The deviation matrix of the $M/M/1/\infty$ and $M/M/1/N$ queue, with applications to controlled queueing models[1]

Ger Koole
Vrije Universiteit
Faculty of Mathematics and Computer Science
De Boelelaan 1081a, 1081 HV Amsterdam
The Netherlands

Abstract

In this paper we give closed expressions for the deviation matrices of the $M/M/1/\infty$ and the $M/M/1/N$ queue. It is applied to controlled queueing models by solving the dynamic programming optimality equation for Bernoulli policies, after which we can execute a single step of policy iteration, by calculating the bias vector of the Bernoulli policies directly from the deviation matrix.

1 Introduction

The main result of this paper is the derivation of the deviation matrices of the $M/M/1/\infty$ and the $M/M/1/N$ queue, see (1) and (2)-(3). The (i, j)th entry of the deviation matrix of a Markov chain can be interpreted as the expected total difference in number of visits to a state j starting from i compared to the number of visits to j starting from equilibrium. The derivation of the formulas and more details on the deviation matrix can be found in Section 2.

Then we apply the formulas to some controlled queueing problems. To explain this, let us first make some general observations about Markov decision chains (MDCs) with the average cost criterion. A general solution method for MDCs is *policy iteration* (see e.g. Puterman [3], Sections 8.6 and 9.2). It consists of the repetition of two steps: For a given policy, compute the average costs and bias vector (to be explained later), and, for a given policy and bias, compute a new (better) policy. The second step is easy; the first however is numerically burdensome, often an iterative method is used. This can take a lot of time, especially for high-dimensional models (the so-called *curse of dimensionality*).

From the deviation matrix of a Markov chain it is very easy to derive its bias vector for a particular choice of cost function. Now suppose that we have a controlled queueing model which behaves as a single server queue for a certain policy or class of policies. Then for this policy we can compute the bias from the formulas, and obtain an improved policy, without iterating. This policy could serve as an approximation for the optimal policy. This procedure is applied to an $M/M/1/\infty$ queue with admission control in Section 3.

The bias vector of a network of independent queues is equal to the sum of the bias vectors of the queues. This opens the possibility to compute bias vectors for network models as well. We do this in Section 3 for a routing model with finite buffers.

The idea of doing a single step of policy iteration, based on a policy with a known bias, was introduced in Ott & Krishnan [2]. For their controlled queueing model (a network of $M/M/s/s$ queues) they compute the difference in bias necessary to obtain the improved policy. The objective is to minimize the average blocking probability. Sassen, Tijms & Nobel [4] apply the method to routing to parallel $M/G/1/\infty$ queues, where minimization of the average number of customers in the system is the objective. Conceptually we generalize this method by starting from the deviation matrix instead of the bias itself. This makes our analysis independent of the choice of objective.

2 The deviation matrices

First we state some general results concerning Markov (decision) chains. Then we derive the deviation matrices of the discrete-time single server queues. After that we make the transition to continuous-time models. We pass by discrete-time models because the deviation matrix is only defined for discrete-time Markov chains.

We study an irreducible aperiodic Markov chain on a countable state space, having a stationary distribution.

[1] Proc. of the 37th IEEE CDC, Tampa, pages 56-59, 1998

For this case we define the stationary matrix and the deviation matrix. For the more general case of multiple classes, see Puterman [3], Appendix A. Define P as the transition matrix, and let P^* be the stationary matrix, i.e.,

$$P^* = \lim_{n\to\infty} P^n.$$

Because P is aperiodic the limit exists. The entry p^*_{ij} of P^* is equal to the stationary probability of j: due to the irreducibility the rows of P^* are identical. We define the deviation matrix D by

$$D = \sum_{n=0}^{\infty}(P^n - P^*).$$

In fact, Puterman [3] only deals with the finite state space case; to prove existence of D for $M/M/1/\infty$ we will rely on a result of Hordijk & Spieksma [1].

Now consider discrete-time single server queues. Later on we give the connection with the standard continuous-time queues. The transitions of the queues are as follows: for all states $i \in \{0,\dots,N\}$ (with N possibly ∞) we have as transition probabilities to other states $p_{ii+1} = \lambda$ if $i < N$ and $p_{ii-1} = \mu$ if $i > 0$. Of course we assume that $\lambda + \mu \le 1$, and let $\rho = \lambda/\mu$. It is clear that this chain is aperiodic and irreducible. The stationary probabilities are as for the equivalent continuous-time queues. If $N = \infty$ we assume that $\lambda < \mu$, thus a stationary distribution exists. Then the stationary matrices, denoted by $P^{*\infty}$ and P^{*N}, have entries $p^{*\infty}_{ij} = (1-\rho)\rho^i$ and $p^{*N}_{ij} = \frac{1-\rho}{1-\rho^{N+1}}\rho^i$ (with the exception that $p^{*N}_{ij} = \frac{1}{N+1}$ if $\lambda = \mu$).

Theorem 2.1 *If $\lambda < \mu$, then the deviation matrix D^∞ of the discrete-time $M/M/1/\infty$ exists and is given by*

$$d^\infty_{ij} = \frac{\rho^{\max\{j-i,0\}} - (i+j+1)(1-\rho)\rho^j}{\mu(1-\rho)}; \qquad (1)$$

the deviation matrix D^N of the discrete time $M/M/1/N$ queue exists and is given by

$$d^N_{ij} = \frac{\rho^{\max\{j-i,0\}}}{\mu(1-\rho)} + \frac{\rho^j}{\mu(1-\rho)(1-\rho^{N+1})} \times \left(\rho^{N-i+1} + \rho^{N-j+1} - (i+j-1)(1-\rho) - 2\right) + \frac{2\rho^{j+1}}{\mu(1-\rho)(1-\rho^{N+1})^2}\left(1-(N+1)\rho^N + N\rho^{N+1}\right) \qquad (2)$$

if $\lambda \neq \mu$ and

$$d^N_{ij} = \frac{-\max\{i-j,0\}}{\mu} + \frac{i(i+1)+(N-j)(N+1-j)}{2\mu(N+1)} + \frac{-N(N+2)}{6\mu(N+1)} \qquad (3)$$

if $\lambda = \mu$.

Proof. First we prove the existence. For the $M/M/1/N$ queue this follows directly from the aperiodicity, see Theorem A.7 in [3]. (Note that the deviation matrix is usually defined as a Cesaro limit.) For the $M/M/1/\infty$ queue we rely on Theorem 2.1 in [1], which states that we need to show *μ-uniform geometric recurrence* (defined on page 500 in [1]). It is readily seen that a function such as $\mu_j = j$ suffices, with $M = \{0\}$.

Having shown the existence, we can now check the formulas (1), (2) and (3). By Lemma 3.1 of [1] D is the unique solution of $P^* + D = I + PD$ and $P^*D = 0$. By some tedious calculations it can be shown that D^∞ and D^N satisfy these equalities for the appropriate P and P^*. □

As stated in the proof of Theorem 2.1, the equation $P^* + D = I + PD$ holds. Multiplying with the vector c gives $P^*c + Dc = c + PDc$, and thus (P^*c, Dc) is a solution to the dynamic programming *optimality equation* $\phi + v = c + Pv$. Theory on Markov decision chains shows us that ϕ is equal to the average long-term costs of the chain, while $v_i - v_k$ can be seen as the total difference in costs between starting in i and starting in k. Because $P^*D = 0$ and thus $P^*Dc = 0$, we can interpret $(Dc)_i$ as the total difference in costs between starting in i and starting in an initially stationary situation. (See [3], in particular Section 8.2.1 and 8.2.3.)

By taking $c_i = \delta_j$ in the optimality equation, with δ the Kronecker delta, we find $(Dc)_i = d_{ij}$, and thus d_{ij} can be interpreted as the difference between the total number of visits to j, starting in i, compared to the number of visits to j starting from a stationary situation.

Note that, in the case of the $M/M/1/\infty$ queue, we need to show that P^*c and Dc exist. This can be shown using the concepts of [1]. It amounts to finding a vector μ for which $||c||_\mu < \infty$ and μ-uniform geometric recurrence holds (for the notation, see [1]). Note that for c linear, i.e., of the form $c_i = iC$, the choice $\mu_j = j$ already suffices.

Finally we pay attention to the relation between the discrete-time models we studied so far and the standard continuous-time models. We study queues with arrival rate λ and service rate μ (without upper bound on the sum), where there are direct costs c_j for each time unit that the system stays in state j. With the help of Serfozo [5] we can compare this system to a discrete-time system with transition probabilities $\lambda' = \lambda/\alpha$ and $\mu' = \mu/\alpha$, and direct costs $c'_j = c_j$, with $\alpha \ge \lambda + \mu$. It follows from [5] that (if the average costs exist) the continuous-time and discrete-time model have the same average costs and the same optimal policy. The optimality equation is equal to

$$\phi + \frac{\lambda+\mu}{\alpha} v_i = c_i + \frac{\lambda}{\alpha} v_{\min\{i+1,N\}} + \frac{\mu}{\alpha} v_{(i-1)^+}.$$

Write $d_{ij}(\lambda,\mu)$ for d_{ij} in (1)-(3), and note that p^*_{ij} depends only on ρ. A solution to the above equation is then given by $(P^*c, \alpha D(\lambda,\mu)c) = (P^*c, D(\lambda/\alpha,\mu/\alpha)c)$, because $(P^*c, D(\lambda,\mu)c)$ is a solution to $\phi + (\lambda+\mu)v_i = c_i + \lambda v_{\min\{i+1,N\}} + \mu v_{(i-1)^+}$ (even if $\lambda+\mu > 1$, as is easily checked).

The uniformization variable α can be seen as the parameter of the time between two epochs, thus $\frac{1}{\alpha}$ is the expected sojourn time. If $\alpha = 1$ and by taking $c_i = \delta_j$, $d_{ij}(\lambda,\mu)$ can be seen as the expected difference in time that is spent in j between starting in i and starting in equilibrium. As α is a scale parameter, this expression should be divided by α if $\alpha \neq 1$. Because $D(\lambda,\mu)/\alpha = D(\lambda\alpha,\mu\alpha)$, we find the following.

Corollary 2.2 *The deviation matrix of the $M/M/1/\infty$ queue (with $\lambda < \mu$) and the $M/M/1/N$ queue, given by (1) and (2)-(3) respectively, has the following interpretation: d_{ij} is the expected total amount of time that the system can be found in state j, starting from i, compared to the expected amount of time that the system is in j, starting from equilibrium.*

3 Applications to controlled queueing problems

In this section we give two simple applications to optimizations problems of the use of the deviation matrix. The general idea is as follows: Assume that a certain control model behaves as a single server queue for a fixed policy, then compute the bias using the expressions for D. This allows us to do one step of policy iteration, which gives a certainly better and hopefully very good policy, without having to compute the bias in a iterative way.

The first example is that of a single queue with admission control. An arriving customer can either be sent to the queue where waiting costs w_j are incurred, while rejecting a customer costs C. The objective is to minimize the average costs. The optimality equation for this system is (after uniformization, see [5]):

$$\phi + (\lambda+\mu)v_i = \lambda \min\{C+v_i, w_i+v_{i+1}\} + \mu v_{(i-1)^+},\ i \geq 0.$$

Solving this system is typically done with an iterative procedure such as dynamic programming. Now suppose that we are allowed to take a randomized action, i.e., the optimality equation is

$$\phi + (\lambda+\mu)v_i = \lambda \min_{a\in[0,\lambda]} \{(\lambda - a)(C+v_i) + a(w_i + v_{i+1})\} + \mu v_{(i-1)^+}.$$

Consider the policy that splits the arrival stream in two, such that there are arrivals at rate λ' at the queue, and arrivals are rejected at rate $\lambda - \lambda'$, for some $0 \leq \lambda' \leq \lambda$. This corresponds to taking $a = \lambda'$ in each state i. We call this a *Bernoulli policy*. The optimality equation consists of

$$\phi' + (\lambda'+\mu)v_i = \lambda' w_i + \lambda' v_{i+1} + \mu v_{(i-1)^+}$$

for some ϕ', and $C(\lambda - \lambda')$ added to both sides of the equation. The formula is the optimality equation for a single server queue with rates λ' and μ, and we can compute ϕ' and v_i from the expressions found in the previous section. We did this, and using value iteration we computed the average costs for all policies, including the optimal one. To make the calculations possible we had to truncate the state space. We did this at a sufficiently high level. In Table 1 we see the results of our numerical experiments. For all cases we took $\lambda = 3$ and $\mu = 4$. In the table we see first the two input parameters that we varied, the initial value of λ' and the value of its Bernoulli policy. After that the improved policy is given. Every policy we found had a threshold form, i.e., there is some number T such that if there are less than T customers then an arriving customer is admitted, at or above T the customer is rejected. Thus T is the maximum number of customers that can be in the queue. With T we denote the threshold level of the improved policy, T^* is the threshold level of the optimal policy. Their values are given by ϕ_T and ϕ_{T^*}. It is interesting to note that a bad Bernoulli policy not always gives a good improved policy.

w_i	C	λ'	$\phi_{\lambda'}$	T	ϕ_T	T^*	ϕ_{T^*}
i	10	0	30.00	11	7.78	5	6.18
i	10	2	12.00	5	6.18	5	6.18
i	10	3	9.00	2	8.27	5	6.18
i	50	2	52.00	25	8.97	15	8.86
i^2	10	2	16.00	2	8.27	3	7.91
i^2	50	2	56.00	4	22.24	5	21.64

Table 1. Results for the admission control model, for $\lambda = 3$ and $\mu = 4$.

The second example consists of two parallel $M/M/1/N$ queues, which have in addition to their own dedicated arrival streams arrivals which have to be routed to one of the queues. The objective is simply to minimize the average number of blocked customers. Again, we derive first the bias for the Bernoulli policy that splits the assignable stream with rate γ into two stream with rates γ_1 and γ_2, where rate γ_k goes to queue k. Furthermore, queue k has arrival rate λ_k, service μ_k, and buffer size N_k. Thus the Bernoulli policy has optimality equation

$$\phi + (\lambda_1 + \lambda_2 + \gamma + \mu_1 + \mu_2)v_{i,j} =$$
$$(\lambda_1 + \gamma_1)I\{i = N_1\} + (\lambda_2 + \gamma_2)I\{j = N_2\} +$$
$$(\lambda_1 + \gamma_1)v_{\min\{i+1,N_1\},j} + (\lambda_2 + \gamma_2)v_{i,\min\{j+1,N_2\}} +$$
$$\mu_1 v_{(i-1)^+,j} + \mu_2 v_{i,(j-1)^+}.$$

Define $\rho_k = (\lambda_k + \gamma_k)/\mu_k$. It is easy to see that the solution of the optimality equation is given by

$$\phi = (\lambda_1 + \gamma_1)p^*_{N_1}(\rho_1, N_1) + (\lambda_2 + \gamma_2)p^*_{N_2}(\rho_2, N_2),$$

with $p^*_i(\rho, N)$ the stationary probability of state i of a queue with load ρ and buffer N, and

$$v_{i,j} = (\lambda_1 + \gamma_1)d_{iN_1}(\lambda_1 + \gamma_1, \mu_1, N_1)+$$
$$(\lambda_2 + \gamma_2)d_{jN_2}(\lambda_2 + \gamma_2, \mu_2, N_2).$$

The minimizing action in the policy improvement step is the one which attains the minimum in $\min\{v_{i+1,j}, v_{i,j+1}\}$. The policies that we find, both the policy resulting from the single step of policy improvement and the optimal one, have increasing switching curves. Numerical results can be found in Table 2. With ϕ_B we denote the average costs under the Bernoulli policy, ϕ' are the costs after one step of policy improvement, and ϕ^* are the minimal costs. We define $\lambda = (\lambda_1, \lambda_2)$, $\gamma = (\gamma_1, \gamma_2)$, $\mu = (\mu_1, \mu_2)$, and $N = (N_1, N_2)$. The results for the admission control model were not that impressive; for the routing model the method works very well, as the numbers show. Note however that the choice of Bernoulli policy plays again an important role.

λ	γ	μ	N	ϕ_B	ϕ'	ϕ^*
(1,1)	(1,1)	(1,1)	(10,5)	2.0163	2.0028	2.0028
(1,1)	(1,1)	(2,2)	(10,5)	0.5151	0.3292	0.3259
(1,1)	(1,1)	(3,3)	(10,5)	0.1079	0.0174	0.0142
(1,1)	(1,1)	(10,10)	(10,5)	0.0005	0.0000	0.0000
(1,1)	(1,1)	(3,3)	(10,10)	0.0233	0.0012	0.0012
(1,1)	(2,0)	(3,3)	(10,10)	0.2727	0.0347	0.0012
(1,0)	(0,1)	(2,2)	(10,10)	0.0009	0.0006	0.0004

Table 2. Results for the assignment model.

Finally we say a few words about the complexity of the methods. As said in the Introduction, computing the optimal policy becomes numerically difficult if there are more than 2 queues: for n $M/M/1/N$ queues the number of states is equal to $(N+1)^n$, thus the complexity is exponential in the number of queues. The single step of policy improvement however has a linear complexity if it is implemented in a smart way (by looking at differences of the form $v_{i+1j} - v_{ij}$).

References

[1] A. Hordijk and F.M. Spieksma. A new formula for the deviation matrix. In F.P. Kelly, editor, *Probability, Statistics and Optimisation*, pages 497–507. Wiley, 1994.

[2] T.J. Ott and K.R. Krishnan. Separable routing: A scheme for state-dependent routing of circuit switched telephone traffic. *Annals of Operations Research*, 35:43–68, 1992.

[3] M.L. Puterman. *Markov Decision Processes.* Wiley, 1994.

[4] S.A.E. Sassen, H.C. Tijms, and R.D. Nobel. A heuristic rule for routing customers to parallel servers. *Statistica Neerlandica*, 51:107–121, 1997.

[5] R.F. Serfozo. An equivalence between continuous and discrete time Markov decision processes. *Operations Research*, 27:616–620, 1979.