

3. Problem 5.1a and 5.1b from *Convex optimization* by Boyd and Vandenberghe, Page 273.

4. Suppose you manage a network of service stations in which customers are served. (Think of a call center, web site, or bureaucratic government service office.) The system consists of 3 stations, and each suppose station can be modeled as a type of queue called an M/M/1 queue. The average number of customers in such a queue is known to converge to
\[
\frac{\lambda}{\mu - \lambda}
\]
where \(\mu\) is the service rate of the station (when it is busy) and \(\lambda\) is the average arrival rate of customers to the station. Note that as \(\lambda\) approaches \(\mu\), the average number of customers in queue approaches \(\infty\). Note the above expression is only valid when \(\lambda \leq \mu\). If \(\lambda > \mu\), the queue grows unbounded.

In your network customers arrive to the network at rate \(a\). Customers who finish service at station 1 go to station 2 with probability \(p\) and go to station 3 with probability \(1 - p\). Customers who finish service at station 2 return to station 1 with probability \(q\) and exit the system with probability \(1 - q\). Customers who finish service at station 3 return to station 1 with chance \(v\) and exit with chance \(1 - v\).

- What is the total arrival rate to each of the 3 stations, assuming that they have sufficient service rates to be stable? Hint: Define \(x\) to be the vector of arrivals to each station. Recognize that the average rate customers depart a stable station is the same as the rate customers arrive. (If it were less, the queue would be blowing up, if it were more, somehow the station would have to be creating customers!) Thus you can write equations for each station that reflect this. For instance, for the first station, you have
\[
x_1 = a + qx_2 + vx_3.
\]
Write 2 more of such equations and solve them for \(x\).

- Suppose the service rates at the stations 1, 2, and 3 are \(\mu_1\), \(\mu_2\), and \(\mu_3\) respectively. Under what conditions will the network be stable?

- It has been shown that the average delay of a customer in a queueing system is equal to the average number of customers in the system divided by the arrival rate (this is called Little’s result.) Thus, for your system, the average delay is
\[
\frac{1}{a} \left[ \frac{x_1}{\mu_1 - x_1} + \frac{x_2}{\mu_2 - x_2} + \frac{x_3}{\mu_3 - x_3} \right].
\]
Suppose now that you can design \(\mu = [\mu_1, \mu_2, \mu_3]^T\) to reduce delay. In particular, suppose each unit of service rate costs $1 and you have $\(C\) to spend. (Thus your constraint is \(\mu_1 + \mu_2 + \mu_3 \leq C\). What choice of service rates minimizes the average delay? Give your answer in terms of \(\lambda\), \(p\), \(q\), and \(v\). Also give your answer numerically for the particular case of \(p = 0.5\), \(\lambda = q = v = 0.1\).