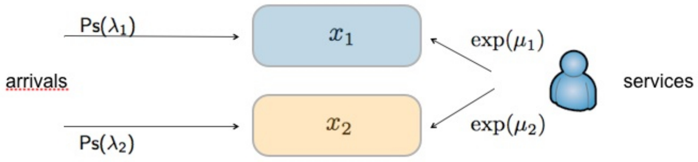
Continuity properties of parametrised and controlled countable state Markov processes

Floske Spieksma

Universiteit Leiden

6eme Journées COSMOS

Motivating example 1: competing queues



Holding cost: $c(x) = c_1 x_1 + c_2 x_2$.

How to allocate server so as to minimise the expected discounted / average expected holding cost?

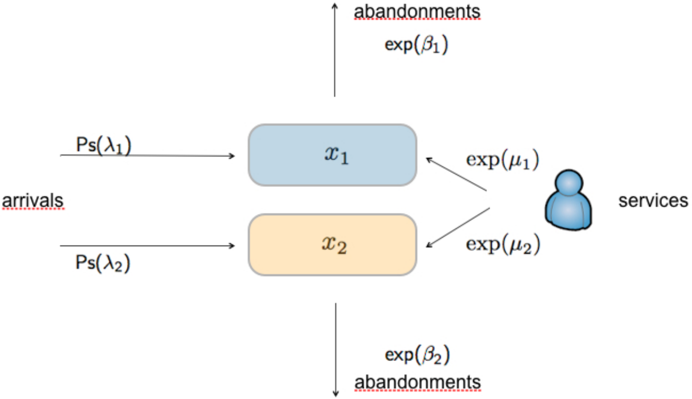
Index policy 'μc-rule' is optimal:

(Baras, Ma, Makowski; Buyukkoc, Varaiya, Walrand 1983, 1985)

if $c_1\mu_1 \geq c_2\mu_2$ then serve type 1 customers if they are present, otherwise type 2.

Proof: (1) via comparison arguments ('general' arrival processes)
(2) via event-based (Koole) dynamic programming (EDP)

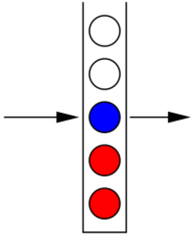
Motivating example 1:
competing queues with abandonments continued (S& Blok)



Question: how to allocate the server so as to minimise expected discounted / average holding cost?

When is μc-rule optimal? Maybe other index policy?

Motivating example 2: server farm (Adan, Kulkarni, van Wijk 2013)



- Busy servers require power
- Idle servers require less power
- Off servers require no power but do require setup cost
- Qn: discounted/average cost switch off optimal policy?

- Customers arrive according to a Poisson (λ) process
- Exponential (μ) service times ($\mathrm{P}\{S > t\} = e^{-\mu t}$)
- Infinitely many servers, can be On, Idle or Off
- Arriving job gets Idle server (if any), otherwise Off server
- Server changes state instantaneously
- Cost: $c(i)$ per unit time, if i Idle servers
- $K^{\mathrm{on}}(K^{\mathrm{off}})$ for switching On (Off) an off (on) server.

Solution

Typical approach using DP: Value Iteration (VI)

Problem: Models in continuous time

Required: Time-discretisation

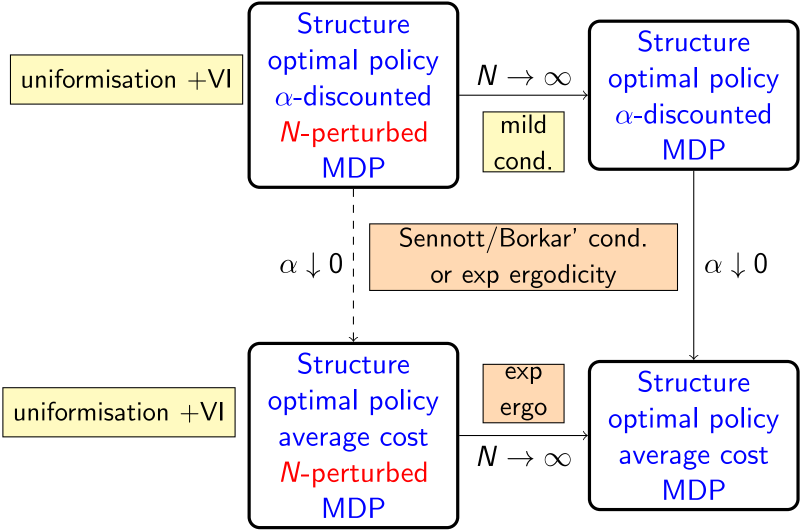
Problem: Unbounded rates $\Rightarrow$ not uniformisable
Structural properties do not propagate for VI on 'jump chain'

Required: Suitable perturbation leaving structural properties intact.

Goal:

- Solution procedure under verifiable conditions;
- Motivation for verifiable conditions

Solution procedure - aim via schematic overview



Overview of talk

- Parametrised Markov Chains in discrete time - discounted cost
- Parametrised Markov Processes in continuous time - discounted cost
- Solution for discounted cost problems of two examples
- Average cost - exponential ergodicity conditions with the desired continuity properties
- Average cost - adapted exponential ergodicity 'conditions' to continuous time

Restriction to stationary policies!

Most proof techniques well-known + extra bits and pieces

(cf. Hernandez-Lerma, Prieto-Rumeau, Feinberg, Piunovskiy, Guo, Zhang ..)

Point of view: motivation of conditions....

Parametrised Markov Chain (PMC) - discrete time

$\mathbf{S}$ is denumerable state space, A parameter set

Definition

A PMC $\{X(a)\}_{a\in A}$ *with transition matrices* $\{P(a)\}_{a\in A}$ *is defined by*

- *a collection of* $\mathbf{S} \times \mathbf{S}$ *transition matrices* $P(a)$, $a \in A$.

Definition

$\{P(a)\}_{a\in A}$ *has the product property if*

- A *has the product property i.e.* $A = \prod_{x\in S} A_x$, *(cf. Hordijk (1974))*;
- $a_x = a'_x$ *implies* $P_{x\cdot}(a) = P_{x\cdot}(a')$, $a, a' \in A$.

$\{P(a)\}_{a\in A}$ has product property, $\implies$ a is *stationary policy*.

Further:

- Perturbation case: $A = \{(N, \pi)\}$ with N perturbation parameter, π stationary or deterministic policy.
- $X : \omega \to \omega$ canonical process on trajectory space $\Omega = \mathbf{S}^\infty$, $P(a)$ defines probability measure P^a on Ω, exp.operator E^a.

Discrete time α-discounted reward

Given: immediate reward function $\{r(a)\}_{a \in \mathrm{A}}$, $r(a) : \mathbf{S} \to \mathbf{R}$. Then

$$v_x^\alpha(a) = \mathrm{E}_x^a \sum_{t \geq 0} \alpha^t r_{X_t}(a) = \sum_{t \geq 0} \alpha^t (P^t(a) r(a))_x$$

is (total expected) α-discounted reward, given starting state x.

Assumption 1

- A *locally compact;*
- *there exists a constant* c*, such that* $\sup_{x \in S, a \in \mathrm{A}} |r_x(a)| \leq c$*;*
- $a \mapsto P(a)$, $a \mapsto r(a)$ *continuous.*

Notation:

- continuity of $a \mapsto P(a)$ etc, is *componentwise continuity*!
- $P(a)f$ is vector with elements $\sum_y P_{xy}(a) f_y$.
- $\mathbf{1}$ is vector of 1's.
- $\ell^\infty(\mathbf{S}) = \{w : \mathbf{S} \to \mathbf{R} \,:\, \sup_x |w_x(a)| < \infty\}$

Discrete time α-discounted reward continued

Theorem

Assumption 1 implies

- $v^\alpha(a)$ *unique sol'n to equation* $w = r(a) + \alpha P(a) w$ *in* $\ell^\infty(\mathbf{S})$;
- $a \mapsto v^\alpha(a)$ *is continuous;*
- *if* $\{P(a)\}_a$, $\{r(a)\}_a$ *have* product property, A *compact, then* $v^\alpha = \sup_a v^\alpha(a)$ *finite and unique solution to* α*-DOE*

$$w = \sup_a \{r(a) + \alpha P(a) w)\}, \quad w \in \ell^\infty(\mathbf{S}),$$

  sup = max *and if* a^* *maximising, then* a^* α*-discount optimal.*
- *VI converges geometrically quickly: i.e. for*

$$v^\alpha_{n+1}(x) = \max_a \{r_x(a) + \alpha(P(a) v^\alpha_n)_x\}, \quad v^\alpha_0 \in \ell^\infty(\mathbf{S})$$

  then, $v^\alpha_n \to v^\alpha$, *and limit points of* $\{a_n\}_n$ *are* α*-disc. optimal*

Note: Continuity of $a \mapsto r(a), a \mapsto P(a)$, $a \mapsto P(a)\mathbf{1} = 1$ plus DOM CONV implies $a \mapsto v^\alpha(a)$ continuous.

Unbounded rewards and drift function

Assumption 2 (Lippman (1975), Wessels (1977))

There exist function $F : \mathbf{S} \to \mathbf{R}$, *constants* γ, ρ, *with for* $a \in A$

- $P(a)F \leq \gamma F$; $a \mapsto P(a)F$ *continuous;*
- $\|r(a)\|_F \leq \rho$, *where*

$\|f\|_F = \sup_x \frac{|f_x|}{F_x} < \infty$, *and* $\ell^\infty(\mathbf{S}, F) = \{f : \mathbf{S} \to \mathbf{R} \mid \|f\|_F < \infty\}$.

Use transformation

$$
P^F_{xy}(a) = \begin{cases} \frac{P_{xy}(a)F_y}{\gamma F_x}, & x, y \in \mathbf{S} \\ 1 - \frac{\sum_y P_{xy}(a)F_y}{\gamma F_x}, & x \in \mathbf{S}, y = \Delta \\ 0, & x = y = \Delta \end{cases}
$$

$$
r^F_x(a) = \begin{cases} \frac{r_x(a)}{F_x}, & x \in \mathbf{S} \\ 0, & x = \Delta. \end{cases}
$$

where Δ is ‘exit’ state, $\mathbf{S}_\Delta = \mathbf{S} \cup \{\Delta\}$.
Note $v^\alpha_x = v^{F,\gamma\alpha}_x F_x$, so it is required that $\alpha\gamma < 1$.

Continuous time: PMP with α-discounted reward

Definition

A PMP $\{X(a)\}_{a\in \mathrm{A}}$ with q-matrices $\{Q(a)\}_{a\in \mathrm{A}}$ is defined by

- *a collection of $\mathbf{S}\times\mathbf{S}$ conservative, stable q-matrices $Q(a)=\{q_{xy}(a)\}_{x,y\in S}$, $a\in \mathrm{A}$, with*
  - $\sum_y q_{xy}(a)=0$, $x\in\mathbf{S}$
  - $q_x(a)=-q_{xx}(a)<\infty.$
- *$P_t(a)$ is minimal transition function.*

Natural extension of discrete time conditions is following.

Assumption 3

- A *locally compact;*
- *reward* rate $r(a)$, $a\in \mathrm{A}$, *satisfies* $\sup_a \|r(a)\|_\infty = c<\infty$;
- *$a\mapsto Q(a)$, $a\mapsto r(a)$ continuous.*

QN: suff't to guarantee analogous properties on continuity, existence?

Assumption appropriate: continuity properties??

Denote: X canonical process on 'trajectory' space Ω,
P^a prob. measure, E^a exp.operator under a

α-discounted reward under parameter a:

$$v_x^\alpha(a) = \mathsf{E}_x^a \int_0^\infty r_{X_t}(a)dt = \int_0^\infty e^{-\alpha t}(P_t(a)r(a))_x dt$$

Then: v^α sol'n to

$$\alpha w = r + Q(a)w.$$

QN:

- $v^\alpha(a)$ unique solution in $\ell^\infty(\mathbf{S})$?
- $a \mapsto v^\alpha(a)$ continuous?

Answer Sol'n unique iff $X(a)$ non-explosive!
Then life is fine: $P_t(a)\mathbf{1} = \mathbf{1}$, and continuity holds.
Note (Reuter (57))
X explosive $\Leftrightarrow$ there exists $w \in \ell^\infty(\mathbf{S})$ with $Qw = \alpha w$.

Digression on explosiveness

Put: $X(a)$ canonical process on 'trajectory' space Ω,
P^a prob. measure, E^a exp.operator under a

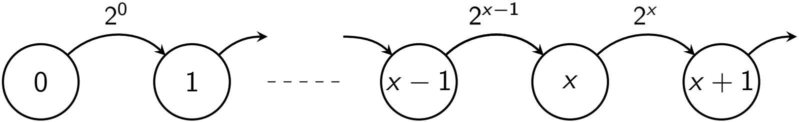
Define:

- $J_1 = \inf\{t \geq 0 \,|\, X_t \neq X_{t-}\} \wedge \infty$;
  $J_{n+1} = \inf\{t > J_n \,|\, X_t \neq X_{t-}\} \wedge \infty$;
- $J_\infty = \lim_{n\to\infty} J_n$;

Then:

- X is explosive for a
  iff $\mathrm{P}^a_x\{J_\infty < \infty\} > 0$ for some $x \in \mathbf{S}$
  iff there exists $\alpha > 0$ and $w \in \ell^\infty(\mathbf{S})$, s.t. $Qw = \alpha w$
  iff for all $\alpha > 0$ there exists ...
  iff there exists $x \in \mathbf{S}$ with $\sum_y P_{t,xy}(a) < 1$, for some $t \geq 0$
  and hence for all t.
- X is non-explosive for a if $\sup_i |Q_{ii}|(a) < \infty$.

Explosive parametrised example



Put $A = \{0, 1, 2, \ldots, \infty\}$,

$$Q_{x,y}(a) = \begin{cases} 2^{x \wedge a}, & y = x+1 \\ -2^{x \wedge a}, & y = x \end{cases}$$

Then $X(a)$ non-explosive for $a < \infty$, but explosive for $a = \infty$.

Then: $\lim_{a \to \infty} P_t(a)\mathbf{1} = \mathbf{1} \neq P_t(\infty)\mathbf{1}$! and
for $r(a) = \mathbf{1}$

- $v^\alpha(a) = \mathbf{1}/\alpha$, $a < \infty$
- $v_0^\alpha(\infty) = \frac{2^0}{2^0+\alpha} + \sum_{y=1}^{\infty} \prod_{z=0}^{y-1} \frac{2^z}{2^z+\alpha} \cdot \frac{1}{2^y+\alpha} \to 2$, as $\alpha \downarrow 0$!

Non-explosiveness: characterisation via time-discretisation

- Resolvent $R^\alpha(a) = \alpha \int_0^\infty e^{-\alpha t} P_t(a) dt$ ($\Longrightarrow$ $v^\alpha(a) = R^\alpha(a) r(a)\alpha$).
- α-jump chain $\{X_n^J(a)\}_n$, Markov chain on $\mathbf{S}_\delta = \mathbf{S} \cup \{\delta\}$, transition matrix $P^J(a)$,

$$\text{where} \quad P^J_{xy}(a) = \begin{cases} \frac{q_{xy}(a)}{\alpha+q_x(a)}, & x, y \in \mathbf{S} \\ \frac{\alpha}{\alpha+q_x(a)}, & x \in \mathbf{S}, y = \delta, \\ p_y > 0, & x = \delta, y \in \mathbf{S} \end{cases}$$

Then

- $R^\alpha_{xy}(a) = \alpha\cdot$(tot.exp.disc. time spent in y starting in x before reaching δ)
  $= \sum_{n=0}^\infty {}_\delta P^{J,n}_{xy}(a) \frac{\alpha}{\alpha+q_y(a)}$
- $\mathsf{P}_x\{\exists n \geq 0 \,|\, X_n^J(a) = \delta\} = \sum_y R^\alpha_{xy}(a)$, $x \in \mathbf{S}$

Theorem

$X(a)$ non-explosive iff $X^J(a)$ recurrent iff there exists a moment function $F : \mathbf{S} \to \mathbf{R}_+$ *and constant γ s.t. $Q(a)F \leq \gamma F$.*

$F : \mathbf{S} \to \mathbf{R}_+$ moment function if there exist $\{\mathbf{S}_n\}_n \subset \mathbf{S}$, with

- $\mathbf{S}_n \uparrow \mathbf{S}$, $\mathbf{S}_n$ finite
- $\inf_{x \in S} F_x \to \infty$, $n \to \infty$

Note: moment function $\Longrightarrow$ non-explosiveness: M.F. Chen (86), recurrence criterion - Mertens, Samual-Cahn, Zamier (78)

Reduction of total exp. discounted reward to time-discretised total exp. rewards

Notation:

- $\tau = \inf\{n \mid X_n = \delta\}$
- $r^J_x = \begin{cases} \frac{r_x}{\alpha + q_x}, & x \in \mathbf{S} \\ 0, & x = \delta \end{cases}$

Then:

$$v^\alpha = \mathsf{E}_x \sum_{n=0}^{\tau-1} r^J_{X^J_n}.$$

Allows value iteration for computing v^α.

cf. also recent work by Feinberg

Finally: α-discounting with bounded rewards

Assumption 4

- A *locally compact;*
- $\sup_a \|r(a)\|_\infty \leq c < \infty$*;*
- $a \mapsto Q(a)$, $a \mapsto r(a)$ *continuous.*
- *there exist a moment function* $F : \mathbf{S} \to \mathbf{R}_+$*, and constant* γ *such that* $Q(a)F \leq \gamma F$ *for all* $a \in \mathrm{A}$*.*

Theorem (many authors, Blok& S (2015))

Under above assumption, same continuity and optimality results hold as for discrete time discounted PMC, where now $\alpha > \gamma$ *and we have the equation*

$$\alpha v^\alpha(a) = r_x(a) + Q(a)v^\alpha(a),$$

and associated α*-discount optimality equation. For optimality part, also require:* $\{Q(a)\}_a, \{r(a)\}_a$ *have product property and* A *compact.*

Unbounded rewards via transformation

Let $Q(a)F \leq \gamma F$, $F\mathbf{S} \to \mathbf{R}_+$, $\gamma \in \mathbf{R}$. Consider F-transformed q-matrices $Q^F(a)$, and rewards $r^F(a)$ with

$$q^F_{xy}(a) = \begin{cases} q_{xy}(a)F_y/F_x, & y \neq x, y, x \in \mathbf{S} \\ q_{xx} - \gamma, & y = x \\ \gamma - \sum_{y \in S} q_{xy}(a)F_y/F_x, & x \in \mathbf{S}, y = \Delta \\ 0 & x = \Delta, \end{cases}$$

$$r^F_x(a) = \begin{cases} \frac{r_x(a)}{F_x}, & x \in \mathbf{S} \\ 0, & x = \Delta, \end{cases}$$

where Δ is an absorbing coffin state, to make Q^F conservative.
Then: This is PMP on $\mathbf{S}_\Delta = \mathbf{S} \cup \{\Delta\}$, with

$$v^\alpha_x(a) = v^{F,\alpha-\gamma}_x(a)F_x, \quad \alpha > \gamma.$$

Provides translation to bounded reward problem!

PMP with unbounded α-discounted reward

Assumption 5

- A *locally compact;*
- *there are a function* $F : \mathbf{S} \to \mathbf{R}_+$, γ *such that*
  - $Q(a)V \leq \gamma F$, $a \in$ A;
  - $\sup_a \|r(a)\|_F = c < \infty$;
  - $a \mapsto q_{xy}(a)_x$, $a \mapsto r_x(a)$ *continuous,* $x \in \mathbf{S}$.
  - *there exists an* F*-moment function* $W : \mathbf{S} \to \mathbf{R}_+$, *constant* θ, *s.t.* $QW(a) \leq \theta W$, $a \in$ A.

Note: X is allowed to be explosive under a.

Theorem (many authors, Blok& S (2015))

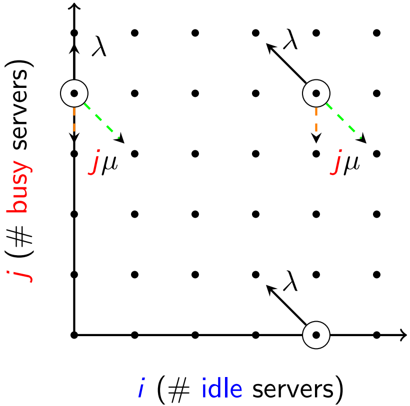
Under Assumption 5 and $\alpha > \gamma$ *same continuity and optimality results hold as for discrete time discounted PMC, where now*

$$\alpha v^\alpha(a) = r_x(a) + Q(a)v^\alpha(a),$$

and associated α*-discount optimality results , if* $\{Q(a)\}_a, \{r(a)\}_a$ *have product property,* A *compact.*

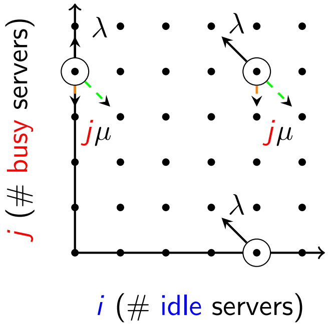
Application to perturbed server farm Adan, Kulkarni, vWijk

- State of system: (i, j) $\sim$ i idle, j busy servers
- Action a in state (i, j), $j > 0$: upon service completion
  switch off (action 0) or keep idle (action 1)
- Holding cost rate: $c(i)$, incr.convex holding cost for i idle servers.
- Lump costs: K^{on}, K^{off}.
  $\Rightarrow$ Total cost rate $c_{(i,j)}(a) = c(i) + K^{\mathrm{on}}\lambda\mathbf{1}_{\{i=0\}} + K^{\mathrm{off}} j\mu\mathbf{1}_{\{a=0\}}$



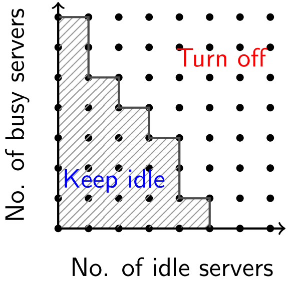
Application: server farm model

- Perturb $j\mu \quad \Rightarrow (j \wedge N)\mu$, then uniformisable.
- Parameter space $\mathsf{A} = \{(N, \pi)\}$:
  π stationary policy, N perturbation parameter of service rates.
- $F_{ij} = e^{\epsilon i + \delta j}$ satisfies $Q(a)F \leq \gamma F$.
- $W_{ij} = e^{\epsilon' i + \delta' j}$, $\epsilon' > \epsilon$, $\delta' > \delta$, satisfies $Q(a)W \leq \theta W$,
  $$W_{ij}/F_{ij} \to \infty \text{ for } (i,j) \to \infty.$$
- $a \mapsto q_{xy}(a), c_{(i,j)}(a) = ci + \lambda K^{\mathrm{on}} \mathbf{1}_{\{i=0\}} + K^{\mathrm{off}} \mathbf{1}_{\{a=0\}}$ continuous,
  $$\sup_a \|c(a)\|_F < \infty.$$



Result server farm (Blok&S 2015)

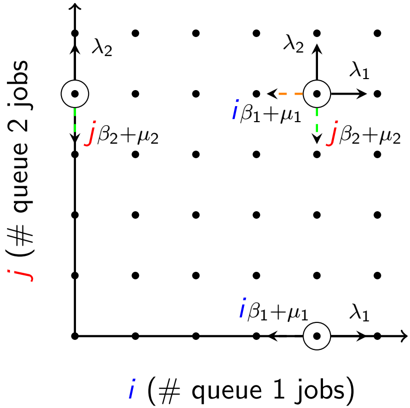
- $\{Q(N,\pi)\}_\pi$ is uniformisable MDP $\Rightarrow$ discrete time results give switching curve optimal policy π^N for N fixed, and necessary structural results of $v^\alpha(N,\pi^N)$ (Adan, Kulkarni, van Wijk (2013))
- for any limit point (∞,π^*) of $\{(N,\pi^N)\}_N$: $v^\alpha(N,\pi^N) \to v^\alpha(\infty,\pi^*)$, $v^\alpha(\infty,\pi^*)$ solves α-DOE, and has same structural properties. Hence, π^* is switching curve policy.



Remark: there exists a number B such that the α-discount optimal policy switches off in (i,j), with $i+j \geq B$, $i>0$, for all $\alpha \leq \alpha_0$.
$\Rightarrow$ Blackwell optimality result.

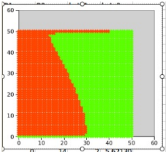
Question:
what about average optimality?

Application to competing queues with abandonments

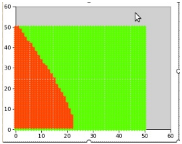


Perturbation possibilities:

1. Normal Perturbation
   $i\beta_1 \quad \Rightarrow \quad (i \wedge N)\beta_1$
   $j\beta_2 \quad \Rightarrow \quad (j \wedge M)\beta_2$
2. Smoothed Rate Truncation (Bhulai, Brooms, S (2013))
   Replace upward rate λ_i by state dependent rate $\lambda_i(x) = (1 - \frac{x_i}{N_i})^+ \lambda_i$.



Normal truncation



SRT

Result for discounted cost

Model satisfies Assumption 3 for $F_{(i,j)} = e^{\epsilon i + \delta j}$, ϵ, δ depend on α

Theorem (Blok& S)

Assume

- $c_1\mu_1 \geq c_2\mu_2$;
- $\mu_1 \leq \mu_2$;
- $c_1\mu_1/\beta_1 \geq c_2\mu_2/\beta_2$.

Serving the customer with the lowest index is α-discount optimal for SRT model for all N, hence for unperturbed model.

Remarks

- Result applies to a general number of queues.
- If conditions on μ, c, β are not satisfied: switching curve is optimal???
- What about average cost?

Average reward optimality

- Analysis via *vanishing discount approach*, setting $\alpha \to 1$ or 0, in discrete and continuous time.
  - Mild conditions introduced by Sennott '86, for negative dynamic programming (minimising non-negative cost).
  - Can be appropriately generalised to continuous time.
  - No need/possibility for parametrised set-up; only need
    - continuity results in the limit α to 0 (or 1), and
    - result that limit yields sol'n to optimality equation, and provides optimal policy.
  - Applies to server farm example and server allocation example.
- 'Direct' analysis, under more restrictive stability conditions, based on Dekker&Hordijk '80s, for discrete time.
  - Convergence of VI holds (e.g. S'90);
  - Parametrised framework has to be included in continuous (and discrete) time - done but results are not yet in paper..
  - Only applies to server allocation example.

Average reward for (discrete time) PMC Dekker, Hordijk (1988, 1992)

Assumption 6

- A *compact;*
- *There exists a function* $F : \mathbf{S} \to [1, \infty)$, $\beta \in (0, 1)$, *a finite set* $M \subset \mathbf{S}$ *such that*
  - ${}_M P(a) F \leq \beta F$ *(equivalent:* $\| {}_M P(a) \|_F \leq \beta$*);*
  - $\sup_a \| r(a) \|_F < \infty$*;*
  - $a \mapsto P(a)$, $a \mapsto (P(a))F$, $a \mapsto r(a)$ *continuous.*
  - *the number of closed classes in Markov chain for* $P(a)$ *is 1.*

${}_M P(a)$ is the taboo matrix with taboo set M defined by

$$
{}_M P_{xy}(a) = \begin{cases} P_{xy}, & y \notin M \\ 0, & y \in M \end{cases}
$$

Note: transient states are allowed.
Condition implies for $|z| < 1/\beta$ that
$\sum_{n \geq 0} \mathsf{P}_x \{ X_n(a) \in M, X_t(a) \notin M, t = 1, \ldots, n-1 \} z^n \leq F_x$!

Theorem (D&H 88,92, S?)

Under Assumption 6

- *the Poisson equation* $w = r(a) - g + P(a)w,\ w \in \ell^\infty(\mathbf{S}, F)$ *has the 'unique' solution pair* $(v(a), g(a))$, *with* $v(a) \in \ell^\infty(\mathbf{S}, F)$ *and*

$$g(a) = \liminf_{N\to\infty} \mathrm{E}_x^a \frac{\sum_{t=0}^N r_{X_t}(a)}{N+1} = \liminf_{N\to\infty} \frac{\sum_{t=0}^N P^t(a) r(a)}{N+1},$$

the expected average cost for parameter a.

- $a \mapsto g(a)$, $a \mapsto v(a)$ *is continuous.*
- *if* $\{P(a)\}_a$, $\{r(a)\}_a$ *have the* product property, A *compact,*
  - $g^* = \sup_a g(a) = \max_a g(a)$ *exists,*

$$w_x = \sup_a \{r(a) - g + \sum_y P_{xy}(a) w_y, \quad w \in \ell^\infty(\mathbf{S}, F)$$

has a 'unique' solution pair (v, g') *with* $g' = g^*$, *and* $v \in \ell^\infty(\mathbf{S}, F)$. *Further,* $\sup = \max$ *and any maximiser* a^* *is average optimal, and* $v(a^*) = v$, $g(a^*) = g^*$.

  - *there exists a minimising parameter yielding a Blackwell optimal policy.*

Connection with exponential ergodicity

Theorem (S'90, Altman, H & S (1997), + S?)

Under Assumption 6 +$P(a)$ aperiodic it holds that

- $P(a)$ *induces a Markov chain with finitely many closed classes, all positive recurrent, + transient states,* $a \in \mathsf{A}$*;*
- *there exist* $d \in \mathbf{R}$, $\phi \in (0,1)$ *s.t., with* $\Pi(a)$ *the stationary matrix,*

$$\sup_a \|P^n(a) - \Pi(a)\|_F \leq d\phi^n, n = 0, 1, \ldots,$$

- $v(a) = \sum_{n=0}^{\infty}(P^n(a) - \Pi(a))r(a).$
- *Product property assumption implies: VI converges, i.o.w. let* $v_{n+1} = \sup_a r(a) + P(a)v_n$, $v_0 \in \ell^\infty(\mathbf{S}, V)$. *Then*

$$v_n(x) - v_n(0) \to V(x), \quad \frac{v_n(x)}{n} \to g^*,$$

any limit point of optimal parameters a_n *is average cost optimal.*

Continuous time extensions and remarks

- Conditions should be adapted to (except for the componentwise continuity conditions)
  - There exist $F : \mathbf{S} \to [1, \infty)$, constant $\gamma > 0$, and finite set $M \subset \mathbf{S}$ s.t. ${}_M Q(a) F \leq -\gamma F$, $a \in$ A;
  - There exists an F-moment function W, and constant θ, such that $Q(a) W \leq \theta W$.

  Pbm *No simple relation between average reward for original PMP and transformed PMP by F.*

  Transformation only used for necessary continuity properties of $g(a), v(a)$.
- Competing queues model with abandonments is 'exponentially stable'. This often holds for abandonment models.
- Server farm model has transient policies $\Rightarrow$ Assumption 6 not applicable. However, maybe theory can be adapted using:
  - There exists $F : \mathbf{S} \to [0, \infty)$, $\inf_x F_x = 0$, and constant $\gamma > 0$, s.t. $Q(a) F \leq -\gamma F$, $a \in$ A; then $X(a)$ has a transient class of states (exponential transience)
- Qn: transformation argument as in discounted case possible? Via reformulation average reward pbm as total reward pbm?