Advanced Topics in Control 2020: Large-Scale Convex Optimization

Exercise 11: Applications

Goran Banjac, Mathias Hudoba de Badyn, Andrea Iannelli,
Angeliki Kamoutsi, Ilnura Usmanova

May 26, 2020

Date due: June 4, 2020 at 23:59.

Please submit your written solutions via Moodle as a PDF with filename `Ex10_Surname.pdf`, replacing `Surname` with your surname. Include any code as separate files (not inside the PDF itself), and please don't zip anything.

---

1 Problem 1

Thanks to your incredible performance in de-noising the distress signals Mathias sent out, the ATIC team has managed to inverse-Laplace transform him back into normal space. Unfortunately, while he was still transcending space and time, he had acausally watched all English-language media ever made. In order for him to survive the quarantine, he has to branch out into other media. Using your optimization skills, you will now have to recommend him the best possible Japanese anime.

You will find two datasets: `anime_upload.csv` contains 11153 different animes, each with a unique ID number that doesn't necessarily go from 1–11153, the anime title, and a list of genres that the anime belongs to. The second dataset, `rating_upload.csv`, contains data from 73516 users of a popular anime streaming website. A user had the option of watching an anime, and then rating it from 1–10. If a user did not rate the anime, a score of '-1' was recorded.

We will model the task of constructing recommendations as a rank-minimization problem. Let X be a matrix whose rows correspond to users, and whose columns correspond to animes. The entry X_{ij} is the ith user's opinion on the anime j. We will be using the nuclear norm as a proxy for the rank of X. Lastly, let $\Omega = \{(i,j)\}$ be the set of ratings that we know. For example, M_{ij} is the ith user's known rating of anime j.
Our optimization problem is thus

$$\begin{array}{ll} \min_X & \|X\|_* \\ \text{s.t.} & X_{ij} = M_{ij},\ \forall\ (i,j) \in \Omega. \end{array} \tag{P1}$$

1. Let M_{ij} be the numerical rating out of 10 that user i assigns to anime j, and solve Problem (P1). Ignore all entries of '-1', i.e. set them to zero. What is the resulting recommendation for user 1337 for the anime *Pokemon: Slowking no Ichinichi*?

Anime	Rating
Steins;Gate	10
Steins;Gate 0	8
Shinsekai yori	5
Serial Experiments Lain	10
Perfect Blue	6.5
Paprika	6

Table 1: Mathias' anime ratings

2. Mathias has watched only a few animes in his life. His ratings for them are in Table 1. Add a new row to X corresponding to Mathias. Given his ratings in Table 1, what are the top 5 animes he should watch?

3. Notice that most of the animes in Table 1 are of the Sci-Fi genre. Solve problem (P1) using data *only* from animes that are tagged as being Sci-Fi. If there are any zero columns or zero rows in your matrix X, make sure to re-define X in such a way that these rows and columns are removed. What are the top 5 animes Mathias should watch?

4. Mathias' otaku friend Sara forced him to watch *Mahou Shoujo Madoka★Magica*, and he liked it so much he gave it a 9. Add this rating to your solution to Part (2) (the case with all the data). Do the recommendations for the top 5 animes he should watch change? Is this a user-friendly or desired feature? Discuss the business consequences of this.

5. What happens if you try a binary classifier, i.e. if a user i rated an anime j 6 or more, then set $M_{ij} = 1$ and zero otherwise. What happens to the recommendations for Mathias? Is this a better classifier method? **Hint:** Think about how subjective ratings can be, and how the scales differ from person to person.