

A queuing model with arrivals according to a two-dimensional diffusion process

Mario Lefebvre

Abstract

In the case of heavy traffic, the arrival process in a queuing model can be approximated by a diffusion process. In this paper, we consider such a model. The number $X(t)$ of customers in the system at time t evolves according to a degenerate two-dimensional diffusion process. In a particular case, the distribution of $X(t)$ is calculated explicitly. Moreover, a stochastic control problem known as a homing problem is formulated, and the equation satisfied by the value function is derived.

Keywords: arrival process, Gaussian process, stochastic optimal control, first-passage time, dynamic programming.

1 Introduction

In classical queuing models, the arrivals of customers constitute a discrete-space stochastic process. For instance, in the case of the $M/M/k$ model, customers arrive according to a Poisson process. However, in some applications, the arrivals occur almost according to a diffusion process; in particular, in the case of *heavy traffic* (for instance, on the Internet).

In this paper, we assume that the number $X(t)$ of customers in the system at time t is such that

$$dX(t) = \rho[Y(t)]dt - kX(t)dt, \quad (1)$$

$$dY(t) = m[Y(t)]dt + \{v[Y(t)]\}^{1/2}dB(t), \quad (2)$$

where $\{B(t), t \geq 0\}$ is a one-dimensional standard Brownian motion. The functions $m(\cdot) \in \mathbb{R}$ and $v(\cdot) > 0$ are such that $\{Y(t), t \geq 0\}$ is a diffusion process. Moreover, the non-negative constant k is the rate at which the customers are served. The function $\rho(\cdot)$ should be such that if $k = 0$, then $X(t)$ will increase with time t .

This type of degenerate two-dimensional diffusion process, which was proposed by Rishel [1], has been used in reliability theory to model the wear of devices. Indeed, wear should be strictly increasing with time.

In the next section, a particular case for the various functions in Eqs. (1) and (2) will be considered. Then, in Section 3, an optimal control problem for the two-dimensional process $\{(X(t), Y(t)), t \geq 0\}$ will be studied.

2 A particular case

Suppose that $m(\cdot) \equiv \mu > 0$ and $v(\cdot) \equiv \sigma^2$. Then, $\{Y(t), t \geq 0\}$ is a Wiener process with positive drift μ and dispersion parameter $\sigma > 0$. A Wiener process, being a Gaussian process, can take both positive and negative values. Therefore, the function $\rho(\cdot)$ should be such that $\rho[Y(t)]$ is a non-negative function. For example, we could take $\rho[Y(t)] = Y^2(t)$. However, if we assume that $y := Y(0)$ and μ are both large enough, and that σ is small, then the probability that $Y(t)$ becomes negative is negligible.

We choose $\rho[Y(t)] = cY(t)$, with c being a positive constant. With this choice, we can appeal to the following proposition to compute the joint probability density function of the random vector $(X(t), Y(t))$.

Proposition 1. (See [2].) *Let $\{\mathbf{X}(t), t \geq 0\}$ be an n -dimensional stochastic process defined by*

$$d\mathbf{X}(t) = (\mathbf{A}\mathbf{X}(t) + \mathbf{a})dt + \mathbf{N}^{1/2}d\mathbf{B}(t), \quad (3)$$

where $\{\mathbf{B}(t), t \geq 0\}$ is an n -dimensional standard Brownian motion, $\mathbf{A}$ is a square matrix of order n , $\mathbf{a}$ is an n -dimensional vector, and

$\mathbf{N}^{1/2}$ is a positive definite square matrix of order n . Then, given that $\mathbf{X}(t_0) = \mathbf{x}$, we may write that

$$\mathbf{X}(t) \sim \mathbf{N}(\mathbf{m}(t), \mathbf{K}(t)) \quad \text{for } t \geq t_0, \quad (4)$$

where

$$\mathbf{m}(t) := \Phi(t) \left(\mathbf{x} + \int_{t_0}^t \Phi^{-1}(u) \mathbf{a} du \right) \quad (5)$$

and

$$\mathbf{K}(t) := \Phi(t) \left(\int_{t_0}^t \Phi^{-1}(u) \mathbf{N} [\Phi^{-1}(u)]' du \right) \Phi'(t), \quad (6)$$

where the symbol prime denotes the transpose of the matrix, and the function $\Phi(t)$ is given by

$$\Phi(t) := e^{\mathbf{A}(t-t_0)} = \sum_{n=0}^{\infty} \mathbf{A}^n \frac{(t-t_0)^n}{n!}. \quad (7)$$

In our case, we have

$$\mathbf{A} = \begin{bmatrix} -k & c \\ 0 & 0 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 0 \\ \mu \end{bmatrix} \quad \text{and} \quad \mathbf{N}^{1/2} = \begin{bmatrix} 0 & 0 \\ 0 & \sigma \end{bmatrix}. \quad (8)$$

Moreover, for $n \geq 1$,

$$\mathbf{A}^n = \begin{bmatrix} (-k)^n & -c(-k)^{n-1} \\ 0 & 0 \end{bmatrix}. \quad (9)$$

Hence, if $t_0 = 0$, the function $\Phi(t)$ is given by

$$\begin{aligned} \Phi(t) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -1 + e^{-kt} & -\frac{c}{k}(-1 + e^{-kt}) \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} e^{-kt} & -\frac{c}{k}(-1 + e^{-kt}) \\ 0 & 1 \end{bmatrix}. \end{aligned} \quad (10)$$

Next, we find that

$$\int_0^t \Phi^{-1}(u) \mathbf{a} du = \begin{bmatrix} \frac{c\mu(-kt-1+e^{kt})}{k^2} \\ \mu t \end{bmatrix}. \quad (11)$$

It follows that

$$\mathbf{m}(t) = \begin{bmatrix} e^{-k t} \left(x + \frac{c \mu (-k t - 1 + e^{k t})}{k^2} \right) - \frac{c (e^{-k t} - 1)(\mu t + y)}{k} \\ \mu t + y \end{bmatrix}. \quad (12)$$

Finally, with

$$\mathbf{N} = \begin{bmatrix} 0 & 0 \\ 0 & \sigma^2 \end{bmatrix}, \quad (13)$$

we calculate

$$\begin{aligned} & \int_0^t \Phi^{-1}(u) \mathbf{N} [\Phi^{-1}(u)]' du = \\ & \begin{bmatrix} -\frac{c^2 \sigma^2 (-2 k t + 4 e^{k t} - e^{2 k t} - 3)}{2 k^3} & -\frac{c \sigma^2 (-k t - 1 + e^{k t})}{k^2} \\ -\frac{c \sigma^2 (-k t - 1 + e^{k t})}{k^2} & \sigma^2 t \end{bmatrix} \end{aligned} \quad (14)$$

and

$$\mathbf{K}(t) = \begin{bmatrix} \frac{c^2 \sigma^2 (2 e^{2 k t} k t - 3 e^{2 k t} + 4 e^{k t} - 1) e^{-2 k t}}{2 k^3} & \frac{c \sigma^2 (k t - 1 + e^{-k t})}{k^2} \\ \frac{c \sigma^2 (k t - 1 + e^{-k t})}{k^2} & \sigma^2 t \end{bmatrix}. \quad (15)$$

Notice that for t large, $X(t)$ has a Gaussian distribution with mean and variance that are both proportional to t .

Making use of the above results, we can compute the probability that $X(t)$ will become equal to zero as a function of time, so that the queue is empty. Similarly, if the system capacity is finite, we can easily compute the probability that the system will become saturated.

The actual number of customers in the system is given by

$$X_r(t) := \begin{cases} 0 & \text{if } X(t) \leq 0, \\ X(t) & \text{if } 0 < X(t) < r, \\ r & \text{if } X(t) \geq r, \end{cases} \quad (16)$$

where r is the system capacity.

3 A homing problem

In this section, we suppose that the constant k in Eq. (1) is replaced by the function $b_0 u[X(t), Y(t)]$, where b_0 is a positive constant and $u(\cdot, \cdot)$ is a control variable that is assumed to be a continuous function.

Let

$$\tau(x, y) := \inf\{t > 0 : X(t) = 0 \text{ or } Y(t) = \gamma \mid X(0) = x, Y(0) = y\}, \quad (17)$$

where $x > 0$ and $y > \gamma$, and $\gamma \geq 0$ is assumed to be small. The random variable $\tau(x, y)$ is called a *first-passage time* in probability theory.

Our aim is to find the control that minimizes the expected value of the cost function

$$J(x, y) := \int_0^{\tau(x, y)} \left\{ \frac{1}{2} q_0 u^2[X(t), Y(t)] + \theta \right\} dt, \quad (18)$$

where q_0 and θ are positive constants. Hence, the optimizer tries to empty the queue as soon as possible, while taking the quadratic control costs into account. We also stop controlling the process if the arrival rate of customers becomes small enough. This type of stochastic control problem is known as a *homing problem*; see Whittle [3] and [1]. The author has written several papers on homing problems; see, for example, [4].

To find the optimal control $u^*[X(t), Y(t)]$, we can try using dynamic programming, which enables us to express u^* in terms of the value function

$$F(x, y) := \inf_{\substack{u[X(t), Y(t)] \\ t \in [0, \tau(x, y))}} E[J(x, y)]. \quad (19)$$

The function $F(x, y)$ gives the smallest expected cost, starting from $X(0) = x$ and $Y(0) = y$.

We can prove the following proposition.

Proposition 2. *The value function $F(x, y)$ satisfies the dynamic programming equation*

$$0 = \inf_{u(x, y)} \left\{ \frac{1}{2} q_0 u^2(x, y) + \theta + [\rho(y) - b_0 u(x, y)] F_x(x, y) + m(y) F_y(x, y) + \frac{1}{2} v(y) F_{yy}(x, y) \right\}, \quad (20)$$

where $F_x = \frac{\partial}{\partial x} F(x, y)$, etc. The equation is valid for $x > 0$ and $y > \gamma$. We have the boundary conditions $F(0, y) = F(x, \gamma) = 0$. Moreover, the optimal control is given by

$$u^*(x, y) = \frac{b_0}{q_0} F_x(x, y). \quad (21)$$

If we substitute the expression for $u^*(x, y)$ into the dynamic programming equation, we find that to obtain the value function, we must solve the second-order non-linear partial differential equation (PDE)

$$\theta - \frac{b_0}{2q_0} [F_x(x, y)]^2 + \rho(y) F_x(x, y) + m(y) F_y(x, y) + \frac{1}{2} v(y) F_{yy}(x, y) = 0. \quad (22)$$

Assume now that instead of the service rate, the optimizer can control the arrival rate of the process, so that the two-dimensional process $(X(t), Y(t))$ is defined by the system of stochastic differential equations

$$dX(t) = \rho[Y(t)]dt - kX(t)dt, \quad (23)$$

$$dY(t) = b_0 u[X(t), Y(t)]dt + m[Y(t)]dt + \{v[Y(t)]\}^{1/2} dB(t). \quad (24)$$

Proceeding as above, we obtain the following corollary.

Corollary 1. *In the case of the controlled process defined in Eqs. (23) and (24), the optimal control is given by*

$$u^*(x, y) = -\frac{b_0}{q_0} F_y(x, y). \quad (25)$$

Furthermore, the value function satisfies the PDE

$$\begin{aligned} 0 = & \theta - \frac{b_0}{2q_0} [F_x(x, y)]^2 + [\rho(y) - kx] F_x(x, y) \\ & + m(y) F_y(x, y) + \frac{1}{2} v(y) F_{yy}(x, y). \end{aligned} \quad (26)$$

Finally, in some cases, Eq. (26) can be linearized.

Proposition 3. *Suppose that $v(y) \equiv \sigma^2$. The function*

$$G(x, y) := e^{-\alpha F(x, y)}, \quad (27)$$

where

$$\alpha := \frac{b_0^2}{q_0 \sigma^2}, \quad (28)$$

satisfies the linear second-order PDE

$$\alpha \theta G(x, y) = [\rho(y) - kx] G_x(x, y) + m(y) G_y(x, y) + \frac{1}{2} \sigma^2 G_{yy}(x, y). \quad (29)$$

The boundary conditions are $G(0, y) = G(x, \gamma) = 1$.

Remark 1. *The linear PDE in (29) is in fact the Kolmogorov backward equation satisfied by the moment-generating function of the random variable $\tau(x, y)$:*

$$M(x, y) := E \left[e^{-s\tau(x, y)} \right], \quad (30)$$

with $s = \alpha \theta > 0$, for the uncontrolled process $(X_0(t), Y_0(t))$ obtained by setting $u[X(t), Y(t)] \equiv 0$ in Eq. (24). Moreover, the boundary conditions are the appropriate ones.

4 Conclusion

In this paper, a queuing model in which customers arrive (approximately) according to a degenerate two-dimensional diffusion process

was studied. The model is such that, in the case of the absence of service, the number of customers in the system is strictly increasing.

In a particular case, we were able to derive the distribution of the number $X(t)$ of customers in the system at time t . It would be interesting to obtain this distribution in other cases.

Then, a stochastic control problem was formulated for the queuing model. We derived the non-linear partial differential equation satisfied by the value function, and we saw that it is sometimes possible to linearize this equation. In a next step, we could try to obtain explicit solutions to special problems.

Acknowledgments. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

- [1] R. Rishel, “Controlled wear processes: modeling optimal control,” *IEEE Trans. Autom. Control.*, vol. 36, no. 9, pp. 1100–1102, 1991. DOI: <https://doi.org/10.1109/CDC.1989.70212>.
- [2] M. Lefebvre, *Applied Stochastic Processes*, New York, NY, USA: Springer, 2007, 382 p.
- [3] P. Whittle, *Optimization Over Time*, Volume I: *Dynamic Programming and Stochastic Control*, Chichester, UK: Wiley, 1982, 317 p.
- [4] M. Lefebvre, “Minimizing or maximizing the first-passage time to a time-dependent boundary,” *Optimization*, vol. 71, no. 2, pp. 387–401, 2022. DOI: <https://doi.org/10.1080/02331934.2021.1914039>.

Mario Lefebvre

ORCID: <https://orcid.org/0000-0002-9451-543X>

Polytechnique Montréal

E-mail: mlefebvre@polymtl.ca