Adaptive Greedy Rules Based on Dynamic Prices

Peter Jacko*

BEM

June 26, 2010

*Basque Center for Applied Mathematics, Spain

Neoclassical Economics

- Standard problem of neoclassical economics:
  - ▷ maximize aggregate utility w.r.t. budget constraint
- Standard assumptions (among others):
  - ▷ goods/services are continuously-divisible
  - ▷ budget (money) is continuously-divisible
  - ▷ goods/services do not change over time
- Standard solution:
  - ▷ marginal utility per unit of money spent must be equal for each good

Motivation

- Resource allocation when the assumptions do not hold:
  - ▷ telecommunications (routing, congestion control)
  - ▷ robotics (tracking, ranking)
  - ▷ marketing (assortment, pricing)
  - ▷ labor economics (job search)
  - ▷ clinical trials (treatment selection)
- Can we still apply marginalism ideas?
- What policies are optimal?
- What policies are simple to implement?

Talk Outline

- Resource allocation problem
- Framework
- Approach and adaptive greedy rules
- Known results
- Challenges

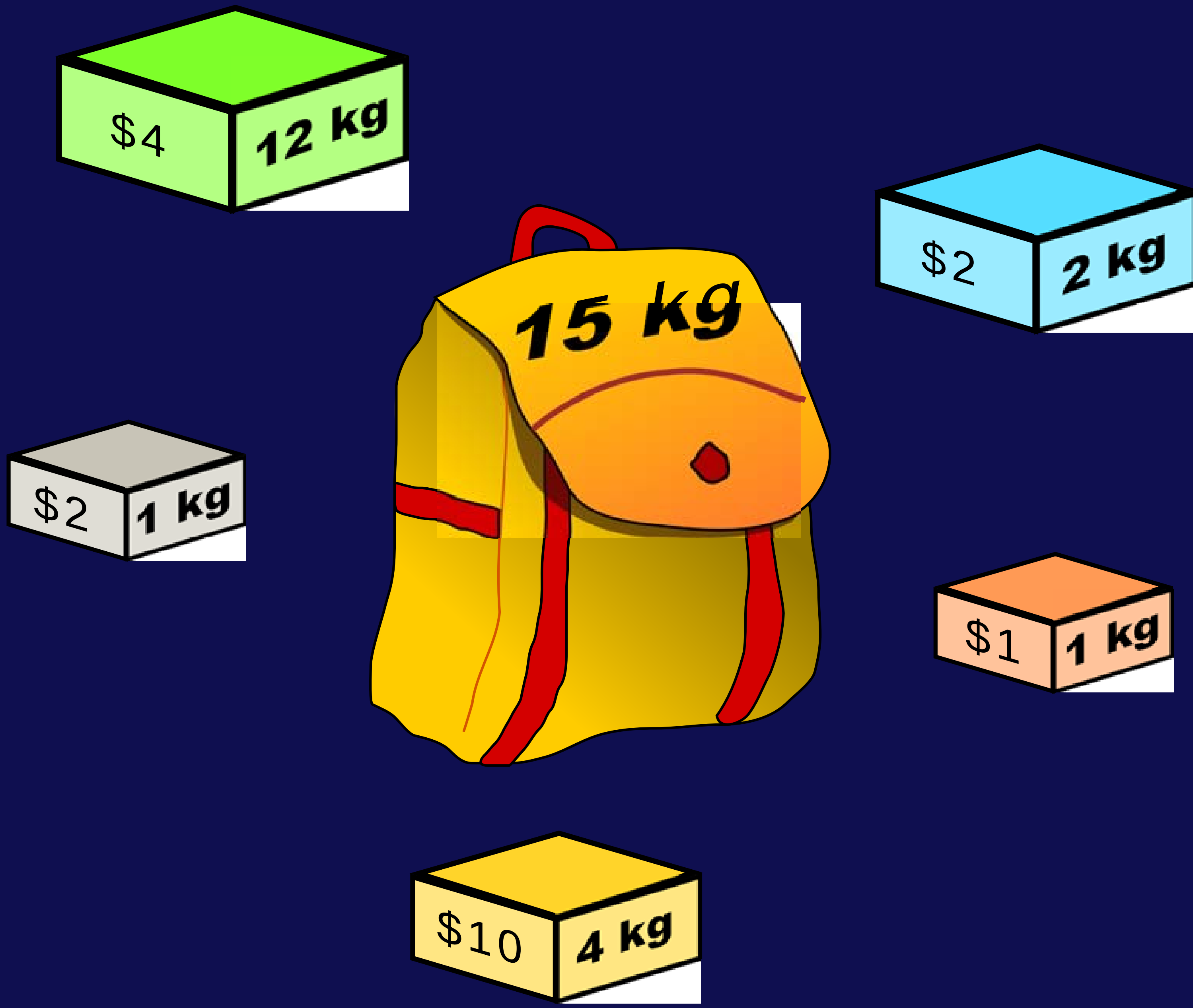
Resource Allocation Problem (RAP)

- Stochastic and dynamic
- There are a number of independent competitors
- Constraint: resource capacity at every moment
- Objective: maximize expected "reward"
- Captures the exploitation vs. exploration trade-off
  - ▷ always exploiting (being myopic) is not optimal
  - ▷ always exploring (being utopic) is not optimal
- This is a model of learning by doing!

Questions to Answer

- [Economic] For a given joint goal, is it possible to define dynamic quantities for each competitor that can be interpreted as prices? And if yes,
- [Algorithmic] How to calculate such prices quickly?
- [Mathematical] Under what conditions is there a greedy rule that achieves optimal resource capacity allocation?
- [Experimental] If greedy rules are not optimal, how close to optimality do they come? And how do they compare to alternative rules?

Static RAP: Knapsack Problem



Stochastic Programming Framework

- Stochastic programming = Markov Decision Processes
- Discrete time model $(t = 0, 1, 2, \ldots)$
- Competitor $k \in \mathcal{K}$ is defined by
  - ▷ state space $\mathcal{N}_k$, action space $\mathcal{A}$
  - ▷ expected one-period capacity consumption $\boldsymbol{W}_k^a$
  - ▷ expected one-period reward $\boldsymbol{R}_k^a$
  - ▷ one-period transition probability matrix $\boldsymbol{P}_k^a$
- State process $X_k(t) \in \mathcal{N}_k$
- Action process $a_k(t) \in \mathcal{A}$ – to be decided

Example: Job Sequencing Problem

- Find a serving sequence minimizing the total cost of waiting of jobs $k \in \mathcal{K}$
  - ▷ c_k = cost of waiting for jobs k
  - ▷ μ_k = service rate for jobs k
- $\mathcal{N}_k := \{\text{'completed'}, \text{'waiting'}\}$, $\mathcal{A}_k := \{\text{'serve'}, \text{'wait'}\}$
- expected one-period capacity consumption

$$W_{k,\text{'completed'}}^{\text{'serve'}} := 1, \qquad W_{k,\text{'waiting'}}^{\text{'serve'}} := 1,$$
$$W_{k,\text{'completed'}}^{\text{'wait'}} := 0, \qquad W_{k,\text{'waiting'}}^{\text{'wait'}} := 0;$$

Example: Job Sequencing Problem

- expected one-period reward

$$R_{k,\text{`completed'}}^{\text{`serve'}} := 0, \qquad R_{k,\text{`waiting'}}^{\text{`serve'}} := -c_k(1-\mu_k),$$
$$R_{k,\text{`completed'}}^{\text{`wait'}} := 0, \qquad R_{k,\text{`waiting'}}^{\text{`wait'}} := -c_k;$$

- one-period transition probability matrices

$$\boldsymbol{P}_k^{\text{`serve'}} := \begin{array}{r} \\ \text{`completed'} \\ \text{`waiting'} \end{array} \begin{array}{c} \begin{array}{cc} \text{`completed'} & \text{`waiting'} \end{array} \\ \begin{pmatrix} 1 & 0 \\ \mu_k & 1-\mu_k \end{pmatrix} \end{array},$$

$$\boldsymbol{P}_k^{\text{`wait'}} := \begin{array}{r} \\ \text{`completed'} \\ \text{`waiting'} \end{array} \begin{array}{c} \begin{array}{cc} \text{`completed'} & \text{`waiting'} \end{array} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{array}.$$

Resource Allocation Problem

- Formulation under the β-discounted criterion:

$$\max_{\pi \in \Pi} \sum_{k \in \mathcal{K}} \mathbb{E}^{\pi} \left[\sum_{t=0}^{\infty} \beta^t R_{k, X_k(t)}^{a_k(t)} \right]$$

subject to $\displaystyle\sum_{k \in \mathcal{K}} W_{k, X_k(t)}^{a_k(t)} \le W,$ for all $t = 0, 1, 2, \ldots$

- This problem is PSPACE-hard
  - ▷ intractable to solve exactly by Dynamic Programming
  - ▷ instead, we relax and decompose the problem

Whittle's Relaxation

- Fill the capacity in expectation
  - ▷ infinite number of constraints is replaced by one
  - ▷ sort of perfect market assumption

$$\max_{\pi\in\Pi} \sum_{k\in\mathcal{K}} \mathbb{E}^{\pi}\left[\sum_{t=0}^{\infty} \beta^t R_{k,X_k(t)}^{a_k(t)}\right]$$

$$\text{subject to} \quad \sum_{k\in\mathcal{K}} \mathbb{E}^{\pi}\left[\sum_{t=0}^{\infty} \beta^t W_{k,X_k(t)}^{a_k(t)}\right] \leq \sum_{t=0}^{\infty} \beta^t W$$

- Provides an upper bound for RAP

Lagrangian Relaxation

- Pay cost λ for using the capacity
  - ▷ the constraint is moved into the objective

$$\max_{\pi \in \Pi} \sum_{k \in \mathcal{K}} \mathbb{E}^{\pi} \left[\sum_{t=0}^{\infty} \beta^t R_{k, X_k(t)}^{a_k(t)} \right] - \lambda \sum_{k \in \mathcal{K}} \mathbb{E}^{\pi} \left[\sum_{t=0}^{\infty} \beta^t W_{k, X_k(t)}^{a_k(t)} \right]$$

- Also provides an upper bound for RAP
- Decomposes due to competitor's independence into single-competitor parametric subproblems
  - ▷ solved by identifying the efficiency frontier

Dynamic Prices

- We will assign each competitor a dynamic price
- They arise in the solution of the parametric subproblem
  - ▷ optimal policy: use capacity iff price lower than λ
- Prices are values of λ when optimal solution changes
- They define an indifference curve
- However, such prices may not exist!
- Price computation:
  - ▷ in general, by parametric simplex method
  - ▷ after math, sometimes obtained in a closed form

Adaptive Greedy Rules

- We are concerned with the following rule
  - ▷ at each moment be greedy:
    prefer competitors with higher current prices
- It is adaptive because the prices are dynamic
- Experiments and simulations suggest that it gives a nearly-optimal solution to RAP
- In some simple problems, it is optimal

Optimality of Adaptive Greedy Rules

- Often in problems with symmetric competitors
- E.g.: in routing to parallel queues, route to:
  - ▷ the Shortest Queue (if min. delays)
- E.g.: sequencing of customers to service:
  - ▷ the Shortest Service Time (if min. waiting time)
  - ▷ the Least Empty Buffer Space (if min. losses)
  - ▷ the Shortest Queue (if min. delays)
- These values are the "dynamic prices"

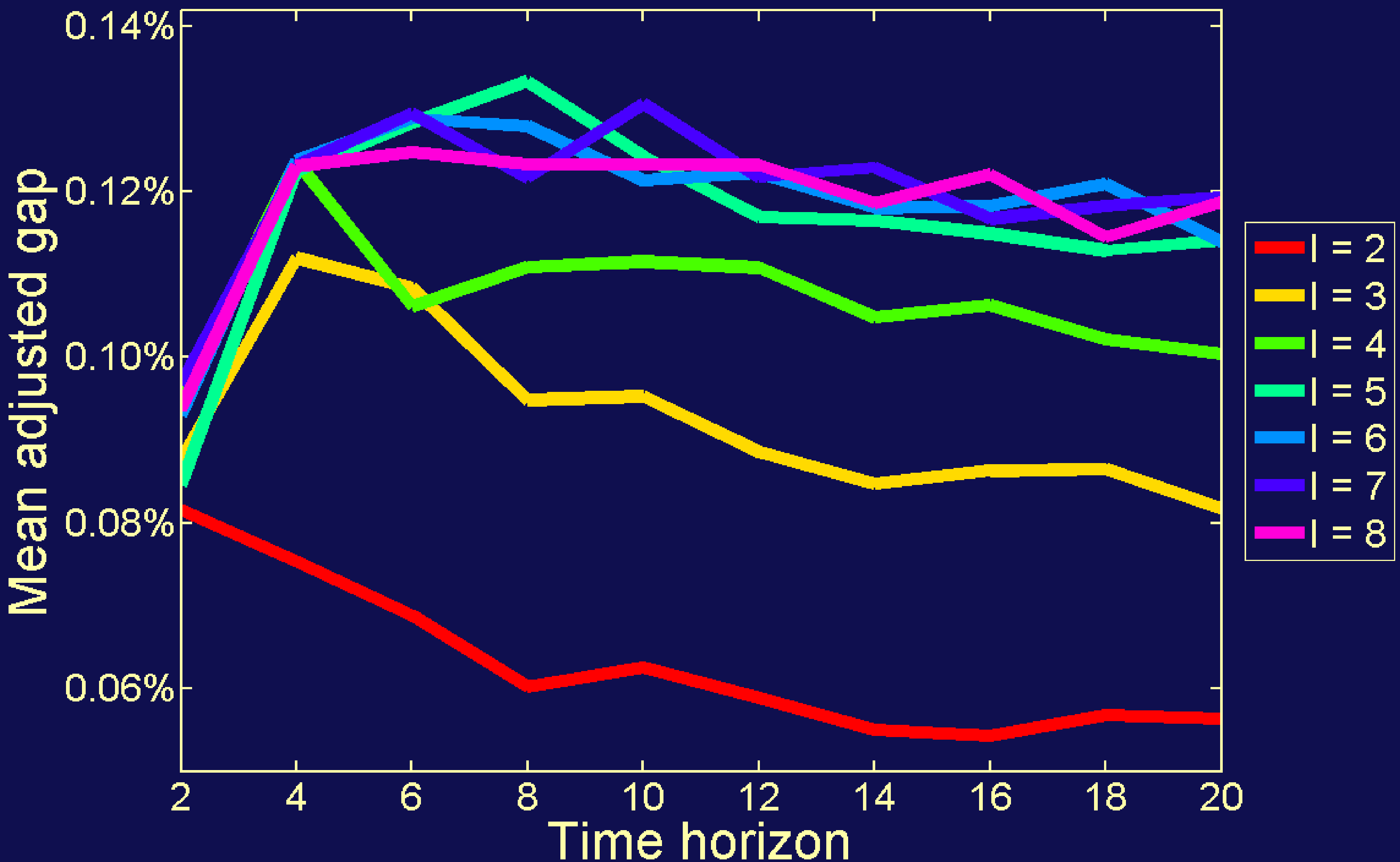
Optimality of Adaptive Greedy Rules

- Also in asymmetric problems with simple dynamics
- $c\mu$-rule (Cox & Smith '61) for job sequencing
  - ▷ assigning priority to the job k with largest $c_k\mu_k$
- Gittins index rule ('72) for multi-armed bandit problem
  - ▷ big surprise: many people believed it was insolvable
- Klimov index rule ('74) for $M/G/1$ model with feedback

Priority Rules in More Complex Problems

- Telecommunications
  - ▷ several variants of routing to parallel queues
  - ▷ congestion control for Internet connections
  - ▷ time sharing in 3G wireless systems
- Marketing
  - ▷ dynamic assortment for “fast-fashion” companies
  - ▷ dynamic promotion for supermarkets

Example: Performance of a Priority Rule



Challenges

- Modeling
  - ▷ ...if modeling were as easy as mathematics...
- Proving (near-)optimality of greedy rules
  - ▷ asymptotic optimality proved for symmetric case, as number of competitors and resource capacity grow
- Incorporation of risk aversion, etc.

Thank you for your attention

Resource Allocation in Telecommunications

- Congestion control
  - ▷ maximize throughput (choose preferred flows)
- Routing
  - ▷ minimize packet losses (choose preferred paths)
- Admission control
  - ▷ minimize delays (choose preferred packets)
- Fairness
  - ▷ maximize users' utilities (choose preferred users)

Resource Allocation in Robotics

- Arises within reinforcement learning
- Behavior coordination & navigation in animats
  - ▷ maximize utility (choose preferred behavior)
- Ranking in web search robots
  - ▷ minimize searching time (choose preferred document)
- Multi-target tracking & environment mapping
  - ▷ maximize map correctness (choose preferred object)

Problems in Telecommunications

- Fall into resource allocation problems
- Advantages:
  - ▷ decentralized control
  - ▷ natural for creating priority tables
  - ▷ dynamic prices yield structural results
  - ▷ nearly-optimal (optimal in expectation)
- Disadvantages
  - ▷ prices may not exist

Problems in Telecommunications

- However, dynamic prices are scarcely used
  - ▷ Niño-Mora '02, '06
  - ▷ Raissi-Dehkordi & Baras '02
  - ▷ Goyal et al. '06
  - ▷ Jacko '09
- Optimality of ad-hoc priority rules is usually analyzed
  - ▷ Glazebrook et al. '04, '04, '07
  - ▷ Ehsan & Liu '04, '05, '06, '07

Problems in Telecommunications

- Niño-Mora '02:
  - ▷ queues with finite buffers
  - ▷ consider a rejection cost as the wage
  - ▷ assume concave nondecreasing service rates
  - ▷ assume convex nondecreasing holding costs
  - ▷ price-based characterization of optimal threshold policy
  - ▷ as rejection cost grows, start rejecting customers under longer queue
  - ▷ priority rule heuristic for routing to parallel queues

Problems in Telecommunications

- Niño-Mora '06:
  - ▷ queues with finite buffers
  - ▷ analyzes loss-sensitive and delay-sensitive queues
  - ▷ rejecting cost vs. discounted holding-forever cost
  - ▷ loss-sensitive: fewest-empty-buffer-spaces rule
  - ▷ delay-sensitive: shorter-queue rule
  - ▷ both converge to $c\mu$ rule for infinite buffers
  - ▷ throughput maximization is special case

Problems in Telecommunications

- Goyal, Kumar & Sharma ’06:
  - ▷ transmissions over polled multiaccess fading channel
  - ▷ voice $\succ$ streaming media $\succ$ files
  - ▷ infinite buffers, delayed information
  - ▷ poll-and-response system
- Raissi-Dehkordi & Baras ’02:
  - ▷ pulling broadcast scheduling (teletext with feedback)
  - ▷ minimize weighted average waiting time

Problems in Telecommunications

- Glazebrook et al. '04:
  - ▷ server allocation to impatient (perishable) tasks
  - ▷ reduces to Gittins indices in a special case
- Glazebrook & Kirkbride '04:
  - ▷ routing of background jobs in distributed PC systems
  - ▷ ad-hoc prices (static policy improvement)
- Glazebrook & Kirkbride '07:
  - ▷ routing to heterogeneous unreliable servers
  - ▷ ad-hoc prices (DP policy improvement)

Problems in Telecommunications

- Ehsan & Liu '04, '05, '06, '07:
  - ▷ wireless server allocation with delays
  - ▷ minimize expected holding costs
  - ▷ ad-hoc prices (myopic)
  - ▷ give sufficient optimality conditions (special cases)