Dynamic Priority Classes (?)

authors (?)

1 Background

- [?] (1960): Nonpreemptive M/M/1 where customers are served in order of the difference between their waiting time and urgency number (that is priorities increasing linearly over time). Solved by considering event probabilities at clock ticks.
- [?] (1971): Similar to the above, but treat each urgency number as a separate customer class, and not considering clock ticks. Upper and lower bounds on the waiting times, based on FIFO and static priorities.
- [?] (1979): Now considers the case where priorities increase non-linearly but concavely over time.
- [?] (1990): Nonpreemptive M/G/1 queue with two classes of customers, where priorities switch if the number from one class exceeds a given threshold. Lower priority customers have a finite waiting capacity, higher have infinite capacity.
- [?] (2003): Similar to the above but with Markovian services and infinite waiting capacities for both customers.
- [?] (2008): Preemptive n-priority-classes M/M/c with exponential upgrades. Customers only upgrade to the priority immediately higher than themselves. Stability considered.
- [?] (2010): Preemptive two-priority-classes M/M/c with exponential upgrades. Customers cannot upgrade if the number of lower priority customers is below a given threshold. Holding costs considered.
- [?] (2012): Extention of the above, allows batch arrivals, multiple classes, phase-type upgrades and services. Customers only upgrade to the priority immediately higher than themselves.
- [?] (2022): Analytical (truncated) expressions for a two class delayed accumulating priority M/G/1 queue. Customer priorities increase linearly over time, at different rates according to class, after an initial fixed delay.

At the moment I don't see any consideration of customers downgrading, or 'skipping' classes, i.e. transfering to a priority class not immediately higher than themselves.

2 The Model

Here we consider an $M/M/c$ queue with K classes of customer. Order and label the customer classes $0, 1, 2, \dots, K-1$, with customer classes with lower labels having priority over customer classes of higher labels. The index k will be used to represent customer classes. Let:

- λ_k be the arrival rate of customers of class k,
- μ_k be the service rate of customers of class k,
- θ_{k_i,k_j} be the rate at which customers of class k_i change to customers of class k_j.

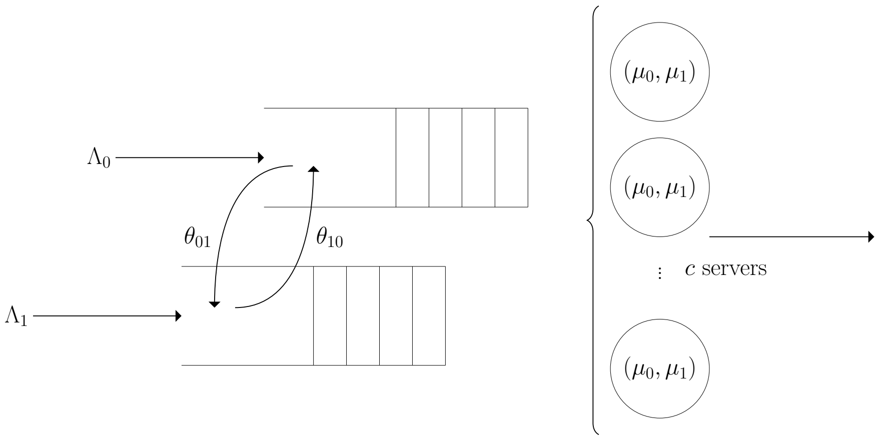

Figure 1: An example of a two-class priority queue.

Figure **??** shows an example with two classes of customer.

3 State Markov Chain Formulation

Let $\underline{\mathbf{s}}_t = (s_{0,t}, s_{1,t}, \dots, s_{K-1,t}) \in \mathbb{N}^K$ represent the state of the system at time step t, where $s_{k,t}$ represents the number of customers of class k present at time step t.

Then the rates of change between $\underline{\mathbf{s}}_t$ and $\underline{\mathbf{s}}_{t+1}$ are given by Equation **??**, where $\underline{\delta} = \underline{\mathbf{s}}_t - \underline{\mathbf{s}}_{t+1}$,

$$q_{\underline{\mathbf{s}}_t, \underline{\mathbf{s}}_{t+1}} = \begin{cases} \lambda_k & \text{if } \delta_k = 1 \text{ and } \delta_i = 0 \; \forall \; i \neq k \\ B_{k,t}\mu_k & \text{if } \delta_k = 1 \text{ and } \delta_i = 0 \; \forall \; i \neq k \text{ and } \sum_{i<k} s_{i,t} < c \\ (s_{k,t} - B_{k,t})\theta_{k_0,k_1} & \text{if } \delta_{k_0} = -1 \text{ and } \delta_{k_1} = 1 \text{ and } \delta_i = 0 \; \forall \; i \neq k_0, k_1 \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

and $B_{k,t}$, representing the number of customers of class k currently in service at time step t, is given by Equation **??**.

$$B_{k,t} = \min\left(c - \min\left(\sum_{i<k} s_{i,t}, c\right), s_{k,t}\right) \tag{2}$$

TODO (?) Unroll the matrix to get a (double?) infinite sum for state probabilities. If possible this would make calculations much more efficient by truncating an infinite sum rather than an infinite matrix.

4 Sojourn Time Markov Chain Formulation

Let $\underline{\mathbf{z}}_t = (z_{0,t}, z_{1,t}, \dots, z_{n,t} \dots, z_{K-1,t}, b_t, n_t) \in \mathbb{N}^{K+2}$ represent the state of a particular customer at time step t, where n_t represents that customer's class at time t; $z_{k,t} \; \forall \; k < n$ represents the number of customers of class k in front of the customer in the queue at time t; $z_{k,t} \; \forall \; n < k < K$ represents the number of customers of class k behind the customer in the queue at time t; and b_t represent the number of customers of class n_t behind the customer in the queue at time t. Also let $\star$ represent an absorbing state, representing the state where that customer has finished service and left the system.

Then the rates of change between $\underline{\mathbf{z}}_t$ and $\underline{\mathbf{z}}_{t+1}$ are given by Equation **??**, where $\underline{\delta} = \underline{\mathbf{z}}_t - \underline{\mathbf{z}}_{t+1}$,

$$q_{\underline{\mathbf{z}}_t,\underline{\mathbf{z}}_{t+1}} = \begin{cases} \mu_n & \text{if } z_{t+1} = \star \text{ and } \sum_{k \leq n} z_{k,t} < c \\ \lambda_n & \text{if } \delta_K = 1 \text{ and } \delta_i = 0 \,\forall\, i \neq K \\ \lambda_k & \text{if } \delta_k = 1 \text{ and } \delta_i = 0 \,\forall\, i \neq k \text{ and } k \neq n \\ A_{k,n,t}\mu_k & \text{if } \delta_k = -1 \text{ and } \delta_i = 0 \,\forall\, i \neq k \text{ and } k < K \\ \tilde{A}_{n,t}\mu_n & \text{if } \delta_K = -1 \text{ and } \delta_i = 0 \,\forall\, i \neq K \\ (z_{k_0,t} - A_{k_0,n,t})\theta_{k_0,k_1} & \text{if } \delta_{k_0} = -1 \text{ and } \delta_{k_1} = 1 \text{ and } \delta_i = 0 \,\forall\, i \neq k_0, k_1 \text{ and } k_0 < K \text{ and } k_1 \neq n, K, K+1 \\ (z_{K,t} - \tilde{A}_{n,t})\theta_{n,k} & \text{if } \delta_K = -1 \text{ and } \delta_k = 1 \text{ and } \delta_i = 0 \,\forall\, i \neq k, n \text{ and } k < K \\ (z_{k,t} - A_{k,n,t})\theta_{k,n} & \text{if } \delta_k = -1 \text{ and } \delta_K = 1 \text{ and } \delta_i = 0 \,\forall\, i \neq k, K \\ \theta_{n,k} & \text{if } \delta_n = z_{K,t} \text{ and } \delta_K = -z_{K,t} \text{ and } \delta_{K+1} = n - k \text{ and } \delta_i = 0 \text{ otherwise, and } \sum_{k \leq n} z_{k,t} < c \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

and $A_{k,n,t}$ and $\tilde{A}_{n,t}$ are given by Equations **??** and **??**.

$$A_{k,n,t} = \begin{cases} \min\left(c, \sum_{i \leq k} z_{i,t}\right) - \min\left(c \sum_{i < k} z_{i,t}\right) & \text{if } k \leq n \\ \min\left(c, \sum_{i \leq k} z_{i,t} + 1 + z_{K,t}\right) - \min\left(c \sum_{i < k} z_{i,t} + 1 + z_{K,t}\right) & \text{if } n < k < K \end{cases} \tag{4}$$

$$\tilde{A}_{n,t} = \min\left(c, \sum_{i \leq n} z_{i,t} + 1 + z_{K,t}\right) - \min\left(c, \sum_{i \leq n} z_{i,t} + 1\right) \tag{5}$$

The expected time to absorption can be calculated from each state. Customers arrive in all states where $z_{K,t} = 0$, and their class can be determined by n. Combining these times to absorption with the state probabilities found in the previous section, the sojourn times for each customer class can be found.

TODO (?) Unroll the matrix to get a (double?) infinite sum for state probabilities. If possible this would make calculations much more efficient by truncating an infinite sum rather than an infinite matrix.

5 Numerical Experiments

TODO: Effect of class transition rates on sojourn times.

TODO: Find stability based on the relationships between θ_{k_i,k_j}, λ_k and μ_k.