Trabalho de Simulação AD 2019/2
Simulando Filas com Prioridades

25 de Novembro de 2019

1 Introdução e Objetivos

Fila com prioridades podem ser analisadas de forma prática e conveniente usando resultados básicos de teoria de filas, simulação ou cadeias de Markov. O objetivo deste trabalho é comparar resultados de simulação com resultados analíticos, obtidos com fórmulas fechadas e cadeias de Markov. Vamos investigar também a propriedade PASTA, e avaliar quando ela vale e quando ela não vale.

2 Como Resolver uma CMTC no Matlab

Atenção! Nesse trabalho, estaremos lidando com cadeias de Markov infinitas. Entretanto, para resolver a cadeia usando o método abaixo, precisamos truncar o espaço de estados. Ou seja, precisamos considerar matrizes finitas. Faz parte do trabalho a identificação do ponto de corte adequado.

A seguir, ilustramos como resolver a cadeia com a seguinte matriz de taxas no Matlab.

$$Q = \begin{pmatrix} -1 & +1 \\ +2 & -2 \end{pmatrix} \tag{1}$$

```
>> Q=[-1,1;2,-2]

Q =

    -1     1
     2    -2

>> Qt=Q'

Qt =
```

```
    -1     2
     1    -2

>> s=size(Qt,1)

s =

     2

>> Qt(2,:)=1

Qt =

    -1     2
     1     1


>> e=zeros(s,1)

e =

     0
     0

>> e(s)=1

e =

     0
     1



>> Qt\e

ans =

    0.6667
    0.3333
```

3 Filas Sem Prioridades (Fila Única)

Considere dois fluxos Poisson com taxas λ_1 e λ_2 chegando a uma fila. O fluxo com taxa λ_i está associado a um tempo médio de serviço $E[X_i] = 1/\mu_i$.

Simule os seguintes cenários:

Cenário	tempo de serviço	λ_1	λ_2	serviço 1 parâmetros	serviço 2 parâmetros
1	exp. com taxa μ_i	$0.05, 0.1, 0.15, \ldots, 0.9$	0	$\mu_1 = 1$	$\mu_2 = 0$
2	exp. com taxa μ_i	$0.05, 0.1, 0.15, \ldots, 0.6$	0.2	$\mu_1 = 1$	$\mu_2 = 0.5$
3	determ., média $1/\mu_i$	$0.05, 0.1, 0.15, \ldots, 0.6$	0.2	$\mu_1 = 1$	$\mu_2 = 0.5$
4	unif., intervalo $[a, b]$	0.08	0.05	$[a_1 = 5, b_1 = 15]$	$[a_2 = 1, b_2 = 3]$

Tabela 1: Cenários a simular

Considere os casos apresentados na Tabela 1. Usando simulação, plote os seguintes gráficos, com intervalo de confiança de 95%

1. para cada cenário, eixo x: λ, eixo y: número médio de clientes na fila de espera (sem considerar servidor). No mesmo gráfico, super imponha uma curva indicando os valores analíticos (resultados M/M/1 com fórmulas vistas em aula). Ainda no mesmo gráfico, super imponha uma curva indicando o valor obtido usando uma cadeia de Markov *finita* com número de estados suficientes para aproximar bem o valor analítico obtido. Mostre que todas as soluções convergem para o mesmo valor: a) simulação, b) analítico fórmula fechada e c) cadeia de Markov. Para resolver a cadeia de Markov, você pode usar eliminação gaussiana (scilab ou matlab, por exemplo). O importante é que você construa a matriz de transição de taxas (finita) e resolva a cadeia numericamente como indicado na seção 2.

2. repita os itens acima, trocando o eixo y para: tempo médio de permanência dos clientes na fila de espera (sem considerar servidor).

4 Filas Com Prioridades Sem Preempção

Repita a questão acima, mas considerando agora filas com prioridades sem preempção. Assuma que a classe 1 tem prioridade sobre a classe 2.

5 Lei da Conservação

A fórmula do trabalho pendente é dada por:

$$E[U] = \rho E[X_r]/(1 - \rho) \tag{2}$$

onde as quantidades acima se referem a uma fila virtual na qual todos os clientes são tratados por igual, ou seja, conforme o sistema analisado na Seção 3.

Obtenha via simulação o trabalho pendente encontrado por um cliente no sistema nos cenários 2 e 3, nos casos a) fila única, b) com preempção e c) sem preempção. Compare sua solução com a solução analítica. Mostre que em suas simulações essa constante acima se mantém inalterada independente do sistema ser de fila única, ou duas filas com ou sem preempção.

Atenção! Tome cuidado ao calcular o trabalho pendente no sistema. Em particular, reflita bem sobre como calcular o intervalo de confiança do trabalho pendente. Compare seus resultados obtidos via simulações com o resultado analítico.

Sugestão. Para calcular o intervalo de confiança do trabalho pendente, você pode considerar uma das seguintes duas soluções

1. solução 1: amostre o tempo de serviço associado a cada cliente quando ele chegar ao sistema. Guarde essa informação junto com o cliente, enquanto ele estiver no sistema. Essa não é a forma padrão de implementar uma fila com prioridades, mas pode simplificar o cálculo do trabalho pendente

2. solução 2: considere uma nova chegada de um cliente. Seja U_S o trabalho pendente do cliente que se encontra no servidor (note que $U_S = 0$ caso o servidor esteja vazio). Seja U_2 o trabalho pendente do cliente do tipo 2 que se encontra na cabeça da fila 2. Toda vez que um novo cliente chegar ao sistema, tire uma amostra da quantidade $N_{q1}E[X_1] + \max(0, N_{q2} - 1)E[X_2] + U_2 + U_S$ lembrando que $E[X_1]$ e $E[X_2]$ são constantes. Note que N_{q1}, N_{q2}, U_2 e U_S são amostras do número de clientes na fila 1, na fila 2, do tempo pendente do cliente na cabeça da fila 2 e do tempo pendente do cliente no servidor, sendo essas quatro quantidades visualizadas pelo cliente que acabou de chegar. Depois, trate as amostras como de costume para obter o intervalo de confiança.

Se quiser, para fins de depuração você pode implementar as duas soluções acima, e mostrar que elas geram o mesmo resultado. Mas isso não é obrigatório. Escolha a opção que julgar mais simples, e invente formas de mostrar que os dados obtidos via simulação estão corretos.

Considere agora a seguinte igualdade:

$$E[U] \overset{(?)}{=} E[N_{q1}]E[X_1] + E[N_{q2}]E[X_2] + \rho E[X_r] \tag{3}$$

Verifique, analiticamente e usando simulações dos cenários 2 e 3, em que casos esta igualdade é válida. Para isto, construa uma tabela, indicando

1. coluna 1: o valor do lado esquerdo, $E[U]$, calculado usando (2)

2. coluna 2: o valor do lado direito de (3), calculado analiticamente

3. coluna 3: o valor de $E[N_{q1}]$, obtido via simulação (com intervalos de confiança)

4. coluna 4: o valor de $E[N_{q2}]$, obtido via simulação (com intervalos de confiança)

5. coluna 5: o valor total do lado direito de (3), sendo $E[N_{q1}]$ e $E[N_{q2}]$ obtidos via simulação e o restante das quantidades obtidas analiticamente. Para calcular o intervalo de confiança nesse caso, faça o seguinte: toda vez que um novo cliente chegar ao sistema, tire uma amostra da quantidade $N_{q1}E[X_1] + N_{q2}E[X_2] + \rho E[X_r]$ lembrando que $E[X_1]$, $E[X_2]$, ρ e $E[X_r]$ são constantes. Depois, trate as amostras como de costume para obter o intervalo de confiança.

Para cada conjunto de parâmetros, construa uma tabela para o caso fila única, uma para o caso sem preempção, e uma para o caso com preempção.

Caso você acabe obtendo muitas tabelas, selecione aquelas que julgar mais ilustrativas para obter suas conclusões, e apresente apenas elas.

Usando as resposta para os itens acima, conclua em que casos podemos igualar (2) e (3) para obter a lei da conservação. Escreva a lei da conservação, e discuta suas implicações.

6 Simulação para Verificar Resultados Analíticos

Resolva a primeira questão da última prova via simulação (a questão também encontra-se abaixo, na Seção 8). Para os cenários em que a resposta correta for "verdadeiro", indique via simulação que o resultado analítico está de acordo com simulação. Para os cenários em que a resposta for "falso", selecione um conjunto de parâmetros que claramente indiquem que a questão está errada. Ajuste os parâmetros conforme julgar adequado – ou seja, você tem liberdade total para determinar os valores de λ_1, λ_2, $E(X_1)$ etc. Nessa questão, você não precisa se ater à Tabela 1. Analise sua solução.

7 Quesitos Que Devem Constar no Relatório

1. Desafios encontrados e soluções

2. Todas as soluções obtidas via simulação devem vir acompanhadas de seus intervalos de confiança

3. Decisões de projeto: breve descrição do simulador, e de como foi implementado

4. Depuração (como o simulador foi depurado). Em particular

   (a) quando possível, comparar analítico com simulação e verificar que analítico está no intervalo de confiança da simulação
   (b) simplificar o problema, e estudar casos básicos
   (c) colocar *printf* no código para rastrear o que ocorre ao longo do processo

5. Resultados

6. Discussão dos resultados

8 Filas com prioridades sem preempção e lei da conservação (questão da prova 2 de 2019/2)

Considere uma chegada de um cliente típico a uma fila que serve clientes de 2 classes (classes 1 e 2). O escalonamento dentro de cada classe é FCFS, com prioridade da classe 1 sobre a classe 2, sem preempção. Seja $E[U]$ o trabalho pendente no sistema, no instante de chegada do cliente.

Seja $E[W]$ o tempo médio de espera na fila de espera de um cliente típico deste sistema.

Seja $E[G]$ o período ocupado do sistema. Seja $\lambda = \lambda_1 + \lambda_2$, $p_1 = \lambda_1/\lambda$ e $p_2 = \lambda_2/\lambda$. Denotamos por $E[X_{r1}]$ e $E[X_{r2}]$ a vida residual do serviço dos clientes do tipo 1 e do tipo 2, respectivamente, e por $E[X_r]$ a vida residual do tempo de serviço do servidor, levando em conta clientes do tipo 1 e do tipo 2. Denotamos por ρ a utilização do sistema. Denotamos por $E[W_1]$ e $E[W_2]$ o tempo médio na fila de espera dos clientes do tipo 1 e do tipo 2, respectivamente.

Certo ou errado, justifique. Atenção! Faz parte da solução completa do problema você definir todas as quantidades de interesse que porventura venha a utilizar, em função de quantidades básicas, ou seja, em função de λ_1, λ_2, $E[X_1]$, $E[X_2]$, $E[X_{r1}]$, $E[X_{r2}]$.

1. (5 pontos) $E[W] = p_1 E[W_1] + p_2 E[W_2]$
2. (5 pontos) $E[W_1] = \rho_1 E[X_{1r}]/(1 - \rho_1)$
3. (5 pontos) $E[W] = \rho E[X_r]/(1 - \rho)$
4. (5 pontos) $E[U] = \rho E[X_r]/(1 - \rho)$
5. (5 pontos) $E[X] = (\rho_1/\rho)E[X_1] + (\rho_2/\rho)E[X_2]$
6. (5 pontos) $E[X_r] = (\rho_1/\rho)E[X_{r1}] + (\rho_2/\rho)E[X_{r2}]$
7. (5 pontos) $E[G] = E[X]/(1 - \rho)$
8. (5 pontos) $E[U] = \rho_1 E[W_1] + \rho_2 E[W_2] + \rho E[X_r] = \lambda_1 E[X_1]E[W_1] + \lambda_2 E[X_2]E[W_2] + \rho E[X_r]$
9. (5 pontos)
$$\rho_1 E[W_1] + \rho_2 E[W_2] + \rho E[X_r] = \frac{\rho_1}{\rho} E[W_1] + \frac{\rho_2}{\rho} E[W_2] = \frac{\rho E[X_r]}{1-\rho} \Rightarrow \rho_1 E[W_1] + \rho_2 E[W_2] = \frac{\rho^2 E[X_r]}{1-\rho}$$
10. (5 pontos) A distribuição do número de clientes no sistema é invariante com relação ao fato de a política de serviço ser FCFS, LCFS sem interrupção, ou FCFS com diferentes classes de clientes, mais uma vez sem interrupção.

11. (5 pontos) A distribuição do tempo dos clientes no sistema é invariante com relação ao fato de a política de serviço ser FCFS, LCFS sem interrupção, ou FCFS com diferentes classes de clientes, mais uma vez sem interrupção.