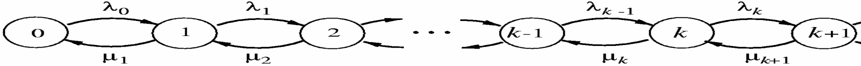
Processus de Naissance et de Mort

- *Une chaîne de Markov est un processus de Naissance et de Mort si et seulement si, pour tout état n, les seules transitions possibles amènent aux états $n-1$ et $n+1$, si ces états existent.*
- La transition de l'état n à l'état $n+1$ est une Naissance.
- La transition de l'état n à l'état $n-1$ est une Mort.
- λ_n (resp. μ_n) est le taux de Naissance (resp. de Mort) à l'état n.

Continuous Time Birth-Death Process (contd.

The state diagram of the birth-death process is given as



The generator matrix Q can be shown to be as

$$
= [q_{i,j}] = \begin{bmatrix}
-\lambda_0 & \lambda_0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 & 0 & \dots \\
\mu_1 & -(\mu_1+\lambda_1) & \lambda_1 & 0 & 0 & \dots & \dots & 0 & 0 & 0 & \dots \\
0 & \mu_2 & -(\mu_2+\lambda_2) & \lambda_2 & 0 & \dots & \dots & 0 & 0 & 0 & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \dots & \dots & & 0 & \mu_{n-2} & -(\mu_{n-2}+\lambda_{n-2}) & \lambda_{n-2} & 0 & \dots \\
0 & 0 & \dots & \dots & & 0 & 0 & \mu_{n-1} & -(\mu_{n-1}+\lambda_{n-1}) & \lambda_{n-1} & \dots \\
0 & 0 & \dots & \dots & & 0 & 0 & 0 & \mu_n & -(\mu_n+\lambda_n) & \lambda_n \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{bmatrix}
$$

Copyright © 2006 by K.S. Trivedi

Equilibre Stationnaire

- Ecrivons l'équation d'équilibre stationnaire pour X_t :

$$\pi(n)[\mu_n 1_{\{n>0\}} + \lambda_n] \;=\; \pi(n-1)\lambda_{n-1}1_{\{n>0\}} + \pi(n+1)\mu_{n+1} \qquad (1)$$

- Pour résoudre, considérons l'équation à l'état 0 pour obtenir $\pi(1)$ en fonction de $\pi(0)$.

$$\pi(0)\lambda_0 \;=\; \pi(1)\mu_1 \qquad (2)$$

- Examinons maintenant l'équation à l'état 1.

$$\pi(1)[\mu_1 + \lambda_1] \;=\; \pi(0)\lambda_0 + \pi(2)\mu_2 \qquad (3)$$

- Après simplification : $\pi(1)\lambda_1 \;=\; \pi(2)\mu_2$

- Par récurence sur n :

$$\pi(n)\lambda_n \;=\; \pi(n+1)\mu_{n+1} \tag{4}$$

- Soit :

$$\pi(n) \;=\; \pi(0)\frac{\prod_{i=0}^{n-1}\lambda_i}{\prod_{i=1}^{n}\mu_i} \tag{5}$$

- Si on peut normaliser :

$$\sum_{n=1}^{\infty}\frac{\prod_{i=0}^{n-1}\lambda_i}{\prod_{i=1}^{n}\mu_i} < \infty \tag{6}$$

- alors,

$$\pi(0) = \left[1 + \sum_{n=1}^{\infty}\frac{\prod_{i=0}^{n-1}\lambda_i}{\prod_{i=1}^{n}\mu_i}\right]^{-1} \tag{7}$$

Processus de Naissance et de Mort simple

- Description par une file d'attente (capacité = stockage dans la file, nombre de serveurs = puissance du service).
- Capacité Infinie, 1 serveur
- Capacité Infinie, m serveurs
- Capacité Finie, 1 serveur
- Capacité Finie, m serveurs
- Capacité Infinie, Infinité de serveurs

La file M/M/1

- Arrivées Poisson, Service iid Exponentiel, 1 Serveur, Capacité infinie pour garder les clients.
- On note μ le taux de service et λ le taux d'arrivées.
- C'est un processus de Naissance et de Mort avec les taux de transition :

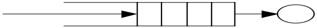
$$\begin{cases} \lambda_i = \lambda & \forall i \\ \mu_i = \mu & \forall i \end{cases} \tag{8}$$

- Posons $\rho = \lambda/\mu$.

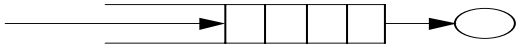
M/M/1 Queue

Arrivals process is Poisson, i.e., *interarrival* times are all $i.i.d$ EXP(λ).

Poisson arrival
Process with rate λ

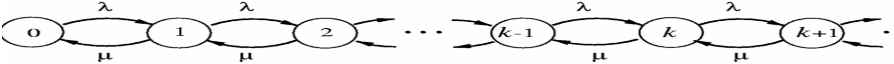


Service times are $i.i.d$, EXP(μ).



$\underbrace{\text{total \#customer at } t}_{N(t)}$

$N(t)$ is a birth-death process, with $\lambda_k = \lambda$; $\mu_k = \mu$.



Define, $\rho = \lambda / \mu$ (traffic intensity, in Erlangs)

Copyright © 2006 by K.S. Trivedi

- Equation d'équilibre global :

$$\pi(x)[\mu 1_{\{x>0\}} + \lambda] \;=\; \pi(x-1)\lambda 1_{\{x>0\}} + \pi(x+1)\mu \qquad (9)$$

- Condition d'ergodicité :
  - la chaine est ergodique si et seulement si ($\rho < 1$). Dans ce cas la solution stationnaire est $\pi(x=k) = (1-\rho)\rho^k$.
  - Lorsque $\rho = 1$ le processus est récurent nul.
  - Si $\rho > 1$ le processus est transitoire.
- Par la suite on suppose que le processus est ergodique

Statistiques Simples

- La moyenne du nombre de clients $E(N)$ dans la file :

$$E(N) = \sum_{i=1}^{\infty} i\pi(x = i) = \frac{\rho}{1 - \rho} \tag{10}$$

- La probabilité que la file soit vide est $(1 - \rho)$
- Le taux d'utilisation de la file U vaut donc ρ.
- Variance $V(N)$ du nombre de clients dans la file :

$$V(N) = E(N^2) - (E(N))^2 = \sum_{i=1}^{\infty} i^2\pi(x = i) - (E(N))^2 = \frac{\rho}{(1 - \rho)^2} \tag{11}$$

- Probabilité qu'il y ait au moins n clients dans la file : ρ^n.
- **Formule de Little :** $E[N] = E[R].\lambda$

$$E[R] = \frac{1}{\mu - \lambda}$$

Effet d'échelle

- Un système avec beaucoup de clients est plus efficace qu'un système avec peu de clients.
- Sur un exemple :
- Arrivées : 5 clients/h. Service moyen de 6mn.
- Donc la charge est 0.5 et le temps moyen de réponse est de $12mn$.
- Mais si on double le taux d'arrivées tout en divisant par deux le temps de service (les clients sont plus petits mais plus nombreux, ou le serveur est plus rapide)
- La charge reste 0.5 mais le temps moyen de réponse devient $6mn$.

Effet d'échelle

- A combien peut-on faire monter le taux d'arrivées tout en gardant le délai initial de $12mn$
$$12 = \frac{1}{\mu - \lambda}$$
  $\lambda = 15$ clients/h
- Un serveur rapide peut fournir des performances supérieures en temps moyen de réponse avec un effet plus que linéaire.

Capacité et Dimensionnement

- Choisir le taux de service pour assurer un temps moyen de service pour un taux d'arrivée donné.
- Soit R le temps de réponse,
- La capacité C est le débit nominal du serveur $= \mu$

$$R = \frac{1}{C - \lambda}$$

- Donc, $C = \lambda + 1/R$
- $C \geq \lambda$ assure la stabilité du système, le terme $1/R$ est le cout pour obtenir les performances demandées sur un système aléatoire simple.

Analyse $M/M/m$

- C'est un processus de Naissance et de Mort caractérisé par :

$$\begin{cases} \lambda_i = \lambda & \forall\ i \\ \mu_i = i\mu \quad si & i \leq m \\ \mu_i = m\mu \quad si & i > m \end{cases} \tag{12}$$

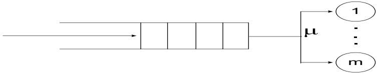
- La probabilité qu'il y ait au moins m clients dans le système est une quantité importante notée γ
- C'est donc aussi la probabilité qu'un client entrant soit contraint d'attendre avant de commencer son service

$$\gamma = \frac{(m\rho)^m}{m!(1-\rho)}\pi_0 \tag{13}$$

M/M/m queue

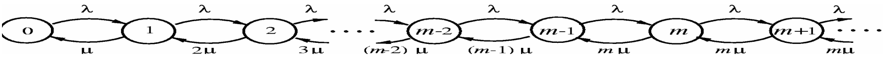
- m-servers service the queue.

Poisson arrivals
(rate λ)



$$\lambda_k = \lambda, \quad k = 0, 1, 2, \ldots,$$

$$\mu_k = \begin{cases} k\mu & k = 0, 1, 2, \ldots, m \\ m\mu & k > m \end{cases}$$



Copyright © 2006 by K.S. Trivedi

La file M/M/m	
Caractéristiques Spatiales	
Charge	$\rho = \lambda/(m\mu)$
Probabilité système vide	$\pi_0 = \left(1 + \frac{(m\rho)^m}{m!(1-\rho)} + \sum_{i=1}^{m-1} \frac{(m\rho)^i}{i!}\right)^{-1}$
Probabilité d'avoir k clients	si $k < m \quad \pi_0 \frac{(m\rho)^k}{k!}$
	sinon $\quad \pi_0 \frac{m^m \rho^k}{m!}$
Probabilité d'attente	$\gamma = \frac{(m\rho)^m}{m!(1-\rho)} \pi_0$
Nombre moyen de clients	$E(N) = m\rho + \frac{\rho\gamma}{1-\rho}$
Variance du nombre de clients	$V(N) = m\rho + \rho\gamma(m + \frac{1+\rho-\rho\gamma}{(1-\rho)^2})$
Nombre moyen de clients en attente	$E(N) = \frac{\rho\gamma}{(1-\rho)}$
Variance du nombre de clients en attente	$V(N) = \frac{\rho\gamma(1+\rho-\rho\gamma)}{(1-\rho)^2}$

La file M/M/m	
Caractéristiques Temporelles	
Temps de réponse Moyen	$E(R) = \frac{1}{\mu}(1 + \frac{\gamma}{m(1-\rho)})$
Variance du temps de réponse	$V(R) = \frac{1}{\mu^2}(1 + \frac{\gamma(2-\gamma)}{m^2(1-\rho)^2})$
Temps d'attente moyen	$E(W) = \frac{\gamma}{m\mu(1-\rho)}$
Variance du temps d'attente	$V(W) = \frac{\gamma(2-\gamma)}{m^2\mu^2(1-\rho)^2}$

La file M/M/m/B

- Il y a B places et m serveurs. Donc $B > m$.
- Le processus du nombre de clients est encore un processus de naissance et de mort dont les taux sont :

$$\left\{\begin{array}{ll} \lambda_i = \lambda & \forall\ i < B \\ \lambda_B = 0 & \\ \mu_i = i\mu & si \quad i \leq m \\ \mu_i = m\mu & si \quad m < i \leq B \end{array}\right. \tag{14}$$

- Puisque la chaîne a un nombre fini d'états, le système est toujours stable.
- Les probabilités stationnaires sont obtenues simplement :

$$\pi_n = \left\{\begin{array}{ll} \frac{\pi_0 \lambda^n}{n!\ \mu^n} & \forall\ n < m \\ \frac{\pi_0 \lambda^n}{m!\ m^{n-m}\ \mu^n} & \forall\ n \geq m \end{array}\right. \tag{15}$$

- et π_0 est obtenu par normalisation.

- Toutes les arrivées qui se produisent pendant que le buffer est dans l'état B sont perdues.
- Charge : $\rho = \lambda/(m\mu)$
- $\pi_0 = \left(1 + \frac{(m\rho)^m(1-\rho^{B+1-m}}{m!(1-\rho)} + \sum_{i=1}^{m-1} \frac{(m\rho)^i}{i!}\right)^{-1}$
- Probabilité d'avoir k clients si $k \leq m \quad \pi_0 \frac{(m\rho)^k}{k!}$
  sinon $\pi_0 \frac{m^m \rho^k}{m!}$
- Nombre moyen de clients : $E(n) = \sum_{i=1}^{B} i\pi_i$
- Nombre moyen de clients en attente : $E(q) = \sum_{i=m+1}^{B} (i-m)\pi_i$
- Taux d'arrivée réel : $\lambda(1-\pi_B)$
- Utilisation moyenne d'un serveur :$\rho(1-\pi_B)$
- Taux de perte : $\lambda\pi_B$

La file M/M/m/B	
Caractéristiques Temporelles	
Temps de réponse moyen	$E(R) = \frac{E(n)}{\lambda(1-\pi_B)}$
Temps d'attente moyen	$E(W) = \frac{E(q)}{\lambda(1-\pi_B)}$

%

Chaînes de Markov en temps discret

- On travaille sur un processus à valeur dans un espace d'états dénombrable.

- **Définition (Markov)**
  Un processus X_t est Markovien si et seulement si

$$Pr(X_{t+1} = j | x_t = i_t, x_{t-1} = i_{t-1} \ldots, x_0 = i_0) = Pr(X_{t+1} = j | X_t = i_t) \quad (16)$$

- On note $Pr(X_{t+1} = j | X_t = i_t) = P_{i,j}(t)$.

- **Définition (Chaîne homogène)**
  Une chaine de Markov est homogène si et seulement si $P_{i,j}(t)$ est indépendant de la date t. Cette probalilité, noté $P_{i,j}$, est la probabilité de transition de i vers j.

DTMC, Matrice Stochastique

Par construction, la matrice P vérifie les propriétés suivantes :

- tous les éléments de P sont positifs ou nuls
- pour tout i, $\sum_j P_{i,j} = 1$

Trajectoire

- Supposons qu'à l'instant 0, le processus soit connu (valeur de la probabilité à l'instant 0 $Pr(X_0 = i)\ \forall i$)
- Appliquons le théorème de conditionnement

$$Pr(X_1 = j) = \sum_i Pr(X_1 = j | X_0 = i) Pr(X_0 = i) = \sum_i P_{i,j} Pr(X_0 = i) \tag{17}$$

- Soit $\pi(t)$ le vecteur des probabilités $Pr(X_t = i)$.

$$\pi(1) = \pi(0)P \quad et \quad \pi(k) = \pi(0)P^k \tag{18}$$

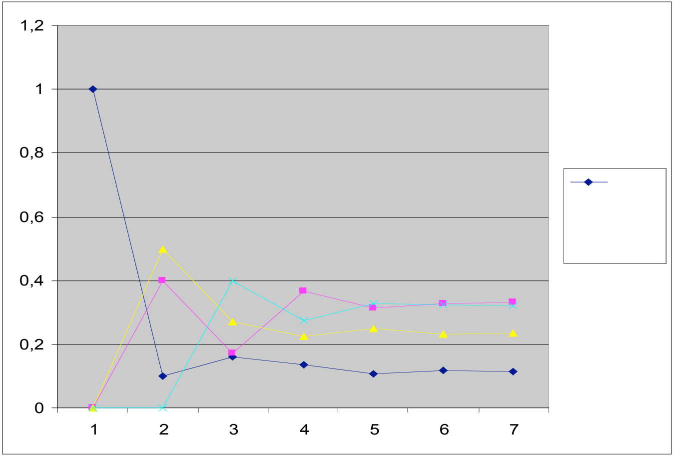
Ce qui donne un algorithme simple de calcul de la distribution à la date k.

Problèmes d'absorption, de transitoires et de stationnaire

- Existence d'une limite au vecteur $\pi(t)$ lorsque le temps t temps vers l'infini ?
- Que vaut $\pi(t)$ à une date quelconque ?
- La chaîne est elle absorbée ?
- La chaîne revient elle toujours dans certains états ?
- Etablir une classification des états et des chaînes.
- Tenir compte de la taille

Premier Exemple : probabilités limites

$$\begin{pmatrix} 0.1 & 0.4 & 0.5 & 0 \\ 0 & 0.2 & 0.3 & 0.5 \\ 0.3 & 0.1 & 0.2 & 0.4 \\ 0.1 & 0 & 0 & 0.9 \end{pmatrix}$$



Convergence des puissances

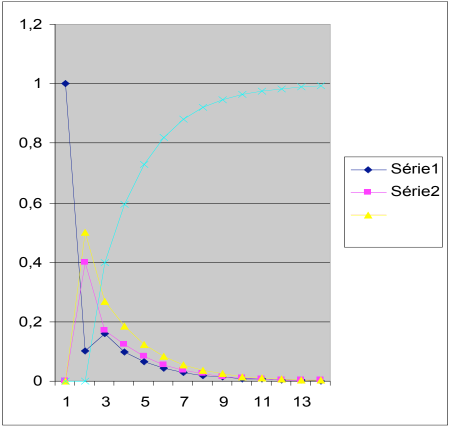
$$A^2 \quad \begin{bmatrix} 0{,}16 & 0{,}17 & 0{,}27 & 0{,}4 \\ 0{,}14 & 0{,}07 & 0{,}12 & 0{,}67 \\ 0{,}13 & 0{,}16 & 0{,}22 & 0{,}49 \\ 0{,}1 & 0{,}04 & 0{,}05 & 0{,}81 \end{bmatrix}$$

$$A^3 \quad \begin{bmatrix} 0{,}137 & 0{,}125 & 0{,}185 & 0{,}553 \\ 0{,}117 & 0{,}082 & 0{,}115 & 0{,}686 \\ 0{,}128 & 0{,}106 & 0{,}157 & 0{,}609 \\ 0{,}106 & 0{,}053 & 0{,}072 & 0{,}769 \end{bmatrix}$$

$$A^{20} = \begin{bmatrix} 0{,}112339 & 0{,}068141 & 0{,}095766 & 0{,}723754 \\ 0{,}112338 & 0{,}068140 & 0{,}095765 & 0{,}723756 \\ 0{,}112339 & 0{,}068141 & 0{,}095765 & 0{,}723755 \\ 0{,}112339 & 0{,}068140 & 0{,}095764 & 0{,}723758 \end{bmatrix}$$

Deuxième exemple : être absorbé en 4

0,1	0,4	0,5	0
0	0,2	0,3	0,5
0,3	0,1	0,2	0,4
0	0	0	1



Classification des Etats

On définit pour chaque état quatre quantités liées aux dates de retour à cet état.

- f_j^n est la probabilité de l'événement "(le premier retour en j à lieu en n transitions)".
- f_j est la probabilité de retour en j : $f_j = \sum_{n=1}^{\infty} f_j^n$
- M_j est le temps moyen de retour en j : $M_j = \sum_{n=1}^{\infty} n \times f_j^n$
- γ est le pgcd des valeurs de n telles que f_j^n est non nulle. γ est la période de retour.

Transience ou Récurence

- Si $f_j < 1$ l'état est transitoire.
- Si $f_j = 1$ l'état est récurent :
  - si $M_j = \infty$, l'état est recurent nul.
  - si $M_j < \infty$, l'état est recurent non nul.

  De plus,
  - si $\gamma > 1$, l'état est périodique de période γ.
  - si $\gamma = 1$, l'état est apériodique.

Classification des chaines

- Nécessaire d'examiner leur structure.

- Définition
  Un sous ensemble A d'états est fermé si et seulement si il n'y a pas de transition entre cet ensemble et son complémentaire.
  $\sum_{i \in A} \sum_{j \in \overline{A}} P_{i,j} = 0$

- Définition (Etat absorbant)
  un état absorbant est un sous ensemble fermé ne comprenant qu'un seul élément.

- Définition
  Une chaine est irréductible si et seulement si le graphe orienté de la chaine est fortement connexe. C'est à dire si il existe un suite de transitions menant de i à j pour tous les états i et j.

- Pour analyser une chaine réductible, il faut la décomposer en composantes fortement connexes.

- Théorème

  Soit une chaine de Markov irréductible, tous les états sont :
    - *tous transitoires*
    - *ou tous récurents non nuls*
    - *ou tous récurents nuls*

  De plus, si un état est périodique, alors tous les états le sont avec la même période.

Existence d'une limite

Théorème

Soit une chaine de Markov, irréductible et apériodique, alors la distribution limite $\pi = lim_{t \to \infty} \pi(t)$ *existe et est indépendante de la distribution initiale* $\pi(0)$ *(on dit qu'elle est ergodique). De plus,*

- *soit tous les états sont transitoires ou récurents nuls et dans ce cas* $\pi(j) = 0$*, pour tout* j*.*
- *soit tous les états sont récurents non nuls et* π *est solution unique des deux équations :*

$$\begin{cases} \pi = \pi P \\ \pi e = 1 \end{cases} \tag{19}$$

  où e *est un vecteur colonne dont tous les éléments sont* 1*. De plus, on a* $\pi(j) = 1/M_j$

Chaîne Finie

- Plus simple : pas de récurrence nulle.

- Théorème
  Toute chaine finie,irréductible et apériodique est ergodique.

Résoudre le stationnaire

- Analytiquement (voir plus loin)
- Numériquement (pourquoi pas ?)
- Simulation (en faisant très attention)

Résoudre états absorbants

- Probabilité d'être absorbée sachant que l'on commence en i
- Temps moyen avant d'être absorbée sachant que l'on commence en i
- Hypothèse : il y a plusieurs points absorbants et pas de classes récurentes (pretraitement pour fusionner chaque classe récurente en un sommet).

Pb des points absorbants

- On suppose que l'espace des états est partitionné : les états absorbants sont en tête.
- La matrice P peut être décomposée en bloc $\left[\begin{array}{c|c} Id & 0 \\ \hline R & Q \end{array}\right]$
- On peut prouver que $P^2 = \left[\begin{array}{c|c} Id & 0 \\ \hline R + QR & Q^2 \end{array}\right]$
- Et que, pour tout n, $P^n = \left[\begin{array}{c|c} Id & 0 \\ \hline (\sum_{j=0}^{n-1} Q^j)R & Q^n \end{array}\right]$

Matrice Fondamentale

- P^n est la matrice de transition pour n sauts successifs.
- $lim_{n\to\infty} P^n[i,j]$ est la probabilité d'être en j à l'infini sachant que l'on a débuté en i.
- La matrice $M = (Id - Q)^{-1}$ existe si Q ne contient pas de classe récurrente et elle vaut $\sum_{j=0}^{\infty} Q^j$.
- Et donc,

$$lim_{n\to\infty} P^n = \left[\begin{array}{c|c} Id & 0 \\ \hline MR & 0 \end{array}\right]$$

- M est la matrice fondamentale.

Probabilités d'être absorbé

- L'entrée $[i, j]$ de la matrice produit $M * R$ donne la probabilité d'être absorbé en j sachant que le point initial est i (point non absorbant).
- Si π_O est la distribution initiale la probabilité d'être absorbé en j s'obtient par conditionnement sur l'état initial

$$\mu(absorbe\ en\ j) = \sum_i (MR)[i, j]\pi_0(i)$$

- Si il y a un seul point, la proba est 1 (inutile de faire des calculs).

Temps Moyen avant d'être absorbé

- Soit $X_{i,j}$ le nombre de visite achant que l'on a débuté en i avant d'être absorbé.
- On calcule dans un premier temps $E[X_{i,j}]$.
- Le temps moyen avant d'être absorbé sachant que l'on a débuté en i est alors : $\sum_j E[X_{i,j}]$.
- Le temps moyen avant d'être absorbé sachant la distribution initiale π_0 est alors : $\sum_i \sum_j \pi_0(i) E[X_{i,j}]$.
- On regoupe tous les points d'absorbtion en un seul (de numéro 1).

$E[X_{i,j}]$

- Théorème
  $E[X_{i,j}] = M(i,j)$
  - On retouve la matrice fondamentale.

Preuve

- Posons $\delta_{i,j}$ le symbole de Kronecker,
- En conditionnant sur la première étape on a

$$X_{i,j} = \begin{cases} \delta_{i,j} & \textit{avec proba } P(i,1) \\ X_{k,j} + \delta_{i,j} & \textit{avec proba } P(i,k) \end{cases}$$

- En passant aux espérances :

$$E[X_{i,j}] = \delta_{i,j} P(i,1) + \sum_{k>1} (E[X_{k,j}] + \delta_{i,j}) P(i,k)$$

- Apres regroupement :

$$E[X_{i,j}] = \delta_{i,j} + \sum_{k>1} E[X_{k,j}] P(i,k)$$

Preuve - suite et fin

- Posons $E[X_{i,j}] = Z[i,j]$.
- En passant à une formulation matricielle :

$$Z = I + QZ$$

- Et donc, $Z = (I - Q)^{-1} = M$