Structured Markov Chains

Ivo Adan and Johan van Leeuwaarden
Book on Analysis of structured Markov processes (arXiv:1709.09060)

I Basic methods
- Basic Markov processes
- Advanced Markov processes
- Queues and transforms

II Basic processes
- Birth-and-death processes
- Queueing networks
- Quasi-birth-and-death-processes
- Processes that are quasi-skip-free in one direction

III Advanced processes
- Single-server priority systems
- Gated single-server systems
- Production systems
- Join the shortest queue
Markov chains in discrete time

A stochastic process \( \{X_n, n = 0, 1, 2, \ldots \} \) with state space \( \{0, 1, 2, \ldots \} \) is a Markov chain if

\[
P(X_{n+1} = i_{n+1} | X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = P(X_{n+1} = i_{n+1} | X_n = i_n)
\]

for all \( i_0, \ldots, i_{n+1} \) and \( n \geq 0 \)

- Assumptions:
  - Time homogeneous: \( P(X_{n+1} = j | X_n = i) = p_{ij} \)
  - Irreducible: all states are reachable
  - Aperiodic

- Equilibrium distribution \( p_0, p_1, \ldots \) satisfy the equilibrium (or balance) equations

\[
p_i = \sum_{j=0}^{\infty} p_j p_{ji}, \quad i = 0, 1, \ldots
\]

- If equilibrium distribution exists then it is equal to the limiting distribution \( \lim_{n \to \infty} P(X_n = i) = p_i \)

- Vector-matrix notation: \( p = pP \) where \( p = (p_0, p_1, \ldots) \) and transition probability matrix \( P = (p_{ij}) \)
Markov chains in continuous time

A stochastic process \( \{X(t), t \geq 0\} \) with state space \( \{0, 1, 2, \ldots\} \) is a Markov chain if

\[
P(X(t + s) = j | X(u) = x(u), 0 \leq u < s, X(s) = i) = P(X(t + s) = j | X(s) = i)
\]

for all \( i, j, x(u) \) and \( s, t \geq 0 \)

- **Assumptions:**
  - **Time homogeneous:** \( P(X(t + s) = j | X(s) = i) = p_{ij}(t) \)
  - **Irreducible:** all states are reachable
- **Transition rates** \( q_{ij} = \lim_{h \to 0} \frac{p_{ij}(h)}{h} \) where we let \( q_{ii} = 0 \) and \( q_i = \sum_{j=0}^{\infty} q_{ij} \)
- **Equilibrium distribution** \( p_0, p_1, \ldots \) satisfy the equilibrium (or balance) equations

\[
p_i q_i = \sum_{j=0}^{\infty} p_j q_{ij}, \quad i = 0, 1, ...
\]

- If equilibrium distribution exists then it is equal to the limiting distribution \( \lim_{t \to \infty} P(X(t) = i) = p_i \)
- **Vector-matrix notation:** \( 0 = pQ \) where \( p = (p_0, p_1, \ldots) \) and generator \( Q \) with \( Q_{ij} = q_{ij}, i \neq j, Q_{ii} = -q_i \)
Balance equations

- Global balance equations:

\[ p_i q_i = \sum_{j=0}^{\infty} p_j q_{ji} \]

or for subset \( A \) and its complement \( A^c \)

\[ \sum_{i \in A} p_i q_{ij} = \sum_{j \in A^c} p_j q_{ji} \]

- Partial balance equations:

\[ p_i \sum_{j \in A} q_{ij} = \sum_{j \in B} p_j q_{ji} \]

for subsets \( A \) and \( B \)

- Detailed (or local) balance equations:

\[ p_i q_{ij} = p_j q_{ji} \]
Balance equations

- Global balance equations:

- Partial balance equations:

- Detailed (or local) balance equations:
Transforms

- Generating function of non-negative discrete random variable $X$ with probability distribution $p_i = P(X = i)$

$$P_X(z) = E(z^X) = \sum_{i=0}^{\infty} p_i z^i, \quad |z| \leq 1$$

- Laplace-Stieltjes transform a non-negative (continuous) random variable $X$ with distribution function $F_X(x)$

$$X(s) = E(e^{-sx}) = \int_0^\infty e^{-sx} dF_X(x), \quad \text{Re}(s) \geq 0$$

- Properties:
  - Probability distribution $p_i$ is uniquely determined by generating function $P_X(z)$
  - Distribution function $F_X(x)$ is uniquely determined by Laplace-Stieltjes $X(s)$
  - $E(X(X - 1) \cdots (X - n + 1)) = P_X^{(n)}(1)$ and $E(X^n) = (-1)^n X^{(n)}(0)$ where $^{(n)}$ is $n$th derivative
Transforms: Example

- Markov chain \( \{X_n, n = 0, 1, 2, \ldots \} \) with state space \( \{0, 1, 2, \ldots, n + 1\} \) and transition probabilities \( p_{ij} \)
- State \( n + 1 \) is single absorbing state

\[
\begin{array}{c}
p_{11} = \frac{1}{2} & p_{22} = \frac{1}{2} \\
p_{21} = \frac{2}{3} & p_{33} = 1
\end{array}
\]

- Random variable \( N_i \) is number of visits to state \( i \) before absorption into \( n + 1 \)
- Time spent in state \( i \) during the \( k \)th visit is equal to \( \alpha_i^{k-1} \) where \( 0 < \alpha_i < 1 \) (learning curve)
- Random variable \( T \) is total time spent in states \( \{1, \ldots, n\} \) before absorption in \( n + 1 \)
- What is \( E(T|X_0 = 1) \)?
Transforms: Example

- Note

\[ T = \sum_{i=1}^{n} \left( 1 + \alpha_i + \cdots + \alpha_i^{N_i-1} \right) = \sum_{i=1}^{n} \frac{1 - \alpha_i^{N_i}}{1 - \alpha_i} \]

so

\[ E(T|X_0 = 1) = \sum_{i=1}^{n} \frac{1 - E(\alpha_i^{N_i}|X_0 = 1)}{1 - \alpha_i} \]

- Introduce for \( i = 1, \ldots, n \) the generating function

\[ P_i(z) = E(z_1^{N_i} \cdots z_n^{N_n}|X_0 = i), \quad |z_1| \leq 1, \ldots, |z_n| \leq 1, \]

where \( z = (z_1, \ldots, z_n) \) so

\[ P_1(1, \ldots, 1, \alpha_i, 1, \ldots, 1) = E(\alpha_i^{N_i}|X_0 = 1) \]
Transforms: Example

- One-step analysis:

\[ P_i(z) = E(z_1^{N_1} \cdots z_n^{N_n}|X_0 = i) \]

\[ = \sum_{j=1}^{n+1} p_{ij} E(z_1^{N_1} \cdots z_i^{1+N_i} \cdots z_n^{N_n}|X_0 = j) \]

\[ = \sum_{j=1}^{n+1} z_i p_{ij} E(z_1^{N_1} \cdots z_i^{N_i} \cdots z_n^{N_n}|X_0 = j) \]

\[ = \sum_{j=1}^{n+1} z_i p_{ij} P_j(z), \quad i = 1, \ldots, n \]

where \( P_{n+1}(z) = 1 \) (since \( N_1 = \cdots = N_n = 0 \) if \( X_0 = n + 1 \))

- Hence \( P_1(1, \ldots, 1, \alpha_i, 1, \ldots, 1) \) can be solved from above equations with \( z = (1, \ldots, 1, \alpha_i, 1, \ldots, 1) \)
Finite Markov chains (no structure)

- Markov process \( \{X(t), t \geq 0\} \) with state space \( \{0, 1, \ldots, N\} \) and transition rates \( q_{ij} \)
- Global balance equations

\[
p_i \sum_{j=0}^{N} q_{ij} = \sum_{j=0}^{N} p_j q_{ji}, \quad i = 0, 1, \ldots, N
\]
Finite Markov chains (no structure)

- Numerical solution through embedding:
  - Remove state $N$ by embedding Markov chain on \{0, 1, ..., $N - 1$\} with transition rates
    \[
    \tilde{q}_{ij} = q_{ij} + q_{iN} \frac{q_{Nj}}{q_{N0} + \cdots + q_{NN-1}}, \quad i, j = 0, ..., N - 1
    \]
  - Equilibrium probability $p_N$ follows from global balance
    \[
    p_N \sum_{j=0}^{N-1} q_{Nj} = \sum_{j=0}^{N-1} p_j q_{jN}
    \]
  - Repeat this for $N - 1, N - 2, \ldots, 1$ and finally normalize $p_0, \ldots, p_N$

- This is Gaussian elimination with multiplication, division and addition of positive numbers
Infinite Markov chains (with structure): Example of impatient customers

- Single server
- Poisson arrivals with rate $\lambda$
- Exponential service times with rate $\mu$
- Exponential patience with rate $\theta$: customer leaves if service does not start before patience expires
- Markov chain with states $i$ where $i$ is number in system
- Transition rates $q_{i,i-1} = \mu + (i - 1)\theta$ and $q_{i-1,i} = \lambda$
Infinite Markov chains (with structure): Birth-and-Death processes

- Markov chain on \{0, 1, \ldots\} with transitions to neighboring states with rates \( q_{i,i-1} = \mu_i \) and \( q_{i-1,i} = \lambda_{i-1} \)

- Global balance

\[
\begin{align*}
p_0 \lambda_0 & = p_1 \mu_1 \\
p_i (\lambda_i + \mu_i) & = p_{i-1} \lambda_{i-1} + p_{i+1} \mu_{i+1}, \quad i = 1, 2, \ldots
\end{align*}
\]

- Global balance for \{0, 1, \ldots, i - 1\}

\[
p_{i-1} \lambda_{i-1} = p_i \mu_i, \quad i = 1, 2, \ldots
\]
Infinite Markov chains (with structure): Birth-and-Death processes

- Iterating yields
  \[ p_i = p_{i-1} \frac{\lambda_{i-1}}{\mu_i} = p_{i-2} \frac{\lambda_{i-2} \lambda_{i-1}}{\mu_{i-1} \mu_i} = \cdots = p_0 \frac{\lambda_0 \cdots \lambda_{i-1}}{\mu_1 \cdots \mu_i} \]

- Normalization
  \[ p_0 = \frac{1}{\sum_{i=0}^{\infty} \prod_{j=1}^{i} \frac{\lambda_{j-1}}{\mu_j}} \]

- Assumption
  \[ \sum_{i=0}^{\infty} \prod_{j=1}^{i} \frac{\lambda_{j-1}}{\mu_j} < \infty \]

- Example of impatient customers
  \[ p_i = p_0 \frac{\lambda^i}{\mu(\mu + \theta) \cdots (\mu + (i-1)\theta)} \]
Infinite Markov chains (with more structure): Constant birth and death rates

- Global balance equations

\[ p_0 \lambda = p_1 \mu \]
\[ p_i (\lambda + \mu) = p_{i-1} \lambda + p_{i+1} \mu, \quad i = 1, 2, \ldots \]

- Assume \( \lambda - \mu < 0 \) and (without loss of generality) \( \lambda + \mu = 1 \)

- Generator \( Q \) of the form

\[
Q = \begin{pmatrix}
-\lambda & \lambda & 0 & 0 & \cdots \\
\mu & -(\lambda + \mu) & \lambda & 0 & \cdots \\
0 & \mu & -(\lambda + \mu) & \lambda & \cdots \\
0 & 0 & \mu & -(\lambda + \mu) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\]
Infinite Markov chains (with more structure): Constant birth and death rates

\[
\begin{array}{ccc}
\lambda & \uparrow & \lambda \\
0 & \mu & 1 \\
\mu & \downarrow & \mu \\
\end{array}
\]

\[
\begin{array}{ccc}
\lambda & \uparrow & \lambda \\
i-1 & \mu & i \\
\mu & \downarrow & \mu \\
\end{array}
\]

- Solution: Generating function \( P(z) = \sum_{i=1}^{\infty} p_i z^i \)
  - Multiply balance equations by \( z^i \) and add all equations

\[
P(z)(\lambda + \mu) - p_0 \mu = P(z)\lambda z + (P(z) - p_0)\mu z^{-1}
\]

- Solving this equation

\[
P(z) = \frac{p_0 \mu (1 - z^{-1})}{(\lambda + \mu) - \lambda z - \mu z^{-1}} = \frac{p_0 \mu (1 - z^{-1})}{(1 - z^{-1})(\mu - \lambda z)} = \frac{p_0}{1 - \rho z}
\]

where \( \rho = \frac{\lambda}{\mu} \)

- Normalization \( P(1) = 1 \) so \( p_0 = 1 - \rho \) and

\[
P(z) = \frac{1 - \rho}{1 - \rho z} = \sum_{i=0}^{\infty} (1 - \rho) \rho^i z^i
\]
Infinite Markov chains (with more structure): Constant birth and death rates

Solution: Geometric

- Ratio $p_{i+1}/p_i$ is expected time spent in $i + 1$ before first return to $i$, given initial state is $i$
- Transition structure implies that ratio $p_{i+1}/p_i$ does not depend on $i$

$$\frac{p_{i+1}}{p_i} = \alpha$$

- Hence $p_i = p_0\alpha^i$ and substitution in global balance equation yields

$$\alpha^i(\lambda + \mu) = \alpha^{i-1}\lambda + \alpha^{i+1}\mu$$

- Dividing by common power $\alpha^{i-1}$

$$\alpha(\lambda + \mu) = \lambda + \alpha^2\mu$$

- This leads to $\alpha = \frac{\lambda}{\mu} = \rho$
Infinite Markov chains (with more structure): Constant birth and death rates

• Solution: Difference equation
  - Balance equations are second-order difference equation with constant coefficients
  - Solution is linear combination of powers $p_i = \alpha^i$
  - Substitution of $p_i = \alpha^i$ in balance equation yields

\[
\alpha(\lambda + \mu) = \lambda + \alpha^2\mu
\]

with roots $\alpha_1 = \rho$ and $\alpha_2 = 1$
  - General solution

\[
p_i = c_1\alpha_1^i + c_2\alpha_2^i = c_1\rho^i + c_2
\]

where $c_1$ and $c_2$ are constants
  - Convergence of $p_i$ implies $c_2 = 0$ and normalization yields $c_1 = 1 - \rho$
Infinite Markov chains: $M/G/1$ structure

- Transition probabilities $P = (p_{ij})$

$$P = \begin{pmatrix} b_0 & b_1 & b_2 & b_3 & \cdots \\ a_0 & a_1 & a_2 & a_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

- Example: Number in system just after departure in $M/G/1$ queue

- Let $A(z) = \sum_{i=0}^{\infty} a_iz^i$ and $B(z) = \sum_{i=0}^{\infty} b_iz^i$ and assume mean step size $= A^{(1)}(1) - 1 < 0$

- Global balance equations

$$p_i = p_{i+1}a_0 + p_ia_1 + \cdots + p_1a_i + p_0b_i, \quad i = 0, 1, 2, \ldots$$
Infinite Markov chains: $M/G/1$ structure

- Solution: Recursion
  - Global balance for $\{0, 1, \ldots, i - 1\}$
    \[ p_0 \bar{b}_i + p_1 \bar{a}_{i-1} + \cdots + p_{i-1} \bar{a}_1 = p_i \bar{a}_0, \quad i = 1, 2, \ldots \]
    where
    \[ \bar{b}_i = \sum_{j=i}^{\infty} b_j, \quad \bar{a}_i = \sum_{j=i}^{\infty} a_j \]
  - Recursively calculate $p_1, p_2, \ldots$ starting with $p_0$
  - $p_0$ follows from (mean displacement is 0 in equilibrium)
    \[ p_0 B^{(1)}(1) + (1 - p_0)(A^{(1)}(1) - 1) = 0 \]
Infinite Markov chains: $M/G/1$ structure

### Solution: Generating function

- Let $P(z) = \sum_{i=0}^{\infty} p_i z^i$, $A(z) = \sum_{i=0}^{\infty} a_i z^i$, $B(z) = \sum_{i=0}^{\infty} b_i z^i$
- Multiply balance equations by $z^i$ and add all equations
  \[ P(z) = z^{-1}(P(z) - p_0)A(z) + p_0B(z) \]
- Solving this equation
  \[ P(z) = \frac{p_0(B(z) - z^{-1}A(z))}{1 - z^{-1}A(z)} = \frac{p_0(zB(z) - A(z))}{z - A(z)} \]
- Normalization $P(1) = 1$ so
  \[ p_0 = \frac{1 - A^{(1)}(1)}{B^{(1)}(1) + 1 - A^{(1)}(1)} \]
Infinite Markov chains: $G/M/1$ structure

- Transition probabilities $P = (p_{ij})$
  \[
P = \begin{pmatrix}
    b_0 & a_0 & 0 & 0 & \cdots \\
    b_1 & a_1 & a_0 & 0 & \cdots \\
    b_2 & a_2 & a_1 & a_0 & \cdots \\
    b_3 & a_3 & a_2 & a_1 & \cdots \\
    \vdots & \vdots & \vdots & \vdots & \ddots
  \end{pmatrix}, \quad b_i = 1 - (a_0 + \cdots + a_i) = \sum_{j=i+1}^{\infty} a_i
  \]

- Example: Number in system just before arrival in $G/M/1$ queue

- Global balance equations
  \[
p_i = p_{i-1}a_0 + p_i a_1 + p_{i+1}a_2 + \cdots = \sum_{j=0}^{\infty} p_{i-1+j}a_j, \quad i = 1, 2, \ldots
  \]
Infinite Markov chains: $G/M/1$ structure

- **Solution: Geometric**
  - Ratio $p_{i+1}/p_i$ is expected time spent in $i + 1$ before first return to $i$, given initial state is $i$
  - Transition structure implies that ratio $p_{i+1}/p_i$ does not depend on $i$
    \[
    \frac{p_{i+1}}{p_i} = \alpha
    \]
  - Hence $p_i = p_0\alpha^i$ and substitution in global balance equation yields
    \[
    \alpha = \sum_{j=0}^{\infty} \alpha^j a_j = A(\alpha)
    \]
  - $\alpha$ is the unique root on $(0, 1)$ of $\alpha = A(\alpha)$ provided mean step size $1 - A^{(1)}(1) < 0$
Examples of Quasi-Birth-and-Death processes: Single server with setup

- Single server
- Poisson arrivals with rate $\lambda$ which need exponential service times with rate $\mu$
- Server switched off when system is empty
- Server switched on when first customer arrives which requires exponential setup time with rate $\theta$
Examples of Quasi-Birth-and-Death processes: Single server with setup

- Markov chain with states \((i, j)\) with \(i\) number in system and \(j\) status of server: 0 means off, 1 means on
- Transition rate diagram

\[
\begin{align*}
\lambda & \\
\mu & \\
\theta & \\
\end{align*}
\]

- Global balance equations

\[
\begin{align*}
p(0, 0)\lambda &= p(1, 1)\mu \\
p(0, 1)\lambda &= 0 \\
p(i, 0)(\lambda + \theta) &= p(i - 1, 0)\lambda \\
p(i, 1)(\lambda + \mu) &= p(i - 1, 1)\lambda + p(i, 0)\theta + p(i + 1, 1)\mu, \quad i = 1, 2, \ldots
\end{align*}
\]
Examples of Quasi-Birth-and-Death processes: Single server with setup

- Order states as \((0, 0), (0, 1), (1, 0), (1, 1), (2, 0), (2, 1), \ldots\) and partition into levels \(i = \{(i, 0), (i, 1)\}\)
- Generator \(Q\) of the form

\[
Q = \begin{pmatrix}
B_1 & A_0 & 0 & 0 & \ldots \\
B_2 & A_1 & A_0 & 0 \\
0 & A_2 & A_1 & A_0 \\
\vdots & 0 & A_2 & A_1 \\
& & \ddots & \ddots
\end{pmatrix}
\]

where

\[
B_1 = \begin{pmatrix}
-\lambda & 0 \\
0 & -\lambda
\end{pmatrix}, \quad B_2 = \begin{pmatrix}
0 & 0 \\
\mu & 0
\end{pmatrix}, \quad A_0 = \begin{pmatrix}
\lambda & 0 \\
0 & \lambda
\end{pmatrix}, \quad A_1 = \begin{pmatrix}
-(\lambda + \theta) & \theta \\
0 & -(\lambda + \mu)
\end{pmatrix}, \quad A_2 = \begin{pmatrix}
0 & 0 \\
0 & \mu
\end{pmatrix}
\]

- Let \(p = (p_0, p_1, \ldots)\) and \(p_i = (p(i, 0), p(i, 1))\)

\[
p_0B_1 + p_1B_2 = 0, \\
p_{i-1}A_0 + p_iA_1 + p_{i+1}A_2 = 0, \quad i = 1, 2, \ldots
\]
Examples of Quasi-Birth-and-Death processes: Single server with setup

- Solution: Matrix-geometric
  - Global balance for level $0, 1, \ldots, i$
    \[ p(i + 1, 1)\mu = (p(i, 0) + p(i, 1))\lambda \]
    or in matrix form
    \[ p_{i+1}A_2 = p_iA_3 \]
  - Substituting in balance equation
    \[ 0 = p_{i-1}A_0 + p_iA_1 + p_{i+1}A_2 = p_{i-1}A_0 + p_i(A_1 + A_3) \]
  - Hence
    \[ p_i = -p_{i-1}A_0(A_1 + A_3)^{-1} = p_{i-1}R \]
  - with
    \[ R = -A_0(A_1 + A_3)^{-1} = \begin{pmatrix} \frac{\lambda}{\lambda + \theta} & \frac{\lambda}{\mu} \\ 0 & \frac{\lambda}{\mu} \end{pmatrix} \]
Examples of Quasi-Birth-and-Death processes: Single server with setup

- **Solution: Matrix-geometric**
  - Iterating yields
    \[ p_i = p_0 R^i, \quad i = 0, 1, 2, \ldots \]
  - Vector \( p_0 \) follows from boundary equation
    \[ p_0 B_1 + p_1 B_2 = p_0 (B_1 + RB_2) = 0 \]
    and normalization
    \[ 1 = \sum_{i,j} p(i,j) = \sum_{i=0}^{\infty} p_i e = p_0 \sum_{i=0}^{\infty} R^i e = p_0 (I - R)^{-1} e \]
    where \( I \) is identity matrix and \( e \) all-one vector
Examples of Quasi-Birth-and-Death processes: Single server with setup

- Solution: Matrix-analytic
  - Let $g_{jk}$ be probability of first passage to level $i - 1$ in $(i - 1, k)$ for initial state $(i, j)$ at level $i$

  $$G = (g_{jk}) = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

  - Then

  $$p_i A_2 = p_i A_0 G$$

  - Substituting in balance equation

  $$0 = p_{i-1} A_0 + p_i A_1 + p_{i+1} A_2 = p_{i-1} A_0 + p_i (A_1 + A_0 G)$$

  - Hence

  $$p_i = -p_{i-1} A_0 (A_1 + A_0 G)^{-1} = p_{i-1} R$$
Examples of Quasi-Birth-and-Death processes: Single server with setup

- **Solution: Spectral expansion**
  - Seek solutions of balance equations \( p_{i-1}A_0 + p_iA_1 + p_{i+1}A_2 = 0 \) of the form
    \[
    p_i = y \cdot x^i, \quad i = 0, 1, 2, \ldots,
    \]
    where \( y = (y(0), y(1)) \) is non-null vector and \( |x| < 1 \)
  - Substitution and dividing by common powers of \( x \) yields
    \[
    y (A_0 + xA_1 + x^2A_2) = 0
    \]
  - Desired values of \( x \) are roots inside the unit circle of
    \[
    \det(A_0 + xA_1 + x^2A_2) = (\lambda - (\lambda + \theta)x)(\mu x - \lambda)(x - 1) = 0
    \]
Examples of Quasi-Birth-and-Death processes: Single server with setup

- Solution: Spectral expansion

  - Roots
  \[ x_1 = \frac{\lambda}{\lambda + \theta}, \quad x_2 = \frac{\lambda}{\mu} \]

  with corresponding non-null vectors

  \[ y_1 = (1, \frac{-\theta x_1}{\lambda - (\lambda + \mu)x_1 + \mu x_2}), \quad y_2 = (0, 1) \]

- Set

  \[ p_i = c_1 y_1 x_1^i + c_2 y_2 x_2^i \]

  where coefficients \( c_1 \) and \( c_2 \) follow from boundary equation

  \[ p_0 B_1 + p_1 B_2 = 0 \]

  and normalization

  \[ 1 = \sum_{i,j} p(i,j) = \frac{c_1 y_1 e}{1 - x_1} + \frac{c_2 y_2 e}{1 - x_2} \]
Examples of Quasi-Birth-and-Death processes: Single server with setup

- Solution: Generating function
  - Let $P(z) = \sum_{i=0}^{\infty} p_i z^i$
  - Multiply balance equations by $z^i$ and add all equations
    \[ zP(z)A_0 + (P(z) - p_0)A_1 + z^{-1}(P(z) - p_0 - p_1 z)A_2 = 0 \]
  - Solving this equation
    \[ P(z)A(z) = p_0(A_1 + z^{-1}A_2) + p_1 A_2 \]
    where
    \[ A(z) = zA_0 + A_1 + z^{-1}A_2 = \begin{pmatrix} \lambda z - (\lambda + \theta) & \theta \\ 0 & \lambda z - (\lambda + \mu) + \mu z^{-1} \end{pmatrix} \]
  - Vectors $p_0$, $p_1$ and $P(1)$ follow from
    \[ p_0 B_1 + p_1 B_2 = 0 \]
    \[ P(1)A(1) = p_0(A_1 + A_2) + p_1 A_2, \quad P(1)e = 1 \]
    \[ P(1)A^{(1)}(1)e = -p_0 A_2 e \]
Examples of Quasi-Birth-and-Death processes: Shortest queue with jockeying

- Two parallel exponential servers with rate $\mu$, each with its own queue
- Customers arrive according to Poisson stream with rate $\lambda$ and join shortest queue on arrival
- Jockeying: If difference in queues exceeds 1, then a customer jumps from longest to shortest queue
Examples of Quasi-Birth-and-Death processes: Shortest queue with jockeying

- Markov chain with states \((i, j)\) with \(i\) number in shortest queue and \(j\) difference between queues
- Transition rate diagram

\[
\begin{align*}
0,0 & \quad \text{to} \quad 1,0 & \lambda \\
0,1 & \quad \text{to} \quad 1,1 & \mu \\
i,0 & \quad \text{to} \quad i+1,0 & 2\mu \\
i,1 & \quad \text{to} \quad i+1,1 & \lambda \\
\end{align*}
\]
Examples of Quasi-Birth-and-Death processes: Shortest queue with jockeying

- Partition into levels $i = \{(i, 0), (i, 1)\}$ for $i = 0, 1, \ldots$
- Generator $Q$ of the form

$$Q = \begin{pmatrix}
B_1 & A_0 & 0 & 0 & \ldots \\
A_2 & A_1 & A_0 & 0 \\
0 & A_2 & A_1 & A_0 \\
0 & 0 & A_2 & A_1 \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}$$

where

$$B_1 = \begin{pmatrix}
-\lambda \\
\mu
\end{pmatrix},
A_0 = \begin{pmatrix} 0 & 0 \\ \lambda & 0 \end{pmatrix},
A_1 = \begin{pmatrix}
-(\lambda + 2\mu) & \lambda \\
2\mu & -(\lambda + 2\mu)
\end{pmatrix},
A_2 = \begin{pmatrix} 0 & 2\mu \\ 0 & 0 \end{pmatrix}$$

- Let $p = (p_0, p_1, \ldots)$ and $p_i = (p(i, 0), p(i, 1))$

$$p_0 B_1 + p_1 A_2 = 0,
p_{i-1} A_0 + p_i A_1 + p_{i+1} A_2 = 0, \quad i = 1, 2, \ldots$$
Examples of Quasi-Birth-and-Death processes: Single server serving $n$ customer types

- Single server serving $n$ types of customers, numbered $1, \ldots, n$
- Type $i$ customers arrive according to Poisson stream with rate $\lambda_i$
- Type $i$ customers require exponential service with rate $\mu_i$
Examples of Quasi-Birth-and-Death processes: Single server serving \( n \) customer types

- Markov chain with states \( (i, j) \) where \( i \) number waiting in the queue and \( j \) is customer type in service
- State \((0, 0)\) is empty state
- Transition rate diagram where \( \lambda = \lambda_1 + \cdots + \lambda_n \) and \( p_i = \lambda_i / \lambda \)
Examples of Quasi-Birth-and-Death processes: Single server serving $n$ customer types

- Partition into level $0 = \{(0, 0), \ldots, (0, n)\}$ and levels $i = \{(i, 1), \ldots, (i, n)\}$ for $i = 1, 2, \ldots$

- Generator $Q$ of the form

$$Q = \begin{pmatrix}
    B_1 & B_0 & 0 & 0 & \ldots \\
    B_2 & A_1 & A_0 & 0 \\
    0 & A_2 & A_1 & A_0 \\
    0 & 0 & A_2 & A_1 \\
    \vdots & \vdots & \vdots & \ddots \\
\end{pmatrix}$$

where

$$A_0 = \begin{pmatrix}
    \lambda & 0 \\
    \vdots & \ddots \\
    0 & \lambda \\
\end{pmatrix}, \quad A_1 = \begin{pmatrix}
    -(\lambda + \mu_1) & \ldots & 0 \\
    \vdots & \ddots & \vdots \\
    0 & \ldots & -(\lambda + \mu_n) \\
\end{pmatrix}, \quad A_2 = \begin{pmatrix}
    \mu_1 p_1 & \cdots & \mu_1 p_n \\
    \mu_2 p_1 & \ddots & \vdots \\
    \mu_n p_1 & \cdots & \mu_n p_n \\
\end{pmatrix}$$

- Let $p = (p_0, p_1, \ldots)$, $p_0 = (p(0, 0), \ldots, p(0, n))$ and $p_i = (p(i, 1), \ldots, p(i, n))$

$$p_0 B_1 + p_1 B_2 = 0$$
$$p_0 B_0 + p_1 A_1 + p_2 A_2 = 0$$
$$p_{i-1} A_0 + p_i A_1 + p_{i+1} A_2 = 0, \quad i = 2, 3, \ldots$$
Quasi-Birth-and-Death processes

- Markov chain with states \((0, 0), (0, 1), \ldots, (0, n), (1, 0), (1, 1), \ldots, (1, n), (2, 0), \ldots\)
- Partition into levels \(i = \{(i, 0), \ldots, (i, n)\}\) for \(i = 0, 1, \ldots\)
- Generator \(Q\) of the form

\[
Q = \begin{pmatrix}
B_1 & A_0 & 0 & 0 & \ldots \\
A_2 & A_1 & A_0 & 0 \\
0 & A_2 & A_1 & A_0 \\
0 & 0 & A_2 & A_1 \\
\vdots & & & & \ddots
\end{pmatrix}
\]

where \(B_1, A_0, A_1\) and \(A_2\) are \(n+1 \times n+1\) matrices
- Let \(p = (p_0, p_1, \ldots)\) and \(p_i = (p(i, 0), p(i, 1), \ldots, p(i, n))\)

\[
p_0 B_0 + p_1 A_2 = 0 \\
p_{i-1} A_0 + p_i A_1 + p_{i+1} A_2 = 0, \quad i = 1, 2, \ldots
\]
Quasi-Birth-and-Death processes

- **Solution**: Matrix-geometric

  - Probability vector $p$ is given by
    $$ p_i = p_0 R^i, \quad i = 0, 1, 2, ... $$
  
  - Rate matrix $R$ is minimal nonnegative solution of
    $$ A_0 + RA_1 + R^2 A_2 = 0 $$
  
  - $R_{jk}$ is expected time spent in $(i + 1, k)$ before first return to level $i$, given initial state is $(i, j)$
  
  - Vector $p_0$ follows from boundary equation and normalization
    $$ p_0(B_1 + RA_2) = 0 $$
    $$ p_0(I - R)^{-1} e = 1 $$

  - Rate matrix $R$ can be determined by successive substitution from fixed point equation
    $$ R = -(A_0 + R^2 A_2)A_1^{-1} $$

  starting with $R = 0$
Quasi-Birth-and-Death processes

- Solution: Matrix-analytic
  - Let $g_{jk}$ be probability of first passage to level $i-1$ in $(i-1, k)$ for initial state $(i, j)$ at level $i$
  - Matrix $G = (g_{jk})$ is is minimal nonnegative solution of
    \[ A_2 + A_1 G + A_0 G^2 = 0 \]
  - Then
    \[ p_{i+1} A_2 = p_i A_0 G \]
  - Substituting in balance equation
    \[ 0 = p_{i-1} A_0 + p_i A_1 + p_{i+1} A_2 = p_{i-1} A_0 + p_i (A_1 + A_0 G) \]
  - Probability vector $p$ is given by
    \[ p_i = -p_{i-1} A_0 (A_1 + A_0 G)^{-1} = p_{i-1} R \]
    with
    \[ R = -A_0 (A_1 + A_0 G)^{-1} \]
Quasi-Birth-and-Death processes

- Solution: Matrix-analytic
  - Vector $p_0$ follows from boundary equation and normalization
    \[
    p_0(B_0 + RA_2) = p_0(B_0 + A_0 G) = 0
    \]
    \[
    p_0(I - R)^{-1}e = 1
    \]
  - Note $RA_2 = A_0 G$
  - Matrix $G$ can be determined by successive substitution from fixed point equation
    \[
    G = -(A_2 + A_0 G^2)A_1^{-1}
    \]
    starting with $G = 0$
Quasi-Birth-and-Death processes

- Solution: Spectral expansion
  - Seek solutions of balance equations \( p_{i-1}A_0 + p_iA_1 + p_{i+1}A_2 = 0 \) of the form
    \[
    p_i = y \cdot x^i, \quad i = 0, 1, 2, \ldots,
    \]
    where \( y = (y(0), \ldots, y(n)) \) is non-null vector and \(|x| < 1\)
  - Substitution and dividing by common powers of \( x \) yields
    \[
    y (A_0 + xA_1 + x^2A_2) = 0
    \]
  - Suppose determinant equation
    \[
    \det(A_0 + xA_1 + x^2A_2) = 0
    \]
    has exactly \( n + 1 \) roots \( x_0, \ldots, x_n \) with \(|x_k| < 1\)
  - Let \( y_0, \ldots, y_n \) be the corresponding non-null vectors
    \[
    y_k (A_0 + x_kA_1 + x_k^2A_2) = 0, \quad k = 0, \ldots, n
    \]
Quasi-Birth-and-Death processes

- Solution: Spectral expansion
  - Set
    \[ p_i = \sum_{k=0}^{n} c_k y_k x_k^i \]
  - Coefficients \( c_0, \ldots, c_k \) follow from boundary equation
    \[ p_0 B_1 + p_1 A_2 = \sum_{k=0}^{n} c_k (y_k B_1 + y_k A_2 x_k) = 0 \]

and normalization

\[ 1 = \sum_{i,j} p(i,j) = \sum_{k=0}^{n} \frac{c_k y_k e}{1 - x_k} \]
Quasi-Birth-and-Death processes

- Solution: Generating function
  - Let $P(z) = \sum_{i=0}^{\infty} p_i z^i$
  - Multiply balance equations by $z^i$ and add all equations

$$zP(z)A_0 + (P(z) - p_0)A_1 + z^{-1}(P(z) - p_0 - p_1 z)A_2 = 0$$

- Solving this equation

$$P(z)A(z) = p_0(A_1 + z^{-1}A_2) + p_1 A_2 = p_0(A_1 - B_1 + z^{-1}A_2)$$

where

$$A(z) = zA_0 + A_1 + z^{-1}A_2$$
Quasi-Birth-and-Death processes

• Solution: Generating function
  
  – Suppose determinant equation
    \[ \det(A(z)) = 0 \]
    has exactly \( n + 1 \) roots \( z_0 = 1, z_1, \ldots, z_n \) with \( |z_k| < 1 \) for \( k = 1, \ldots, n \)
  
  – Let \( v_0 = e, v_1, \ldots, v_n \) be non-null vectors satisfying
    \[ A(z_k)v_k = 0 \]
  
  – Vector \( p_0 \) follows from
    \[
    p_0(A_1 - B_1 + z_k^{-1}A_2)v_k = 0, \quad k = 1, \ldots, n
    \]
    \[
    P(1)A(1) = p_0(A_1 - B_1 + A_2), \quad P(1)e = 1
    \]
    \[
    P(1)A^{(1)}(1)e = -p_0A_2e
    \]
Quasi-Birth-and-Death processes: More structure

- $A_0$, $A_1$ and $A_2$ are upper triangular: $A_0(i,j) = 0$ if $j < i$
- Implication: $R$ is also upper triangular
- Example:
  
  $A_0 = \begin{pmatrix} a_0 & a_1 \\ 0 & a_2 \end{pmatrix}$, $A_1 = \begin{pmatrix} b_0 & b_1 \\ 0 & b_2 \end{pmatrix}$, $A_2 = \begin{pmatrix} c_0 & c_1 \\ 0 & c_2 \end{pmatrix}$,

- Then

  $R = \begin{pmatrix} r_0 & r_1 \\ 0 & r_2 \end{pmatrix}$

  where

  \[
  \begin{align*}
  a_0 + r_0 b_0 + r_0^2 c_0 &= 0 \\
  a_2 + r_2 b_2 + r_2^2 c_2 &= 0 \\
  a_1 + r_0 b_1 + r_1 b_2 + r_0^2 c_1 + (r_0 r_1 + r_1 r_2) c_2 &= 0
  \end{align*}
  \]
Quasi-Birth-and-Death processes: More structure

- Transition rates $A_2$ from level $i$ to level $i-1$ have the form

\[ A_2 = rp = \begin{pmatrix} r_0 \\ \vdots \\ r_n \end{pmatrix} \begin{pmatrix} p_0 & \cdots & p_n \end{pmatrix} = \begin{pmatrix} r_0p_0 & \cdots & r_0p_n \\ \vdots & \ddots & \vdots \\ r_np_0 & \cdots & r_np_n \end{pmatrix} \]

- Example: Single server with setup

\[ A_2 = \begin{pmatrix} 0 & 0 \\ 0 & \mu \end{pmatrix} = \begin{pmatrix} 0 \\ \mu \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} \]

- Example: Shortest queue with jockeying

\[ A_2 = \begin{pmatrix} 0 & 2\mu \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2\mu \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} \]

- Example: Single server serving $n$ customer types for which

\[ A_2 = \begin{pmatrix} \mu_1p_1 & \cdots & \mu_1p_n \\ \vdots & \ddots & \vdots \\ \mu_np_1 & \cdots & \mu_np_n \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \begin{pmatrix} p_1 & \cdots & p_n \end{pmatrix} \]
Quasi-Birth-and-Death processes: More structure

- Transition rates $A_2$ from level $i$ to level $i - 1$ have the form

$$A_2 = rp = \begin{pmatrix} r_0 \\ \vdots \\ r_n \end{pmatrix} \begin{pmatrix} p_0 & \cdots & p_n \end{pmatrix} = \begin{pmatrix} r_0p_0 & \cdots & r_0p_n \\ \vdots & \ddots & \vdots \\ r_np_0 & \cdots & r_np_n \end{pmatrix}$$

- Interpretation: Rate of jumping from $(i, j)$ to level $i - 1$ is $r_j$, this jump is to $(i - 1, k)$ with probability $p_k$

- Important: Jump probability does not depend on $j$

- This implies

$$G = \begin{pmatrix} p_0 & \cdots & p_n \\ \vdots & \ddots & \vdots \\ p_0 & \cdots & p_n \end{pmatrix} = ep$$

and

$$R = -A_0(A_1 + A_0G)^{-1} = -A_0(A_1 + A_0ep)^{-1}$$
Quasi-Birth-and-Death processes: More structure

- Transition rates $A_0$ from level $i$ to level $i + 1$ have the form

$$A_0 = rp = \begin{pmatrix} r_0 \\ \vdots \\ r_n \end{pmatrix} \begin{pmatrix} p_0 & \cdots & p_n \end{pmatrix} = \begin{pmatrix} r_0 p_0 & \cdots & r_0 p_n \\ \vdots & \ddots & \vdots \\ r_n p_0 & \cdots & r_n p_n \end{pmatrix}$$

- Example: Shortest queue with jockeying

$$A_0 = \begin{pmatrix} 0 & 0 \\ \lambda & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix}$$
Quasi-Birth-and-Death processes: More structure

- Transition rates $A_0$ from level $i$ to level $i+1$ have the form

$$A_0 = rp = \begin{pmatrix} r_0 \\ \vdots \\ r_n \end{pmatrix} \begin{pmatrix} p_0 & \cdots & p_n \end{pmatrix} = \begin{pmatrix} r_0p_0 & \cdots & r_0p_n \\ \vdots & \ddots & \vdots \\ r_np_0 & \cdots & r_np_n \end{pmatrix}$$

- Interpretation: Rate of jumping from $(i,j)$ to level $i+1$ is $r_j$, this jump is to $(i+1,k)$ with probability $p_k$

- This implies

$$R = rq = \begin{pmatrix} r_0 \\ \vdots \\ r_n \end{pmatrix} \begin{pmatrix} q_0 & \cdots & q_n \end{pmatrix} = \begin{pmatrix} r_0q_0 & \cdots & r_0q_n \\ \vdots & \ddots & \vdots \\ r_nq_0 & \cdots & r_nq_n \end{pmatrix}$$

for some vector $q$

- Hence

$$R^i = (qr)^{i-1}R, \quad p_i = p_0R^i = (qr)^{i-1}p_1, \quad i = 1, 2, \ldots$$

- Number $x = qr$ is spectral radius of $R$ which is unique root on $(0, 1)$ of

$$\det(A_0 + xA_1 + x^2A_2) = 0$$
Quasi-Birth-and-Death processes: Linear rates

- Transition rates have linear form to neighboring phases
  - Rate from \((i, j)\) to \((i + k, j + 1)\) is \(a_k(n - j)\)
  - Rate from \((i, j)\) to \((i + k, j)\) is \(b_k(n - j) + c_k j\)
  - Rate from \((i, j)\) to \((i + k, j - 1)\) with rate \(d_k j\)

- So \(A_0\) is of tri-diagonal form

\[
A_0 = \begin{pmatrix}
  b_1 n & a_1 n & 0 \\
  d_1 & b_1(n - 1) + c_1 & a_1(n - 1) \\
  0 & d_1 2 & b_1(n - 2) + c_1 2 \\
  & \ddots & \ddots & \ddots \\
  b_1 2 + c_1(n - 2) & a_1 2 & 0 \\
  d_1(n - 1) & b_1 + c_1(n - 1) & a_1 \\
  0 & d_1 n & c_1 n
\end{pmatrix}
\]
Quasi-Birth-and-Death processes: Linear rates

- Transition rate diagram
Quasi-Birth-and-Death processes: Example with linear rates

- $n$ parallel servers serving two customer types 1 and 2
- Type 1 customers require exponential service with rate $\mu_1$, type 2 with rate $\mu_2$
- Customers arrive as Poisson stream with rate $\lambda$, fraction $p_1$ is of type 1 and $p_2$ of type 2
- States $(i, j)$ with $i$ number of customers in the queue and $j$ number of type 1 customers in service
- Transition rate diagram

- Non-zero coefficients $a_{-1} = \mu_2 p_1$, $b_1 = c_1 = \lambda/n$, $b_{-1} = \mu_2 p_2$, $c_{-1} = \mu_1 p_1$, $d_{-1} = \mu_1 p_2$
Quasi-Birth-and-Death processes: Example with linear rates

- $n$ parallel servers and Poisson arrivals with rate $\lambda$ which need exponential service times with rate $\mu$
- Server switched off when queue is empty
- Idle server switched on when customer arrives which requires exponential setup time with rate $\theta$
- States $(i, j)$ with $i$ number of customers in the queue and $j$ number of active servers
- Transition rate diagram

Non-zero coefficients $a_0 = \theta$, $b_1 = c_1 = \lambda/n$, $c_{-1} = \mu$
Quasi-Birth-and-Death processes: Example with linear rates

- Solution: Spectral expansion
  - Seek solutions of balance equations of the form
    \[ p_i = y \cdot x^i, \quad i = 0, 1, 2, \ldots, \]
    where \( y = (y(0), \ldots, y(n)) \) is non-null vector and \( |x| < 1 \)
  - Substitution and dividing by common powers of \( x \) yields following equations for \( j = 0, \ldots, n \)
    \[
    0 = (j + 1)D(x)y(j + 1) + ((n - j)B(x) + jC(x))y(j) \\
    -((n - j)(A(1) + B(1)) + j(C(1) + D(1)))xy(j) + (n - j + 1)A(x)y(j - 1)
    \]
    where
    \[
    A(x) = a_1 + a_0x + a_{-1}x^2, \quad B(x) = b_1 + b_0x + b_{-1}x^2, \quad C(x) = c_1 + c_0x + c_{-1}x^2, \quad D(x) = d_1 + d_0x + d_{-1}x^2
    \]
  - Solve equations for \( y \) by generating function
    \[
    Y(z) = \sum_{i=0}^{n} y(j)z^j
    \]
Quasi-Birth-and-Death processes: Example with linear rates

- Solution: Spectral expansion
  - Multiply equations by $z^j$ and add all equations

\[
D(x) Y'(z) + nB(x)Y(z) + (C(x) - B(x))zY'(z) - (A(1) + B(1))n x Y(z) \\
+ (A(1) + B(1) - C(1) - D(1))xzY'(z) + nA(x)zY(z) - A(x)z^2 Y'(z) = 0
\]

- This equation can be rewritten as

\[
\frac{Y'(z)}{Y(z)} = \frac{n[A(x)z + B(x) - (A(1) + B(1))x]}{A(x)z^2 - ((A(1) + B(1) - C(1) - D(1))x + C(x) - B(x))z - D(x)} \\
= \frac{E(x)}{z - z_1(x)} + \frac{n - E(x)}{z - z_2(x)}
\]

where $z_1(x)$ and $z_2(x)$ are the roots of the denominator and $E(x)$ satisfies

\[
2E(x) - n = n \cdot \frac{B(x) + C(x) - x(A(1) + B(1) + C(1) + D(1))}{\sqrt{F(x)^2 + 4A(x)D(x)}},
\]

where $F(x) = (A(1) + B(1) - C(1) - D(1))x + C(x) - B(x)$
Quasi-Birth-and-Death processes: Example with linear rates

- Solution: Spectral expansion
  - Differential equation
    \[
    \frac{Y'(z)}{Y(z)} = \frac{E(x)}{z - z_1(x)} + \frac{n - E(x)}{z - z_2(x)}
    \]
  - General solution
    \[
    Y(z) = C(z - z_1(x))^E(x)(z - z_2(x))^{n - E(x)}
    \]
- Key observation: \( Y(z) \) is a polynomial in \( z \), so exponent \( E(x) = k \) for \( k = 0, \ldots, n \)
- For each \( k = 0, \ldots, n \) the equation
  \[
  2k - n = n \cdot \frac{B(x) + C(x) - x (A(1) + B(1) + C(1) + D(1))}{\sqrt{F(x)^2 + 4A(x)D(x)}},
  \]
  has a unique solution \( x = x_k \) in the interval \((0, 1)\)
Quasi-Birth-and-Death processes: Example with linear rates

- Solution: Spectral expansion
  - For each $k = 0, \ldots, n$ the equation
    \[
    2k - n = n \cdot \frac{B(x) + C(x) - x (A(1) + B(1) + C(1) + D(1))}{\sqrt{F(x)^2 + 4A(x)D(x)}},
    \]
  has a unique solution $x = x_k$ in the interval $(0, 1)$
Two-dimensional Markov chain: Re-entrant line (1)

- Two stations, station 1 with two servers, labeled 1 and 3 and station 2 with single server, labeled server 2
- Queue $i$ is served at exponential rate $\mu_i$
- Queue 1 has infinite supply and customers subsequently go through queue 1, 2 and 3
Two-dimensional Markov chain: Re-entrant line (1)

- Markov chain with states \((i, j)\) where \(i\) number in queue 2 and \(j\) number in queue 3
- Global balance equations

\[
\begin{align*}
p(i, j)(\mu_1 + \mu_2 + \mu_3) &= p(i - 1, j)\mu_1 + p(i, j + 1)\mu_3 + p(i + 1, j - 1)\mu_2, \quad i > 0, j > 0 \\
p(i, 0)(\mu_1 + \mu_2) &= p(i - 1, 0)\mu_1 + p(i, 1)\mu_3, \quad i > 0 \\
p(0, j)(\mu_1 + \mu_3) &= p(0, j + 1)\mu_3 + p(1, j - 1)\mu_2, \quad j > 0 \\
p(0, 0)\mu_1 &= p_{0,1}\mu_3
\end{align*}
\]
Two-dimensional Markov chain: Re-entrant line (1)

- Substitution of trial solution $p(i, j) = \alpha^i \beta^j$ in global balance equations

\[
\begin{align*}
\alpha \beta (\mu_1 + \mu_2 + \mu_3) &= \beta \mu_1 + \alpha \beta^2 \mu_3 + \alpha^2 \mu_2 \\
\alpha (\mu_1 + \mu_2) &= \mu_1 + \alpha \beta \mu_3 \\
\beta (\mu_1 + \mu_3) &= \beta^2 \mu_3 + \alpha \mu_2
\end{align*}
\]

- Equations solved by $\alpha = \frac{\mu_1}{\mu_2}$ and $\beta = \frac{\mu_1}{\mu_3}$

- Normalization yields

$$p(i, j) = (1 - \alpha) \alpha^i (1 - \beta) \beta^j$$
Two-dimensional Markov chain: Re-entrant line (1)

- Global balance equations

\[ p(i,j)(\mu_1 + \mu_2 + \mu_3) = p(i-1,j)\mu_1 + p(i,j+1)\mu_3 + p(i+1,j-1)\mu_2, \quad i > 0, j > 0 \]

- Partial balance
  - Rate out of \((i,j)\) due to arrival in queue 2 is equal to rate into \((i,j)\) due to departure from queue 3

\[ p(i,j)\mu_1 = p(i+1,j)\mu_3 \]

  - Rate out of \((i,j)\) due to departure from queue 2 is equal to rate into \((i,j)\) due to arrival in queue 2

\[ p(i,j)\mu_2 = p(i-1,j)\mu_1 \]

  - Rate out of \((i,j)\) due to departure from queue 3 is equal to rate into \((i,j)\) due to arrival in queue

\[ p(i,j)\mu_3 = p(i+1,j-1)\mu_2 \]

- Partial balance yields candidate solution

\[ p(i,j) = \left(\frac{\mu_2}{\mu_3}\right)^j p(i+j,0) = \left(\frac{\mu_2}{\mu_3}\right)^j \left(\frac{\mu_1}{\mu_2}\right)^{i+j} p(0,0) = \left(\frac{\mu_1}{\mu_2}\right)^i \left(\frac{\mu_1}{\mu_3}\right)^j p(0,0) \]
Two-dimensional Markov chain: Re-entrant line (2)

- Two stations, both stations with single server
- Queue $i$ is served at exponential rate $\mu_i$
- Queue 1 has infinite supply and customers subsequently go through queue 1, 2 and 3
- Customers in queue 3 have preemptive priority over customers in queue 1 (Last Buffer First Served)
- Assumption $\frac{1}{\mu_1} + \frac{1}{\mu_3} > \frac{1}{\mu_2}$
Two-dimensional Markov chain: Re-entrant line (2)

- Markov chain with states \((i, j)\) where \(i\) number in queue 2 and \(j\) number in queue 3
- Global balance equations

\[
\begin{align*}
p(i, j)(\mu_2 + \mu_3) &= p(i, j + 1)\mu_3 + p(i + 1, j - 1)\mu_2, \quad i > 0, j > 0 \\
p(i, 0)(\mu_1 + \mu_2) &= p(i - 1, 0)\mu_1 + p(i, 1)\mu_3, \quad i > 0 \\
p(0, j)\mu_3 &= p(0, j + 1)\mu_3 + p(1, j - 1)\mu_2, \quad j > 0 \\
p(0, 0)\mu_1 &= p(0, 1)\mu_3
\end{align*}
\]
Two-dimensional Markov chain: Re-entrant line (2)

- Substitution of trial solution $p(i, j) = \alpha^i \beta^j$ in global balance equations for $i > 0$

  \[
  \beta(\mu_2 + \mu_3) = \beta^2 \mu_3 + \alpha \mu_2 \\
  \alpha(\mu_1 + \mu_2) = \mu_1 + \alpha \beta \mu_3
  \]

- Equations solved by $\alpha = \frac{\mu_1}{\mu_2} (1 - \beta)$ and $\beta$ is the root on $(0, 1)$ of $\beta^2 \mu_3 - \beta (\mu_1 + \mu_2 + \mu_3) + \mu_1 = 0$

  \[
  \beta = \frac{\mu_1 + \mu_2 + \mu_3 - \sqrt{\left(\mu_1 + \mu_2 + \mu_3\right)^2 - 4 \mu_1 \mu_3}}{2 \mu_3}
  \]
Two-dimensional Markov chain: Re-entrant line (2)

- Substitution of $p(i, j) = \alpha^i \beta^j$ for $i > 0$ in global balance equations for $i = 0$
  \[
p(0, j) \mu_3 = p(0, j + 1) \mu_3 + \alpha \beta^{j-1} \mu_2, \quad j > 0
  \]
- Equations are solved by $p(0, j) = c \beta^{j-1}$ where $c = \frac{\mu_1}{\mu_3}$
- Normalization yields
  \[
p(i, j) = \begin{cases} 
  \frac{\mu_3}{\mu_1 + \mu_3} (1 - \alpha), & i = 0, j = 0 \\
  \frac{\mu_1}{\mu_1 + \mu_3} (1 - \alpha) \beta^{j-1}, & i = 0, j > 0 \\
  \frac{\mu_3}{\mu_1 + \mu_3} (1 - \alpha) \alpha^i \beta^j, & i > 0, j \geq 0
  \end{cases}
  \]
Two-dimensional Markov chain: Single server priority queue

- Single server serving type 1 and type 2 customers
- Type $i$ customers arrive according to Poisson stream with rate $\lambda_i$
- Type $i$ customers require exponential service with rate $\mu_i$
- Type 1 customers have preemptive priority over type 2 customers
Two-dimensional Markov chain: Single server priority queue

- Markov chain with states \((i, j)\) where \(i\) number of type 1 and \(j\) number of type 2
- Global balance equations

\[
\begin{align*}
\lambda_1 + \lambda_2 + \mu_1 \quad & p(i, j)(\lambda_1 + \lambda_2 + \mu_1) = p(i - 1, j)\lambda_1 + p(i + 1, j)\mu_1 + p(i, j - 1)\lambda_2, \quad i > 0, j > 0 \\
\lambda_1 + \lambda_2 + \mu_1 \quad & p(i, 0)(\lambda_1 + \lambda_2 + \mu_1) = p(i - 1, 0)\lambda_1 + p(i + 1, 0)\mu_1, \quad i > 0 \\
\lambda_1 + \lambda_2 + \mu_2 \quad & p(0, j)(\lambda_1 + \lambda_2 + \mu_2) = p(1, j)\mu_1 + p(0, j - 1)\lambda_2 + p(0, j + 1)\mu_2, \quad j > 0 \\
\lambda_1 + \mu_2 \quad & p(0, 0)(\lambda_1 + \mu_2) = p(0, 1)\mu_2 + p(1, 0)\mu_1
\end{align*}
\]

- Note that \(p(0, 0) = 1 - \rho_1 - \rho_2\) where \(\rho_i = \frac{\lambda_i}{\mu_i}\)
Two-dimensional Markov chain: Single server priority queue

- Solution: Difference equation
  - Balance equations for $j = 0$ are homogeneous second-order difference equation
    \[
    p(i, 0)(\lambda_1 + \lambda_2 + \mu_1) = p(i - 1, 0)\lambda_1 + p(i + 1, 0)\mu_1, \quad i > 0
    \]
  - General solution
    \[
    p(i, 0) = c_{0,0}x_0^i + c_{0,1}x_1^i, \quad i \geq 0
    \]
    where $0 < x_0 < 1 < x_1$ are roots of quadratic equation
    \[
    x(\lambda_1 + \lambda_2 + \mu_1) = \lambda_1 + x^2\mu_1
    \]
  - Convergence of $p(i, 0)$ implies $c_{0,1} = 0$ so
    \[
    p(i, 0) = c_{0,0}x_0^i
    = (1 - \rho_1 - \rho_2)x_0^i, \quad i \geq 0
    \]
Two-dimensional Markov chain: Single server priority queue

- Solution: Difference equation
  - Balance equations for \( j = 1 \) are inhomogeneous second-order difference equation
    \[
    p(i, 1)(\lambda_1 + \lambda_2 + \mu_1) = p(i - 1, 1)\lambda_1 + p(i + 1, 1)\mu_1 + p(i, 0)\lambda_2 \\
    = p(i - 1, 1)\lambda_1 + p(i + 1, 1)\mu_1 + c_{0,0}x_0^i\lambda_2, \quad i > 0
    \]
  - General solution
    \[
    p(i, 1) = c_{1,0}x_0^i + c_{1,1}(i + 1)x_0^i \\
    = c_{1,0}\binom{i + 1}{0}x_0^i + c_{1,1}\binom{i + 1}{1}x_0^i, \quad i \geq 0
    \]
    where
    \[
    c_{1,1} = \frac{\lambda_2c_{0,0}}{\lambda_1 + \lambda_2 - 2\mu_1x_0}
    \]
  - Coefficient \( c_{1,0} \) follows from balance equation in \((0, 0)\)
    \[
    c_{1,0} = \frac{\lambda_1 + \lambda_2 - \mu_1x_0}{\mu_2}c_{0,0} - c_{1,1}
    \]
Two-dimensional Markov chain: Single server priority queue

- Solution: Difference equation
  - Repeating procedure for balance equations for $j = 2, 3, \ldots$ leads to

  $p(i, j) = \sum_{k=0}^{j} c_{j,k} \binom{i+j}{k} x_0^i, \quad i, j \geq 0$

  where coefficients $c_{j,k}$ can be determined recursively starting with $c_{0,0} = 1 - \rho_1 - \rho_2$
Two-dimensional Markov chain: Single server priority queue

- Solution: Generating function
  - Let $P(x, y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} p(i,j)x^iy^j$ for $|x|, |y| \leq 1$
  - Multiply balance equations by $x^iy^j$ and add all equations

$$h_1(x, y)P(x, y) = h_2(x, y)P(0, y) + h_3(x, y)P(0, 0)$$

with

$$
\begin{align*}
  h_1(x, y) &= \lambda_1xy(1-x) + \lambda_2xy(1-y) - \mu_1y(1-x) \\
  h_2(x, y) &= -\mu_1y(1-x) + \mu_2x(1-y) \\
  h_3(x, y) &= -\mu_2x(1-y)
\end{align*}
$$

and $P(0, 0) = p(0, 0) = 1 - \rho_1 - \rho_2$
Two-dimensional Markov chain: Single server priority queue

- **Solution: Generating function**
  - For every $|y| \leq 1$ let $x = \xi(y)$ be the unique root with $|x| \leq 1$ of the quadratic equation
    \[
    0 = h_1(x, y) = \lambda_1 xy(1 - x) + \lambda_2 xy(1 - y) - \mu_1 y(1 - x)
    \]
  - Then
    \[
    x = \xi(y) = \frac{\lambda_1 + \mu_1 + \lambda_2(1 - y) - \sqrt{(\lambda_1 + \mu_1 + \lambda_2(1 - y))^2 - 4\lambda_1\mu_1}}{2\lambda_1}
    \]
  - Substituting $x = \xi(y)$ and $P(0, 0) = 1 - \rho_1 - \rho_2$ in
    \[
    h_1(x, y)P(x, y) = h_2(x, y)P(0, y) + h_3(x, y)P(0, 0)
    \]
  gives
    \[
    P(0, y) = \frac{-h_3(\xi(y), y)(1 - \rho_1 - \rho_2)}{h_2(\xi(y), y)}
    \]
Two-dimensional Markov chain: Single server priority queue

- Solution: Matrix geometric
  - Partition states space into levels \( i = \{(i, 0), (i, 1), \ldots\} \)
  - Generator \( Q \) of the form
    \[
    Q = \begin{pmatrix}
    B_1 & A_0 & 0 & 0 & \ldots \\
    A_2 & A_1 & A_0 & 0 & \ldots \\
    0 & A_2 & A_1 & A_0 & \ldots \\
    0 & 0 & A_2 & A_1 & \ldots \\
    \vdots & \vdots & \vdots & \vdots & \ddots
    \end{pmatrix}
    \]

where \( \lambda = \lambda_1 + \lambda_2 \)

\[
A_0 = \begin{pmatrix}
\lambda_1 \\
\lambda_1 \\
\vdots
\end{pmatrix}, \quad A_1 = \begin{pmatrix}
-\lambda - \mu_1 & \lambda_2 \\
-\lambda - \mu_1 & \lambda_2 \\
\lambda_2 \\
\vdots & \ddots & \ddots
\end{pmatrix}, \quad A_2 = \begin{pmatrix}
\mu_1 & & \\
& \mu_1 & \\
& & \ddots
\end{pmatrix}
\]

\[
B_1 = \begin{pmatrix}
-\lambda & \lambda_2 \\
\mu_2 & -\lambda - \mu_2 & \lambda_2 \\
\vdots & \ddots & \ddots
\end{pmatrix}
\]
Two-dimensional Markov chain: Single server priority queue

- Solution: Matrix geometric
  - \(A_0, A_1\) and \(A_2\) are upper triangular and have repeating structure
  - Implication: \(R\) is also upper triangular and of the form
    \[
    R = \begin{pmatrix}
    r_0 & r_1 & r_2 & \cdots \\
    r_0 & r_1 & \cdots \\
    r_0 & \cdots \\
    \cdots
    \end{pmatrix}
    \]

- Elements can be determined recursively from matrix equation \(A_0 + RA_1 + R^2A_2 = 0\)
  \[
  \mu_1 r_0^2 - (\lambda + \mu_1)r_0 + \lambda_1 = 0
  \]
  \[
  \mu_1 \sum_{l=0}^{k} r_{k-l}r_l - (\lambda + \mu_1)r_k + \lambda_2 r_{k-1} = 0, \quad k \geq 1
  \]
Two-dimensional Markov chain: Single server priority queue

- Solution: Matrix geometric
  - Vector $p_0 = (p(0,0), p(0,1), ...)$ follows from boundary equations
    \[
    0 = p_0B_1 + p_1A_2 = p_0(B_1 + RA_2)
    \]
    yielding the recursion
    \[
    p(0, 1)\mu_2 = p(0, 0)(\lambda - \mu_1 r_0)
    \]
    \[
    p(0, j + 1)\mu_2 = (\lambda + \mu_2)p(0, j) - \lambda_2p(0, j - 1) - \mu_1 \sum_{k=0}^{j} p(0, j - k)r_k, \quad j \geq 1
    \]
    starting with $p(0, 0) = 1 - \rho_1 - \rho_2$
  - Vector $p_{i+1} = (p(0,0), p(0,1), ...)$ can be obtained from $p_{i+1} = p_iR$ yielding
    \[
    p(i + 1, j) = \sum_{k=0}^{j} p(i, j - k)r_k
    \]
Two-dimensional Markov chain: Single server priority queue

- Solution: Matrix analytic
  - $A_0$, $A_1$ and $A_2$ are upper triangular and have repeating structure
  - Implication: $G$ is also upper triangular and of the form
    \[
    G = \begin{pmatrix}
    g_0 & g_1 & g_2 & \\
    g_0 & g_1 & \\
    g_0 & \\
    \end{pmatrix}
    \]
  - $g_k$ is probability of first passage to level $i-1$ in $(i-1, j+k)$ for initial state $(i, j)$ at level $i$
  - One-step analysis

\[
\begin{align*}
g_0 &= \frac{\mu_1}{\lambda + \mu_1} + \frac{\lambda_1}{\lambda + \mu_1}g_0^2 \\
g_k &= \frac{\lambda_2}{\lambda + \mu_1}g_{k-1} + \frac{\lambda_1}{\lambda + \mu_1} \sum_{l=0}^{k} g_l g_{k-l}, \quad k \geq 1
\end{align*}
\]
Two-dimensional Markov chain: Single server priority queue

- Solution: Matrix analytic
  - From $p_{i+1} A_2 = p_i A_0 G$ follows

  $$p(i + 1, j) \mu_1 = \sum_{k=0}^{j} p(i, j - k) \lambda_1 g_k, \quad i, j \geq 0$$
Two-dimensional Markov chain: Single server priority queue

- **Solution: Matrix analytic**
  - Vector $p_0$ follows from boundary equations
    \[ 0 = p_0 B_1 + p_1 A_2 = p_0 (B_1 + A_0 G) \]
  - $B_1 + A_0 G$ is generator of Markov chain embedded on level 0

- Global balance for $\{(0, 0), (0, 1), \ldots, (0, j)\}$
  \[ p(0, j + 1) \mu_2 = p(0, j) \lambda_2 + \sum_{k=0}^{j} p(0, j - k) \lambda_1 \left( 1 - \sum_{l=0}^{k} g_l \right) \]
Two-dimensional Markov chain: Shortest queue

- Two parallel exponential servers with rate $\mu$, each with its own queue
- Customers arrive according to Poisson stream with rate $\lambda$ and join shortest queue on arrival
Two-dimensional Markov chain: Shortest queue

- Markov chain with states \((i, j)\) where \(i\) number in shortest queue and \(j\) difference between queues
- Global balance equations

\[
\begin{align*}
p(i, j)(\lambda + 2\mu) &= p(i - 1, j + 1)\lambda + p(i, j + 1)\mu + p(i + 1, j - 1)\mu, \quad i > 0, j > 1 \\
p(i, 1)(\lambda + 2\mu) &= p(i - 1, 2)\lambda + p(i, 2)\mu + p(i, 0)\lambda + p(i + 1, 0)2\mu, \quad i > 0 \\
p(i, 0)(\lambda + 2\mu) &= p(i - 1, 1)\lambda + p(i, 1)\mu, \quad i > 0 \\
p(0, j)(\lambda + \mu) &= p(0, j + 1)\mu + p(1, j - 1)\mu, \quad j > 1
\end{align*}
\]
Two-dimensional Markov chain: Shortest queue

- Markov chain with states $(i, j)$ where $i$ number in shortest queue and $j$ difference between queues
- Global balance equations

\[
\begin{align*}
p(i, j)(\lambda + 2\mu) &= p(i - 1, j + 1)\lambda + p(i, j + 1)\mu + p(i + 1, j - 1)\mu, \quad i > 0, j > 1 \\
p(i, 1)(\lambda + 2\mu) &= p(i - 1, 2)\lambda + p(i, 2)\mu \\
                      &\quad + (p(i - 1, 1)\lambda + p(i, 1)\mu)\frac{\lambda}{\lambda + 2\mu} + (p(i, 1)\lambda + p(i + 1, 1)\mu)\frac{2\mu}{\lambda + 2\mu}, \quad i > 0 \\
p(0, j)(\lambda + \mu) &= p(0, j + 1)\mu + p(1, j - 1)\mu, \quad j > 1
\end{align*}
\]
Intermezzo: Method of images

- What is the potential $\Phi$ in point $p$?
Intermezzo: Method of images

- What is the potential $\Phi$ in point $p$?
- Add image charge $-q$:
- Potential $\Phi$ in point $p$ is

$$
\Phi(p) = \frac{q}{r_1} - \frac{q}{r_2}
$$

where $r_1$ and $r_2$ are distances from point $p$ to charge $q$ and $-q$
Intermezzo: Method of images

- For two planes at right angles

$$\Phi(p) = \frac{q}{r_1} - \frac{q}{r_2} + \frac{q}{r_3} - \frac{q}{r_4}$$
**Intermezzo: Method of images**

- What is the potential $\Phi$ in any point $p$ outside the spheres $A$ and $B$?
Intermezzo: Method of images

- What is the potential $\Phi$ in any point $p$ outside the spheres $A$ and $B$?
- Add image charge $\alpha_0$ in center of $A$ with $\alpha_0 = a\Phi_a$: Then
  \[ \Phi(p) \approx \frac{\alpha_0}{r_0} \]
  where $r_0$ is distance from $p$ to charge $\alpha_0$ (correct if $c = \infty$)
- Potential of $\alpha_0$ does not vanish on $B$
Intermezzo: Method of images

- Place new image charge $\beta_0$ in $B$ to compensate for this error

$$\Phi(p) \approx \frac{\alpha_0}{r_0} + \frac{\beta_0}{r_1}$$

where $r_0$ and $r_1$ are distances from $p$ to charges $\alpha_0$ and $\beta_0$

- Charge $\beta_0$ alters potential on $A$
Intermezzo: Method of images

- Place new image charge $\beta_0$ in $B$ to compensate for this error

  $$\Phi(p) \approx \frac{\alpha_0}{r_0} + \frac{\beta_0}{r_1}$$

  where $r_0$ and $r_1$ are distances from $p$ to charges $\alpha_0$ and $\beta_0$

- Charge $\beta_0$ alters potential on $A$

- Place new image charge $\alpha_1$ in $A$ to compensate for this error

  $$\Phi(p) \approx \frac{\alpha_0}{r_0} + \frac{\beta_0}{r_1} + \frac{\alpha_1}{r_2}$$
Intermezzo: Method of images

- Potential problem of two spheres is solved by infinite sequence of image charges in A and B

\[ \Phi(p) = \frac{\alpha_0}{r_0} + \frac{\beta_0}{r_1} + \frac{\alpha_1}{r_2} + \frac{\beta_1}{r_3} + \cdots \]
Two-dimensional Markov chain: Shortest queue

• Image charges are products $\alpha^i \beta^j$ satisfying the global balance equations

\[ p(i,j)(\lambda + 2\mu) = p(i-1,j+1)\lambda + p(i,j+1)\mu + p(i+1,j-1)\mu, \quad i > 0, j > 1 \]

• Substituting $\alpha^i \beta^j$ in this equation and dividing by common powers gives the curve

\[ \alpha\beta(\lambda + 2\mu) = \beta^2\lambda + \alpha\beta^2\mu + \alpha^2\mu \]
Two-dimensional Markov chain: Shortest queue

- Quadratic curve

\[ \alpha \beta (\lambda + 2\mu) = \beta^2 \lambda + \alpha \beta^2 \mu + \alpha^2 \mu \]
Two-dimensional Markov chain: Shortest queue

- Unique product form satisfying global balance equations for $i > 0, j > 1$ and $i > 0, j = 1$:
  \[
  p(i, j) \approx c_0 \alpha_0^i \beta_0^j
  \]
  where $\alpha_0 = \rho^2$, $\beta_0 = \rho^2/(2 + \rho)$ with $\rho = \lambda/2\mu$

- Product form $c_0 \alpha_0^i \beta_0^j$ violates global balance equations for $i = 0$
Two-dimensional Markov chain: Shortest queue

- Add new product form $c_1 \alpha_1^i \beta_1^j$ such that sum satisfies balance equations for $i = 0$

$$p(i, j) \approx c_0 \alpha_0^i \beta_0^j + c_1 \alpha_1^i \beta_1^j$$

- Then $\beta_1 = \beta_0$ and $\alpha_1$ is companion root of $\alpha_0$ and

$$c_1 = -\frac{\alpha_1 - \beta_0}{\alpha_0 - \beta_0} c_0$$

- Product form $c_1 \alpha_1^i \beta_1^j$ violates global balance equations for $j = 1$
Add new product form \( c_2 \alpha_2^i \beta_2^j \) such that sum satisfies balance equations for \( j = 1 \)

\[
\rho(i, j) \approx c_0 \alpha_0^i \beta_0^j + c_1 \alpha_1^i \beta_1^j + c_2 \alpha_2^i \beta_2^j
\]

Then \( \alpha_2 = \alpha_1 \) and \( \beta_2 \) is companion root of \( \beta_1 \) and (with \( \rho = \lambda/(2\mu) \))

\[
c_2 = -\frac{(\rho + \alpha_1^i) / \beta_2 - (1 + \rho)}{(\rho + \alpha_1^i) / \beta_1 - (1 + \rho)} c_1
\]

Product form \( c_2 \alpha_2^i \beta_2^j \) violates global balance equations for \( i = 0 \)
Two-dimensional Markov chain: Shortest queue

- Iterating yields infinite sequence

\[ p(i, j) = c_0^i \alpha_0^i \beta_0^i + c_1^i \alpha_1^i \beta_1^i + c_2^i \alpha_2^i \beta_2^i + c_3^i \alpha_3^i \beta_3^i + \cdots \]
Two-dimensional Markov chain: Random walk in quarter plane

- Equilibrium probabilities can be expressed as infinite sequence of product forms

\[ p(i,j) = c_0 \alpha_0^i \beta_0^j + c_1 \alpha_1^i \beta_1^j + c_2 \alpha_2^i \beta_2^j + \cdots \]
Two-dimensional Markov chain: Random walk in quarter plane

- Equilibrium probabilities can be expressed as infinite sequence of product forms
  \[ p(i, j) = c_0 \alpha_0^i \beta_0^j + c_1 \alpha_1^i \beta_1^j + c_2 \alpha_2^i \beta_2^j + \cdots \]

- Provided there are no transitions to North, North-East and East
  \[ q_{0,1} = q_{1,1} = q_{1,0} = 0 \]
Two-dimensional Markov chain: Random walk in quarter plane

- Condition $q_{0,1} = q_{1,1} = q_{1,0} = 0$ implies that curve for $\alpha$ and $\beta$ is of the above form.
- This form guarantees that $\alpha_k \to 0$ and $\beta_k \to 0$ as $k \to \infty$. 
Multi-skilled system

- Customer types $C = \{a, b, \ldots\}$
- Independent Poisson arrivals with rates $\lambda_c$, $c \in C$
- Exponential service requirements with mean 1
- $J$ servers, $\mathcal{S} = \{m_1, \ldots, m_J\}$
- Server $m_j$ works at rate $\mu_{m_j}$
- Skill-based service:
  - Server $m_i$ can serve customer types $C(m_j)$
  - Type $c$ customer can be served by servers $\mathcal{S}(c)$
- Service discipline combination of FCFS and ALIS:
  - Server picks longest waiting compatible customer
  - Customer assigned to longest idle compatible server
Multi-skilled system: Example

- $\mathcal{C} = \{a, b, c\}$ and $\mathcal{S} = \{m_1, m_2, m_3\}$

- Skill-based service:
  - $\mathcal{C}(m_1) = \{a, b\}$, $\mathcal{C}(m_2) = \{a, c\}$, $\mathcal{C}(m_3) = \{a\}$
  - $\mathcal{S}(a) = \{m_1, m_2, m_3\}$, $\mathcal{S}(b) = \{m_1\}$, $\mathcal{S}(c) = \{m_2\}$
Multi-skilled system: Markov chain

- Markov chain with states \((M_1, n_1, M_2, \ldots, M_i, n_i, M_{i+1}, \ldots, M_J)\)
  - \(i\) busy servers \(M_1, \ldots, M_i\)
  - \(J - i\) idle servers \(M_{i+1}, \ldots, M_J\) with increasing idle times
  - \(M_1, \ldots, M_J\) is a permutation of \(m_1, \ldots, m_J\)
  - \(n_j\) number of waiting jobs between server \(M_j\) and \(M_{j+1}\)
Multi-skilled system: Examples of states

(i) All servers busy, servers move left to right, arrivals come from the right

(ii) Two servers busy, no possible customers between $m_3$ and $m_1$

(iii) Two servers are idle, $m_3$ has been idle longest time

(i) $(m_1, 3, m_2, 4, m_3, 2)$

(ii) $(m_3, 0, m_1, 3, m_2)$

(iii) $(m_1, 3, m_2, m_3)$
Multi-skilled system: N-system

- Markov chain with states
  - \((m_1, n_1, m_2, n_2)\) with \(n_1, n_2 \geq 0\)
  - \((m_2, 0, m_1, n_2)\) with \(n_2 \geq 0\)
  - \((m_1, n_1, m_2)\) with \(n_1 \geq 0\)
  - \((m_2, 0, m_1)\)
  - \((m_1, m_2)\) and \((m_2, m_1)\) (empty system)
- Assume \(\lambda_b < \mu_1\) and \(\lambda_a + \lambda_b < \mu_1 + \mu_2\)
Multi-skilled system: N-system

- State transition from \((m_1, n_1, m_2, n_2)\) to \((m_1, n_1 + j - 1, m_2, n_2 - j)\) with rate \(\mu_2 p_j\) where \(p_j = (1 - \gamma)\gamma^{j-1}\)

\[
\gamma = \frac{\lambda_b}{\lambda_a + \lambda_b}
\]
Multi-skilled system: N-system

- Global balance equation in \((m_1, n_1, m_2, n_2)\) with \(n_2 > 0\)

\[
(\lambda_a + \lambda_b + \mu_1 + \mu_2)p(m_1, n_1, m_2, n_2) = (\lambda_a + \lambda_b)p(m_1, n_1, m_2, n_2 - 1) + \mu_1p(m_1, n_1 + 1, m_2, n_2) \\
+ \mu_2(1 - \gamma) \sum_{j=0}^{n_1} \gamma^j p(m_1, n_1 - j, m_2, n_2 + j + 1) \\
+ \mu_2(1 - \gamma)\gamma^{n_1} p(m_2, 0, m_1, n_1 + 1 + n_2)
\]
Multi-skilled system: Partial balance

(i) Rate out of state \((M_1, n_1, M_2, \ldots, M_i, n_i, M_{i+1}, \ldots, M_J)\) due to an arrival activating a server equals rate into that state due to a departure deactivating \(M_{i+1}\)

(ii) Rate out of state \((M_1, n_1, M_2, \ldots, M_i, n_i, M_{i+1}, \ldots, M_J)\) due to an arrival that is queueing equals rate into that state due to a departure from a server staying active

(iii) Rate out of state \((M_1, n_1, M_2, \ldots, M_i, 0, M_{i+1}, \ldots, M_J)\) with \(n_i = 0\) due to a departure equals rate into that state due to an arrival activating \(M_i\)

(iv) Rate out of state \((M_1, n_1, M_2, \ldots, M_i, n_i, M_{i+1}, \ldots, M_J)\) with \(n_i > 0\) due to a departure equals rate into that state due to an arrival that is queueing
Multi-skilled system: N-system

(iv) Rate out of state \((m_1, n_1, m_2, n_2)\) with \(n_i > 0\) due to departure equals rate into that state due to an arrival that is queueing

\[
(\lambda_a + \lambda_b + \mu_1 + \mu_2)p(m_1, n_1, m_2, n_2) = (\lambda_a + \lambda_b)p(m_1, n_1, m_2, n_2 - 1) + \mu_1 p(m_1, n_1 + 1, m_2, n_2)
\]

\[
+ \mu_2(1 - \gamma) \sum_{j=0}^{n_1} \gamma^j p(m_1, n_1 - j, m_2, n_2 + j + 1)
\]

\[
+ \mu_2(1 - \gamma) \gamma^{n_1} p(m_2, 0, m_1, n_1 + 1 + n_2)
\]
Multi-skilled system: N-system

(ii) Rate out of state \((m_1, n_1, m_2, n_2)\) due to an arrival that is queueing equals rate into that state due to a departure from a server staying active

\[
(\lambda_a + \lambda_b + \mu_1 + \mu_2)p(m_1, n_1, m_2, n_2) = (\lambda_a + \lambda_b)p(m_1, n_1, m_2, n_2 - 1) + \mu_1 p(m_1, n_1 + 1, m_2, n_2)
+ \mu_2 (1 - \gamma) \sum_{j=0}^{n_1} \gamma^j p(m_1, n_1 - j, m_2, n_2 + j + 1) \\
+ \mu_2 (1 - \gamma) \gamma^{n_1} p(m_2, 0, m_1, n_1 + 1 + n_2)
\]
Multi-skilled system: N-system

- Partial balance dictates solution

\[ p(m_1, n_1, m_2, n_2) \stackrel{(iv)}{=} p(m_1, n_1, m_2, n_2 - 1) \frac{\lambda_a + \lambda_b}{\mu_1 + \mu_2} \]

\[ p(m_1, n_1, m_2, 0) \left( \frac{\lambda_a + \lambda_b}{\mu_1 + \mu_2} \right)^{n_2} \]

\[ p(m_1, n_1, m_2) \frac{\lambda_a}{\mu_1 + \mu_2} \left( \frac{\lambda_a + \lambda_b}{\mu_1 + \mu_2} \right)^{n_2} \]

\[ p(m_1, 0, m_2) \left( \frac{\lambda_b}{\mu_1} \right)^{n_1} \frac{\lambda_a}{\mu_1 + \mu_2} \left( \frac{\lambda_a + \lambda_b}{\mu_1 + \mu_2} \right)^{n_2} \]

\[ p(m_1, m_2) \frac{\lambda_a + \lambda_b}{\mu_1} \left( \frac{\lambda_b}{\mu_1} \right)^{n_1} \frac{\lambda_a}{\mu_1 + \mu_2} \left( \frac{\lambda_a + \lambda_b}{\mu_1 + \mu_2} \right)^{n_2} \]
Multi-skilled system: N-system

- Candidate solution

\[
p(m_1, n_1, m_2, n_2) = p(m_1, m_2) \frac{\lambda_a + \lambda_b}{\mu_1} \left( \frac{\lambda_b}{\mu_1} \right)^{n_1} \frac{\lambda_a}{\mu_1 + \mu_2} \left( \frac{\lambda_a + \lambda_b}{\mu_1 + \mu_2} \right)^{n_2}
\]

\[
p(m_2, 0, m_1, n_2) = p(m_1, m_2) \frac{\lambda_a}{\mu_2} \left( \frac{\lambda_a + \lambda_b}{\mu_1 + \mu_2} \right)^{n_2+1}
\]

\[
p(m_1, n_1, m_2) = p(m_1, m_2) \frac{\lambda_a + \lambda_b}{\mu_1} \left( \frac{\lambda_b}{\mu_1} \right)^{n_1}
\]

\[
p(m_2, 0, m_1) = p(m_1, m_2) \frac{\lambda_a}{\mu_2}
\]

\[
p(m_2, m_1) = p(m_1, m_2) \frac{\lambda_a}{\lambda_a + \lambda_b}
\]
Multi-skilled system: N-system

- Set $C = p(m_1, m_2)(\lambda_a + \lambda_b)\lambda_a$ then candidate solution

$$p(m_1, n_1, m_2, n_2) = C \frac{\lambda_b^{n_1}}{\mu_1^{n_1+1}} \frac{(\lambda_a + \lambda_b)^{n_2}}{(\mu_1 + \mu_2)^{n_2+1}}$$

$$p(m_2, 0, m_1, n_2) = C \frac{1}{\mu_2 (\mu_1 + \mu_2)^{n_2+1}} (\lambda_a + \lambda_b)^{n_2}$$

$$p(m_1, n_1, m_2) = C \frac{\lambda_b^{n_1}}{\mu_1^{n_1+1}} \frac{1}{\lambda_a}$$

$$p(m_2, 0, m_1) = C \frac{1}{\mu_2 (\lambda_a + \lambda_b)}$$

$$p(m_2, m_1) = C \frac{1}{\lambda_a + \lambda_b (\lambda_a + \lambda_b)}$$

$$p(m_1, m_2) = C \frac{1}{\lambda_a + \lambda_b \lambda_a}$$

- Checking partial balance equations proves that this is solution to the global balance equations
Multi-skilled system

- $\mathcal{C}$ is set of customer types \{a, b, $\ldots$\}
- $\mathcal{S}$ is set of servers \{m$_1$, m$_1$, $\ldots$, m$_J$\}
- $\lambda_{X} = \sum_{c \in X} \lambda_c$ where $X \subset \mathcal{C}$
- $\mu_Y = \sum_{M \in Y} \mu_M$ where $Y \subset \mathcal{S}$
- $\mathcal{S}(X)$ is set of server types that can handle customers in $X$
- $\mathcal{C}(Y)$ is set of customer types that can be handled by servers in $Y$
- $\mathcal{U}(Y)$ is set of customer types that can be uniquely handled by servers in $Y$
- Assume (stability) $\lambda_{\mathcal{U}(Y)} < \mu_Y$ for every subset of servers $Y \subset \mathcal{S}$
Multi-skilled system

• Solution to the global balance equations

\[
p(M_1, n_1, \ldots, M_i, n_i, M_{i+1}, \ldots, M_J) = C \prod_{j=1}^{i} \frac{\lambda_{n_j}^{n_j}}{\mu_{\{M_1, \ldots, M_j\}}^{n_j+1}} \prod_{j=i+1}^{J} \lambda_{C((M_j, \ldots, M_J))}^{-1}
\]

where \( C \) is a normalizing constant

• Proof by checking partial balance equations
Multi-skilled system: Alternative Markov chain

- Markov chain with states \((C_1, \ldots, C_L, M_1, \ldots, M_K)\)
- \(L\) customers waiting for service \(C_1, \ldots, C_L\) with decreasing waiting times \((L \geq 0)\)
- \(K\) idle servers \(M_1, \ldots, M_K\) with decreasing idle times \((K \leq J)\)
- Note that \(L = n_1 + \ldots + n_i\) and \(M_1, \ldots, M_K\) correspond to \(M_J, \ldots, M_{i+1}\)
- Solution to the global balance equations

\[
p(C_1, \ldots, C_L, M_1, \ldots, M_K) = C \prod_{l=1}^{L} \frac{\lambda_{C_l}}{\mu \delta(C_1, \ldots, C_l)} \prod_{k=1}^{K} \frac{\mu M_k}{\lambda \epsilon(M_1, \ldots, M_k)}
\]

where \(C\) is a normalizing constant

- Proof by checking partial balance equations
- Note that

\[
p(C_1, \ldots, C_L|\text{all servers busy}) \propto \prod_{l=1}^{L} \frac{\lambda_{C_l}}{\mu \delta(C_1, \ldots, C_l)}
\]
FCFS matching queue

- Independent Poisson arrivals of customers with rates $\lambda_c, c \in C$
- Independent Poisson arrivals of servers with rates $\mu_{m_j}, m_j \in S$
- Arriving server $m_j$ scans queue of customers
  - Matches with the longest waiting compatible customer and the two leave the system immediately
  - If no match is possible he leaves immediately without customer
- Note that this system is equivalent with a redundancy service system
**FCFS matching queue: Markov chain**

- Markov chain with states \((C_1, \ldots, C_L)\)
- \(L\) customers waiting for service \(C_1, \ldots, C_L\) with decreasing waiting times \((L \geq 0)\)
- Solution to the global balance equations

\[
p(C_1, \ldots, C_L) = C \prod_{l=1}^{L} \frac{\lambda C_l}{\mu S(\{C_1, \ldots, C_l\})}
\]

where \(C\) is a normalizing constant