to data

We model the (random) number of times that some type of event occurs in each hour of the day as independent Poisson variables, with

$$\Pr(k \text{ events occur}) = e^{-\lambda_t} \frac{\lambda_t^k}{k!}, \quad k = 0, 1, \ldots$$

with parameter $\lambda_t \geq 0$, $t = 1, \ldots, 24$. (For $\lambda_t = 0$, $k = 0$ events occur with probability one.) Here, t denotes the hour, with $t = 1$ corresponding to the hour from midnight to 1AM, and $t = 24$ the hour between 11PM and midnight. (This is the periodic Poisson distribution in the title.) The parameter λ_t is the expected value of the number of events that occur in hour t; it can be thought of as the rate of occurrence of the events in hour t. Over one day, we observe the numbers of events $N_1, \ldots, N_{24}$.

1. What is the maximum likelihood estimate of the parameters $\lambda_1, \ldots, \lambda_{24}$? Hint: There is a simple analytical solution. You should consider the cases $N_t > 0$ and $N_t = 0$ separately.

2. In many applications, it is reasonable to assume that λ_t varies smoothly over the day; for example, the rate of occurrence of events for 3PM–4PM is not too different from the rate of occurrence for 4PM–5PM. To obtain a smooth estimate of λ_t, we maximize the log-likelihood minus the regularization term

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

   where $\rho \geq 0$. Explain how to find the values $\lambda_1, \ldots, \lambda_{24}$ using convex optimization. If you change variables, explain.

3. What happens as $\rho \to \infty$? You can give a very short answer with an informal argument. Hint: As in part (1), there is a simple analytical solution.

4. Over one day, we observe

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

   Find the regularized maximum likelihood parameters for $\rho \in \{0.1, 1, 10, 100\}$ using CVXPY, and plot λ_t versus t for each value of ρ.

5. One way to choose the value of ρ is to see which of the models found in part (4) has the highest log-likelihood on a test set, i.e., another day's data that was not used to create the model. For each of the four values of the parameters you estimated in part (4), evaluate the log likelihood of another day's number of events,

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

   Which hyper-parameter value ρ would you choose?