Rolling Horizon Procedures in Nonhomogeneous Markov Decision Processes

Jeffrey M. Alden
General Motors Research Laboratories
Warren, MI 48090

Robert L. Smith
Department of Industrial and Operations Engineering
The University of Michigan
Ann Arbor, MI 48109

Rolling Horizon Procedures in Nonhomogeneous Markov Decision Processes

Jeffrey M. Alden

General Motors Research Laboratories

Warren, MI 48090

Robert L. Smith

Department of Industrial and Operations Engineering

The University of Michigan

Ann Arbor, MI 48109

August 1987

Abstract

By far the most common planning procedure found in practice is to approximate the solution to an infinite horizon problem by a series of rolling finite horizon solutions. Although a great number of empirical studies have been done, the so-called rolling horizon procedure has heretofore been the subject of few analytical studies. We provide a cost error bound for a general rolling horizon algorithm when applied to infinite horizon non-homogeneous Markov Decision Processes, both in the discounted and average cost cases. We show that a Doeblin coefficient of ergodicity acts much like a discount factor to reduce this error. In particular, we show that the error goes to zero for any *fixed* rolling horizon as this Doeblin measure of uncertainty about the future increases. The theory is illustrated through an application to vehicle deployment.

The classical approach for solving homogeneous Markov Decision Processes over an infinite horizon is to employ either policy or value iteration (Denardo [1982]) to solve the corresponding dynamic programming functional equation. Policy iteration (a variant of the simplex method) solves the functional equation in a finite number of steps while value iteration involves a series of finite horizon approximations. The non-homogeneous case, while not even finitely expressable, can also be solved by a form of value iteration that recursively recovers a sequence of optimal policies by solving a sequence of successively longer horizon problems (Hopp, Bean, and Smith [1984], Bhaskaran and Sethi [1985]). The sequence of finite horizons employed to recover a sequence of optimal policies are however *variable* and difficult to discover (Bean and Smith [1986]). In practice, planners employ a *fixed* horizon procedure to iteratively recover a sequence of policies. The procedure is to fix a horizon T, solve the corresponding T period problem, implement the initial policy found, roll forward one period and repeat from the new current state. This *rolling horizon procedure* is by far the most common method employed in practice for generating and implementing solutions to non-homogeneous problems (see for example Baker [1977]).

Because the horizon employed by a rolling horizon procedure is fixed, the sequence of policies (or strategy) generated may fail to be optimal. Bean, Birge and Smith [1987] have determined the error in cost obtained as a function of horizon employed for a variant of the standard rolling horizon procedure. They require a full roll of T periods so that the entire subsequence of policies identified in solving the first T period problem must be committed to. In related work, Lee and Denardo [1983] provide for the dynamic lot size problem a bound on the error from infinite horizon optimal under the assumption that one continues optimally from T onwards.

In this paper, we bound the error from infinite horizon optimal as a function of horizon employed for the standard rolling horizon procedure. The analysis, provided in Section 1, is complicated in this standard case by the policy interactions induced by overlapping subproblems. In Section 2, we reduce the bound on error found in Section 1 by exploiting the mitigating effects of uncertainty as measured by a coefficient of ergodicity we call the decoupling coefficient. We show that, as this measure of uncertainty about the future grows, the error between the infinite horizon optimal value and that attained by a rolling horizon strategy goes to zero *regardless of horizon used*. This result provides rigorous support for the commonly held view that uncertainty about the future justifies short term planning. In Section 3, we consider an application to vehicle deployment.

1 The Rolling Horizon Procedure and Its Associated Error

We begin with a formulation of the problem to be solved.

1.1 The Markov Decision Process

Consider a system which, during all time intervals or periods, can be described as being in exactly one of a finite set of states $S = \{1, 2, \cdots, n\}$. Corresponding to each state $i \in S$ and period $k = 1, 2, 3, \cdots$ is a finite set of feasible decisions, D^k. If the system is currently in state i during period k and decision $x_i^k \in D_i^k$ is chosen, then the system makes a transition into state j during period $k + 1$ with probability $P_{ij}^k(x_i^k)$. In general, the cost of selecting the decision and making the transition may depend on the state we transition into as well as the decision. However, by calculating the expected cost of making a transition over all possible such states and using the additivity of expectation, we may, without loss of generality, assume that the expected cost $C_i^k(x_i^k)$ depends only on the current state i, period k, and decision taken x_i^k. We assume that all costs are non-negative and uniformly bounded from above, i.e., for all i and k

$$0 \le C_i^k(x_i^k) \le \bar{C} < \infty \quad \text{for all} \quad x_i^k \in D_i^k.$$

Costs incurred during period k are discounted by a factor α^{k-1} where $0 < \alpha < 1$.

The column vector $x^k = (x_1^k, x_2^k, x_3^k, \cdots, x_n^k)^t$ that associates feasible decisions to all possible states in period k is called a feasible *policy* for period k. A sequence of feasible policies, $x = (x^1, x^2, x^3, \cdots)$ is called a feasible *strategy*. A *Markov Decision Process* (MDP) is the problem of determining a feasible strategy that minimizes discounted expected costs over the infinite horizon.

To express this problem mathematically, we adopt the following additional notation where $1 \le l < k$ throughout. $x^{lk} = (x^l, x^{l+1}, \cdots, x^k) \in D^{lk} \equiv \prod_{i=l}^k D^i$ is the set of feasible policies for period i where $D^i = (D_1^i, D_2^i, D_3^i, \cdots, D_n^i)^t$. $D = \prod_{k=1}^\infty D^k$ is the space of feasible strategies which we assume to be closed (Bean and Smith[1984]). $P^k(x^k) = (P_{ij}^k(x_i^k))$ is the kth period one step transition probability matrix under policy x^k. $P^{lk}(x^{lk}) = \prod_{j=l}^k P^j(x^j)$ is the $(k-l+1)$-step transition probability matrix from periods l through k under policies x^l through x^k where $P^{l,l-1} \equiv I$. $C^k(x^k) = (C_1^k(x_1^k), C_2^k(x_2^k), \cdots, C_n^k(x_n^k))^t$ is the column vector of expected costs under policy x^k during period k. $C^k(x^k)$, as with other expected costs considered

2

later, is a column vector whose i^{th} component represents a conditional expected cost given we enter state i at the beginning of period k.

The MDP is then

$$(P) \qquad \min_{x \in D} C(x)$$

where $C(x) = \sum_{k=1}^{\infty} \alpha^{k-1} P^{1,k-1}(x^{1,k-1}) C^k(x^k)$. Because all costs are uniformly bounded, the discount factor is less than one, and the set of feasible strategies is compact, the minimum above is attained (Bean and Smith[1984]).

Let x^* be an optimal strategy and C^* its corresponding optimal cost. Note that x^* incorporates an optimal policy $(x^*)^1$ for the first period that provides an optimal decision for all possible initial states. Further, the fact that subsequent policies can only depend on the state entered during that period and not on previously visited states involves no loss of optimality by the principle of optimality.

1.2 The Rolling Horizon Algorithm

A rolling horizon procedure generates a (suboptimal) strategy for (P) by solving a sequence of subproblems, each with a fixed finite horizon. The initial periods of these subproblems start at one and increment by one for each subproblem in the sequence. The optimal initial policies of these subproblems form a suboptimal strategy for the original infinite horizon problem (P).

Let

$$x^*(l,k) = \arg \min_{x^{lk} \in D^{lk}} C^{lk}(x^{lk})$$

where $C^{lk}(x^{lk}) = \sum_{j=l}^{k} \alpha^{j-1} P^{l,j-1}(x^{l,j-1}) C^j(x^j)$, so that $x^*(l,k)$ is an optimal substrategy for periods l through k. Let $C^*(l,k) = C^{lk}(x^*(l,k))$ be the optimal expected cost for the subproblem spanning periods l through k.

Rolling Horizon Algorithm

1. Set $l = 1$ and $k = l + T - 1$.

2. Solve for $x^*(l,k)$ and set $\hat{x}^l = (x^*(l,k))^1$.

3. Let $l = l + 1$ and go to 2.

The algorithm above recursively generates the suboptimal infinite horizon strategy

$$\hat{x} = (\hat{x}^1, \hat{x}^2, \hat{x}^3, \cdots).$$

The strategy $\hat{x}$ is called a *rolling horizon strategy*. Let $\hat{C} = C(\hat{x})$ be the expected cost vector of $\hat{x}$.

In an implementation of the rolling horizon algorithm, the random state I_l occurring in period l from implementation of the previous $l - 1$ policies would lead to implementation of decision $\hat{x}^l_{I_l}$ during period l after iteration l.

Because the determination of $\hat{x}^l$ is, for all l, based upon a truncated version of the original infinite horizon problem, the resulting strategy $\hat{x}$ is myopic and may not be optimal. Consider, for example, the following deterministic special case of (P), represented in Figure 1 as a decision network. The numbers above all of the decision arcs correspond to the discounted cost of making that decision all of which have a duration of one period. The problem is to make an infinite sequence of decisions at minimum total cost. It is clear that the minimum cost strategy x^* is to move downward and then across resulting in an optimal cost of $1 + \frac{1}{2} + \frac{1}{4} + \cdots = 2$. However, the rolling horizon strategy $\hat{x}$ generated by a horizon of T equal to 2 periods
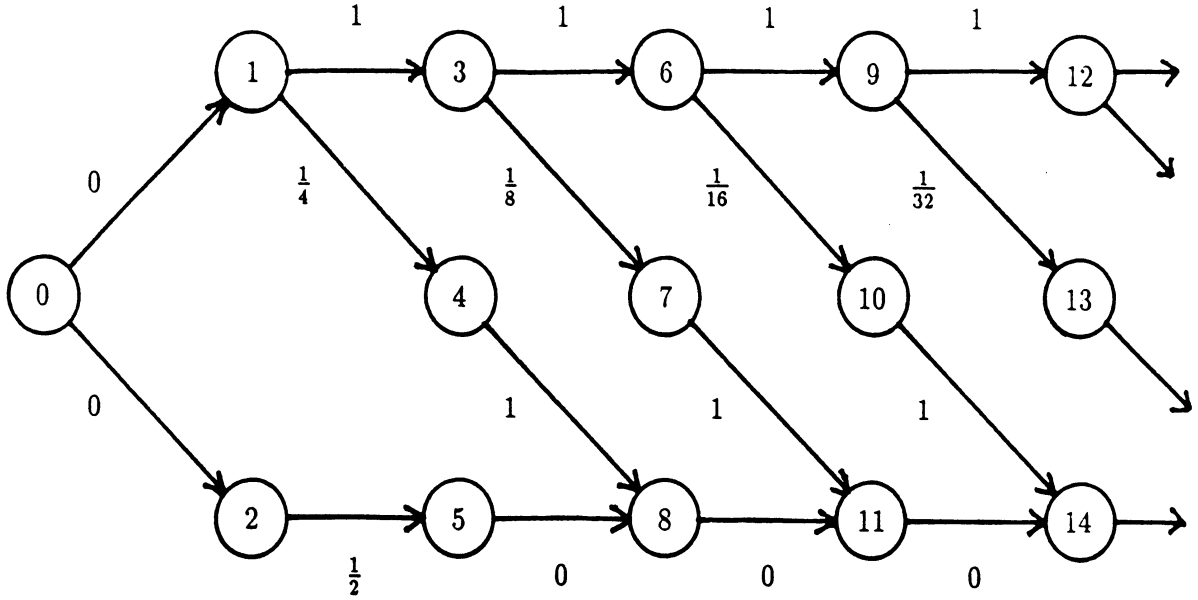
3

Figure 1: An example of problem (P) where a rolling horizon procedure (using a horizon of 2 periods) generates a poor solution.

when applied to the initial state node 0 is to move up and then across for a total cost of $1 + 1 + 1 + \cdots$ or infinity.

The rolling horizon strategy is in fact the only infinite cost strategy! It is interesting that a full roll solution achieves a nearly optimal finite cost even though most decisions are made with less information.

As this example illustrates, the departure of a rolling horizon solution from optimal can be dramatic. We turn, in the next section, to an analysis of the magnitude of error involved for less pathological instances of (P).

1.3 An Error Bound for Rolling Horizon Procedures

In this section, we derive a tight upper bound on the difference in expected cost between a rolling horizon strategy $\hat{x}$ and an optimal strategy x^*. We begin with some definitions.

Let $C_l^* = C^*(l, l + T - 1)$ be the optimal expected cost of the lth subproblem and $z(l)$ be any strategy such that $z^{l,l+T-1}(l) = x^*(l, l + T - 1)$ is its optimal substrategy. Let $\hat{C}(l, k) = C^{lk}(\hat{x}^{lk})$ be the expected cost of the rolling horizon strategy spanning periods l through k.

We turn now to introducing the important notion of an extension cost. Note that the optimal substrategy in $z(l)$ obtained from solving the lth subproblem of a rolling horizon algorithm can always be extended to form a feasible substrategy for the $l + 1^{st}$ subproblem. In particular, any strategy $\xi(l)$ with a substrategy of the form $\xi^{l+1,l+T}(l) = (z^{l+1}(l), z^{l+2}(l), \cdots, z^{l+T-1}(l), x^{l+T})$ for $x^{l+T} \in D^{l+T}$ (formed by dropping the initial policy of $z^{l,l+T-1}(l)$ and adding any feasible policy for period $l + T$) is feasible for the $l + 1^{st}$ subproblem.

Let

$$e_l = \min_{x^{l+T} \in D^{l+T}} \alpha^{l+T-1} P^{l,l+T-1}(z^{l+1,l+T-1}(l)) C^{l+T}(x^{l+T})$$

be the expected cost vector whose i^{th} component is the minimum conditional expected cost of extending substrategy $z(l)$ given the state entered beginning period l. This is called the lth extension cost.

4

We now establish in the following lemma an inequality that bounds from above the expected cost of a rolling horizon strategy in terms of its extension costs.

Lemma 1.1 $\hat{C} \leq C_1^* + \sum_{l=1}^{\infty} E(\underline{e}_l)$ *where* $\underline{e}_l$ *is a random vector under strategy* $\hat{x}$ *whose* i^{th} *component,* $i = 1, 2, \cdots, n,$ *takes on the value* $(e_l)_{I_l}$ *when the initial state of* (P) *beginning period 1 is* i *and* I_l *is the state entered beginning period* l.

Proof: See Appendix. ∎

Using Lemma 1.1, we now establish an upper bound on the cost error between $\hat{x}$ and x^*.

Theorem 1.2 *Let* $\varepsilon = \sum_{l=1}^{\infty} \alpha^{l-1} \bar{C}^l e$ *where* $\bar{C}^l = \max_{i=1,2,\cdots,n} \min_{x_i^l \in D^l} C^l(x_i^l)$ *and* $e = (1, 1, \cdots, 1)^t$. *The error in cost between a rolling horizon strategy* $\hat{x}$ *and an optimal strategy* x^* *for problem* (P) *is uniformly bounded from above by* $\alpha^T \varepsilon$, *i.e.,*

$$\hat{C} - C^* \leq \alpha^T \varepsilon.$$

In particular,

$$\hat{C} - C^* \leq \alpha^T \frac{\bar{C}e}{1-\alpha}.$$

Proof: Since y^{l+T} is feasible,

$$e_l = \min_{x^{l+T} \in D^{l+T}} \alpha^{l+T-1} P^{l,l+T-1}(z^{l+1,l+T-1}(l)) C^{l+T}(x^{l+T})$$

$$\leq \alpha^{l+T-1} P^{l,l+T-1}(z^{l+1,l+T-1}(l)) C^{l+T}(y^{l+T})$$

where $y_i^{l+T} = \arg\min_{x_i^{l+T} \in D_i^{l+T}} C^{l+T}(x_i^{l+T})$ for $i = 1, 2, \cdots, n$. Since $C^{l+T}(y^{l+T}) \leq \bar{C}^l$,

$$e_l \leq \alpha^{l+T-1} P^{l,l+T-1}(z^{l+1,l+T-1}(l)) \bar{C}^l e$$

$$\leq \alpha^{l+T-1} \bar{C}^l e.$$

Hence $E(\underline{e}_l) \leq \alpha^{l+T-1} \bar{C}^l e$. Also $C_1^* \leq C^*(1, T) \leq C^*$ since $C_i^k(x_i^k) \geq 0$ for all i and k by the assumption of Section 1.1. Therefore

$$\hat{C} \leq C^* + \alpha^T \sum_{l=1}^{\infty} \alpha^{l-1} \bar{C}^l e = C^* + \alpha^T \varepsilon.$$

Moreover,

$$\varepsilon \leq \sum_{l=1}^{\infty} \alpha^{l-1} \bar{C} e = \frac{\bar{C}e}{1-\alpha},$$

since $C_i^l \leq \bar{C}$ for all i and l by the assumption of Section 1.1 ∎

It is interesting to note that the error of the second bound of the theorem is the present worth of an annuity paid every period after period T while the Bean, Birge, and Smith model incorporates a smaller error corresponding to an annuity paid once every T periods. Although the full roll error bound is smaller than that for decision rolls, the former is seldom used in practice. One can speculate that the reason is lack of accurate forecasting of the parameters of the MDP, encouraging updating every period.

5

2 An Improved Error Bound Exploiting Stochastic Effects

The error bound derived in the last section relies exclusively on discounting to bound the error involved in a rolling horizon procedure. In fact, for fixed horizon T, as the discount factor α goes to 1, the error bound goes to infinity. In this section, we improve this bound by exploiting the effect of uncertainty about the future on the error involved in failing to consider immediate effects beyond the horizon T. The procedure is to formulate an equivalent MDP whose discount factor is scaled by a factor β that measures the sensitivity of future state distributions to current decisions. As β goes to 0, a myopic strategy becomes closer to optimal. The error involved is quantitatively assessed by applying the error bound of Theorem 1 to this equivalent MDP. It is shown that, for fixed horizon T and fixed discount factor α, the bound on error goes to zero as $\beta \rightarrow 0$.

2.1 An Equivalent Markov Decision Process

We say two Markov Decision Processes (P) and $(\tilde{P})$ are *equivalent* if the sets of optimal substrategies $x^*(l, k)$ and $\tilde{x}^*(l, k)$ are equal for all l, k. Whenever (P) and $(\tilde{P})$ are equivalent, it is easy to see that the rolling horizon strategies $\hat{x}$ and $\hat{\tilde{x}}$ will be the same provided that ties for optimal substrategies are broken according to the same rules. Before we define an MDP $(\tilde{P})$ equivalent to our problem (P) of Section 1.1, we need to introduce the notion of a coefficient of ergodicity.

Lemma 2.1 *Every one step transition probability matrix in* (P) *can be expressed as a convex combination of a stochastic matrix and a stable matrix (i.e., a stochastic matrix of equal rows) dependent only on period where the multiplier of the former, β is given by*

$$\beta = \max_k (1 - \min_{x^k \in D^k} \sum_j \min_i P_{ij}^k(x^k)).$$

That is, for all k and all $x^k \in D^k$,

$$P^k(x^k) = \beta \tilde{P}^k(x^k) + (1 - \beta)\tilde{L}^k$$

where $\tilde{P}^k(x^k)$ is a stochastic matrix, $\tilde{L}^k$ is a stable stochastic matrix independent of x^k, and $0 \le \beta \le 1$ is given as above.

Proof: See Appendix. ∎

The multiplier, β is the largest Doeblin coefficient of all the one-step transition matrices (Seneta[1981]) and is a measure of ergodicity. That is, the closer β is to 0, the closer the $P^k(x^k)$ are to a stable matrix. In the proof of the lemma above, β_k is a coefficient of ergodicity for the transition matricies of the kth period alone, where $\beta = \max_k \beta_k$. In the extreme case where $\beta_k = 0$, $\{P^k(x^k), x^k \in D^k\}$ for that k are all stable. The MDP is then effectively decoupled at the end of period k; that is, the optimal substrategy for a horizon equal to k is an optimal substrategy for the infinite horizon problem. On the other hand, if β_k is one, then every state in period $k + 1$ is impossible to attain from some state and decision pair in period k. For these reasons, we refer to β_k or more generally β as *coupling coefficients*.

We are now in a position to define an equivalent MDP $(\tilde{P})$. $(\tilde{P})$ is formed from (P) by replacing $P^k(x^k)$ by $\tilde{P}^k(x^k)$ for $x^k \in D^k$ and α by $\alpha\beta$ respectively. Specifically, $(\tilde{P})$ is the MDP

$$(\tilde{P}) \qquad \min_{x \in D} \tilde{C}(x)$$

where $\tilde{C}(x) = \sum_{k=1}^{\infty} (\alpha\beta)^{k-1} \tilde{P}^{1,k-1}(x^{1,k-1}) C^k(x^k)$ with $\tilde{P}^{lk}(x^{lk}) = \prod_{j=l}^{k} \tilde{P}^j(x^j)$ for $1 \le l \le k$ and $\tilde{P}^{l,l-1}(x^{l,l-1}) = I$ for $x^{lk} \in D^{lk}$.

Lemma 2.2 (P) *and* $(\tilde{P})$ *are equivalent Markov Decision Processes.*

Proof: See Appendix. ∎

We now use the equivalent MDP $(\tilde{P})$ to establish a tighter bound on the error associated with the rolling horizon procedure when applied to (P).

2.2 An Improved Error Bound for the Rolling Horizon Procedure

Although the MDP's (P) and $(\tilde{P})$ were demonstrated to be equivalent in optimal strategies, their optimal costs are not equal. In order to relate cost error bounds between them, we establish the relationship between the expected costs generated by a rolling horizon procedure when applied to the two formulations.

Let

$$\hat{\tilde{C}}(l) = \sum_{i=l}^{\infty} (\alpha\beta)^{i-1} \tilde{P}^{l,i-1}(\hat{x}^{l,i-1}) C^i(\hat{x}^i)$$

be the expected cost for problem $(\tilde{P})$ of the rolling horizon strategy $\hat{x}$ from period l onwards, and

$$\tilde{C}^*(l) = \sum_{i=l}^{\infty} (\alpha\beta)^{i-1} \tilde{P}^{l,i-1}((x^*)^{l,i-1}) C^i((x^*)^i)$$

be the expected cost for problem $(\tilde{P})$ of the optimal strategy x^* from period l onwards.

Lemma 2.3 *The error in cost between a rolling horizon strategy $\hat{x}$ and an optimal strategy x^* for problem (P) is given in terms of problem $(\tilde{P})$ by*

$$\hat{C} - C^* = \hat{\tilde{C}} - \tilde{C}^* + (1-\beta) \sum_{l=1}^{\infty} \beta^{-l} \tilde{L}^l \left(\hat{\tilde{C}}(l+1) - \tilde{C}^*(l+1) \right).$$

Proof: See Appendix. ∎

Note that $\tilde{C}^*(l+1)$ in the above lemma also represents the optimal cost vector of beginning period $l+1$ in all possible initial states since $((x^*)^{l+1}, (x^*)^{l+2}, \cdots)$ must, by the principle of optimality, be optimal for that problem. That is,

$$((x^*)^{l+1}, (x^*)^{l+2}, \cdots) = \arg\min \{ \sum_{i=l+1}^{\infty} (\alpha\beta)^{i-1} \tilde{P}^{l,i-1}(x^{l,i-1}) C^i(x^i), x^i \in D^i \text{ for } i = l+1, l+2, \ldots \}.$$

Also the rolling horizon algorithm if started at the beginning of period $l+1$ would generate the strategy $(\hat{x}^{l+1}, \hat{x}^{l+2}, \cdots)$. Therefore, we may apply Theorem 1.2 to bound the terms $\tilde{C}(l+1) - \tilde{C}^*(l+1)$ in Lemma 2.3 to obtain a bound on the rolling horizon error $\hat{C} - C^*$.

Theorem 2.4 *Let*

$$\varepsilon(\beta) = \sum_{l=1}^{\infty} (\alpha\beta)^{l-1} \bar{C}^l e$$

where

$$\bar{C}^l = \max_{i=1,\cdots,n} \min_{x_i^l \in D_i^l} C^l(x_i^l)$$

7

and $e = (1, 1, \cdots, 1)^t$. The error in cost between a rolling horizon strategy $\hat{x}$ and an optimal strategy x^* for problem (P) is bounded from above by $((\alpha\beta)^T(1 - \alpha\beta)/(1 - \alpha))\varepsilon(\beta)$, i.e.,

$$\hat{C} - C^* \leq \frac{(\alpha\beta)^T(1 - \alpha\beta)}{1 - \alpha}\varepsilon(\beta).$$

In particular,

$$\hat{C} - C^* \leq \frac{(\alpha\beta)^T}{1 - \alpha}\bar{C}e.$$

Proof: From Theorem 1.2,

$$\tilde{C}(l + 1) - \tilde{C}^*(l + 1) \leq (\alpha\beta)^{T+l}\varepsilon(\beta)$$

where $\varepsilon(\beta) = \sum_{l=1}^{\infty}(\alpha\beta)^{l-1}\bar{C}^l e$ and $\bar{C}^l = \max_{i=1,\cdots,n} \min_{x_i^l \in D_i^l} C^l(x_i^l)$. Hence from Lemma 2.3,

$$
\begin{aligned}
\hat{C} - C^* &\leq (\alpha\beta)^T\varepsilon(\beta) + (1 - \beta)\sum_{l=1}^{\infty}\beta^{-l}\tilde{L}^l(\alpha\beta)^{T+l}\varepsilon(\beta) \\
&= (\alpha\beta)^T\varepsilon(\beta) + (1 - \beta)\sum_{l=1}^{\infty}\beta^{-l}(\alpha\beta)^{T+l}\varepsilon(\beta)
\end{aligned}
$$

since $\tilde{L}^l e = e$. Hence

$$
\begin{aligned}
\hat{C} - C^* &\leq (\alpha\beta)^T\varepsilon(\beta) + (1 - \beta)(\alpha\beta)^T(\alpha/(1 - \alpha))\varepsilon(\beta) \\
&= ((\alpha\beta)^T(1 - \alpha\beta)/(1 - \alpha))\varepsilon(\beta).
\end{aligned}
$$

In particular, since $\varepsilon(\beta) \leq (\bar{C}e)/(1 - \alpha\beta)$, we have

$$\hat{C} - C^* \leq \frac{(\alpha\beta)^T(1 - \alpha\beta)}{1 - \alpha}\frac{\bar{C}e}{1 - \alpha\beta} = \frac{(\alpha\beta)^T}{1 - \alpha}\bar{C}e. \quad\blacksquare$$

Note that the error bound provided by Theorem 2.4 is always better than that of Theorem 1.2 and that they are equal when β equals one. Unlike the bound of Theorem 1.2, the error bound here goes to zero for the horizon finite and fixed when β goes to zero. Also since

$$\hat{C} - C^* \leq (\alpha\beta)^T\varepsilon(\beta),$$

$\alpha' = \alpha\beta$ may be viewed as an effective discount rate that acts at least as strongly as conventional discounting. For example, an α of 0.87 corresponding to a real interest rate of 15% together with an equal β of 0.87 lead to an effective discount factor of 0.76, or an effective interest rate of 32%. The conclusion is that a real interest rate of 15% together with stochastic effects corresponding to a coupling coefficient of 0.87 reduces the rolling horizon error over the deterministic case by a factor at least as great as doubling the real interest rate produces.

Theorem 1.2 provides a bound on rolling horizon error as a function of the horizon used. By inverting this bound, we can obtain a formula for the horizon T_ϵ^* necessary to guarantee a rolling horizon solution within a predetermined error ϵ of optimal. In particular,

$$T_\epsilon^* = \frac{\ln((1 - \alpha)\epsilon/\bar{C})}{\ln(\alpha\beta)}.$$

8

Thus a horizon guaranteed to achieve an error of at most ϵ depends only on the error ϵ, the discount factor α, the coupling coefficient β, and a uniform bound on single period costs, $\bar{C}$. Note that T_ϵ^* goes to zero as uncertainty about the future increases.

We turn now to the average cost case where, in the absence of discounting, the criterion is to minimize the average cost per period, $A(x)$ over all feasible strategies x.

Corollary 2.5 *Suppose*

$$\bar{\beta} = \max_k(1 - \min_{x^k \in D^k} \max_j \min_i P_{ij}^k(x^k)) < 1.$$

Then the error in average cost *between a rolling horizon strategy $\hat{x}$ and an average optimal strategy x^* is bounded by $\beta^T \bar{C}e$, i.e.*

$$\hat{A} - A^* \leq \beta^T \bar{C}e$$

for all T, where $\hat{A} = A(\hat{x})$, $A^* = A(x^*)$, and

$$A(x) = \liminf_{H \to \infty} \frac{1}{H} \sum_{k=1}^{H} P^{1,k-1}(x^{1,k-1})C^k(x^k).$$

Proof: Analogous to the proof of Theorem 2.4, it is possible to obtain the following error bound formula for the problem (P) with finite horizon $H < \infty$ (see Alden and Smith [1986])

$$\hat{C}_H - C_H^* \leq \frac{(\alpha\beta)^T}{1 - \alpha\beta} \left[1 + \frac{(1 - \beta)(\alpha - \alpha^{H-T})}{1 - \alpha}\right] \bar{C}e.$$

Setting $\alpha = 1$, we get

$$\hat{C}_H - C_H^* \leq \frac{\beta^T}{1 - \beta}[1 + (1 - \beta)(H - T - 1)]\bar{C}e.$$

Then $\liminf_{H \to \infty} \hat{C}_H/H - \liminf_{H \to \infty} C_H^*/H = \liminf_{H \to \infty}(\hat{C}_H - C_H^*)/H \leq \liminf_{H \to \infty}[\beta^T(1 + (1 - \beta)(H - T - 1))\bar{C}e]/((1 - \beta)H) = \beta^T \bar{C}e$. The first two liminfs exist (i.e., they are finite) since period costs are uniformly bounded by $\hat{C} < \infty$. Now

$$\liminf_{H \to \infty} \frac{\hat{C}_H}{H} = A(\hat{x})$$

since $\hat{x}_H^{1,H-T} = \hat{x}^{1,H-T}$ where $\hat{x}_H$ is the rolling horizon strategy generated for the H horizon problem. Also

$$\liminf_{H \to \infty} \frac{C_H^*}{H} = A(x^*)$$

where x^* is a periodic forecast horizon optimal strategy since for the case $\bar{\beta} < 1$, x^* is average optimal (see Hopp, Bean, and Smith[1984]) and, by definition, $x_{H_n}^* \to x^*$ for some $H_n \to \infty$ where x_H^* is an optimal strategy for horizon H. Hence $\hat{A} - A^* \leq \beta^T \bar{C}e$ for all T. ∎

Corollary 2.5 states that even in the absence of discounting, when $\bar{\beta} < 1$ and the rolling horizon is fixed, the error between a rolling horizon strategy and an average cost optimal strategy goes to zero as the coupling coefficient goes to zero. The form of the bound makes a precise analogy to discounting with β adopting the role of an effective discount rate. Corollary 2.5 is particularly important from a practical point of view since average cost criteria are frequently employed in practice.

We apply the bound of Corollary 2.5 to a problem of vehicle deployment in the next section.
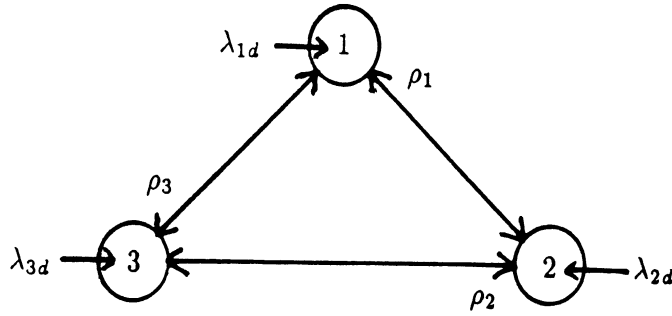
Figure 2: A three station dynamic vehicle deployment problem.

3 An Application to a Vehicle Deployment Problem

The following model is motivated by a real life truck deployment problem considered by Powell [1986]. Consider the problem of assigning a fixed number of vehicles M to a given set of stations numbered $1, 2, \cdots, K$ in order to move loads between them to maximize revenues minus costs. Loads randomly arrive at each station at the beginning of each day forming a queue. Vehicles assigned to a station at the end of the previous day serve this queue with each vehicle carrying exactly one load until there are either no more loads or no more vehicles at a station. The number of loads to be moved at a station is random depending on both the station and the day. The destination of a load is also random depending on the station at which the load arrived. Since a vehicle can carry exactly one load per day, the loads in excess of the number of vehicles at a station will be lost with no costs or revenues incurred. Alternatively, the vehicles in excess of the number of loads at a station will not carry any loads during the day and remain at the station unless reassigned to some other station later in the day. There is a fixed profit, R, for moving a load from a station to its destination. After all vehicles are finished moving their loads, any number of vehicles at a station can be assigned and moved to other stations to move loads the next day. There is a fixed cost, C, to assign and move an empty vehicle to another station.

The period in this problem is a day. The state of a period is the number of vehicles at each station after all vehicles have moved their loads for the day. A decision for a state is an assignment of vehicles to station in preparation for the next day. A policy for a period is a specification of an assignment for each possible state of the period. The optimization problem is to select a strategy, which is a vehicle assignment for each state of each period, that minimizes the average expected net costs per day. The rolling horizon procedure is simple: find a least cost strategy for the problem over a fixed horizon, roll forward one day using the first policy of that strategy, and repeat the process indefinitely. We now turn to the question of determining a rolling horizon T sufficiently large to produce an acceptable error from optimum.

For our particular example, we will consider the three station problem of Figure 2. We assume that the number of loads that arrive during the morning at a station has a Poisson distribution with parameter λ_{id} for station $i = 1, 2, 3$ and day $d = 1, 2, \cdots, 7$ of the week. We assume for purposes of exposition that no seasonal variation is present. Let N_{id} be the random variable for the number of vehicles that arrive at station i during the dth day of a week. The probability distribution of N_{id}, for $n = 0, 1, 2, \cdots$, is Poisson:

$$P\{N_{id} = n\} = \frac{\lambda_{id}^n e^{-\lambda_{id}}}{n!}.$$

Assume that a load at station i will require a move to the adjacent station in a clockwise direction with probability p_i and will require a move in the counterclockwise direction with probability $1 - p_i$. A state of a

10

period is represented by the vector (k_1, k_2, k_3) where k_i is the number of vehicles at station i. An assignment of vehicles is represented in a similar manner as (a_1, a_2, a_3) where a_i is the number of vehicles assigned to station i. The transition probabilities between states of two consecutive periods depend on the assignment of vehicles in the former period and are driven by the probabilities $P\{N_{id} = n\}$ and p_i. For example, if all M vehicles are assigned to station 1, corresponding to the assignment $(M, 0, 0)$, then the probability that they all go to station 2, or the state $(0, M, 0)$ as a result of moving loads on day d of the week is

$$P\{(M,0,0),(0,M,0)\} = \left(1 - \sum_{n=0}^{M-1} \frac{\lambda_{1d}^n e^{-\lambda_{1d}}}{n!}\right) p_1^M.$$

The expected net cost of an assignment is C times the number of empty vehicles moved in the assignment less R times the expected number of loads moved the next day. In this initial formulation, the expected cost of an assignment can be negative. However, by subtracting a lower bound $\underline{C} > -\infty$ over all expected net costs, we obtain an equivalent problem where all costs are nonnegative as required. A simple lower bound on the expected cost of an assignment is provided by assuming that no vehicles are moved in a assignment and that all vehicles happen to move a load in the following day. This gives a lower bound of $\underline{C} = -MR$.

We turn now to computing an upper bound on costs, $\bar{C}$, for use in Corollary 2.5 to bound the error from optimal. For the Poisson distribution, the expected number of loads moved is minimized by the assignment of placing all vehicles at the station having the lowest expected number of arrivals. Such an assignment will minimize the expected profits of moving loads. Assuming the worst case situation, where this assignment requires all vehicles to move empty to other stations, we obtain the following bound on the maximum expected cost of an assignment.

$$\bar{C} = M(C + R) - R \min_{i,d} \left[\sum_{n=1}^{M-1} nP\{N_{id} = n\} + M \sum_{n=M}^{\infty} P\{N_{id} = n\}\right].$$

Typical values for truck deployment (Powell [1986]) are: C =$100 to move an empty vehicle, R =$400 profit per load moved, routing probabilities of $p_1 = p_2 = p_3 = 0.5$, and expected number of arrivals per station per day of the week as given in Table 3. For these data and for the case of five vehicles ($M = 5$), the above bound is $\bar{C} = \$1350$.

We illustrate the calculation of the coupling coefficient β, for this problem containing only three vehicles. Later, we will consider a larger problem with five vehicles when actually applying Corollary 2.5. The reason we consider the larger problem is to show that sometimes, even if the coupling coefficient is near one for the problem as stated, by eliminating decisions that are clearly non-optimal in any strategy or substrategy we may find a much smaller coupling coefficient.

Recall that the coupling coefficient of a period is calculated from the set of transition matrices of the period, and that a transition matrix for a period is defined by specifying the decision for each state of the period. Specifying a decision for one state effectively defines a transition vector from the state to states of the next period and so defines a row of the transition matrix for the period. In our example, the decision set of a state, that is, the set of possible vehicle assignments, is state and period independent. Moreover, the transition probabilities given a particular vehicle assignment is also state independent (but period dependent). Hence , the set of candidate transition vectors for any row of a period's transition matrix is row independent. This means the coupling coefficient for a period in this problem is equal to the coupling coefficient of the matrix whose rows represent the set of unique candidate transition vectors for a single row state. We call this matrix the *candidate* matrix for a state. In this example, the candidate matrix for a state is the same for each state of a period, hence the matrix representing the set of unique transition vectors taken from the union of all candidate matrices for each state of a period, or the candidate matrix of a period, is equal to the candidate matrix for any given state of the period.

Table 4 contains the candidate matrix for the first day of a week in our example with three vehicles. The last line of the table contains the column minimums whose sum is 0.3025. Hence, the coupling coefficient is 0.6975 for the first day of a week.

Now we consider the case of five vehicles. Calculations of the coupling coefficient for each day of the week were made using the approach illustrated above and the results are given in the column headed by the number 5 in Table 5. Ignore, for the moment, the remaining columns of Table 5. The maximum coupling coefficient over each day of the week, and hence over all days in the problem, is 0.9606. This number is near one which means that a small error bound from Corollary 2.5 will require a large rolling horizon. However, the coupling coefficient we need is only for the set of transition matrices that could be a part of an optimal substrategy in the rolling horizon procedure or part of an optimal strategy since the formulation of (P) may clearly include constraints that fail to exclude an optimal solution. We argue in Alden and Smith [1984] that it is suboptimal to assign more than 3 vehicles to any station. The column headed 3 in Table 5 contains the results of using this limit. Thus, we can use $\beta = 0.7106$ for this particular problem, which is significantly lower than our initial value of 0.9606. The other columns of Table 5 are included only for completeness sake in showing the sensitivity of the coupling coefficient to the assignment limit. We turn now to Corollary 2.5. Since we are interested in selecting a horizon T to guarantee the average error of a rolling horizon solution is small, say no more than ϵ, we set the right hand side of the expression in the corollary to less than or equal to ϵ and solve for a lower bound on the horizon. Letting $f = \epsilon/\bar{C}$ represent the maximum expected relative error,

$$T \geq \frac{\ln(f)}{\ln(\beta)}.$$

For our β of 0.7106, we find that a fixed horizon of 14 days or more will guarantee the average error of a rolling horizon procedure is no more than one percent of the maximum expected cost of a decision (i.e., $f = 0.01$). Note that this error is no more than $13.50 per day which is one percent of the bound $\bar{C} \leq \$1350$ we calculated above. Table 6 contains the horizon for this and other values of β found in Table 5.

The bound on the rolling horizon of 14 days is remarkably small and illustrates the power of the coupling coefficient in reducing error. A horizon bound to achieve a comparable error in the presence of discounting is on the order of years. We suspect that the 14 day bound grows as the number of vehicles in the problem increases, since the stochastic effects correspondingly diminish.

References

1. Alden, J. and R.L. Smith [1986], "Error Bounds for Rolling Horizon Procedures in Markov Decision Processes", Technical Report 86-10, Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, Michigan 48109.

2. Baker, K. [1977], "An Experimental Study of the Effectiveness of Rolling Schedules in Production Planning," Decision Science Vol. 8, pp.19-27.

3. Bean, J., Birge, J., and R.L. Smith [1987], "Aggregation in Dynamic Programming,"Operations Research Vol. 35, pp. 215-220.

4. Bean, J.and R.L. Smith [1984], "Conditions for the Existence of Planning Horizons," Mathematics of Operations Research, Vol. 9, 1984, pp. 391-401.

5. Bean, J. and R.L. Smith [1986], "Conditions for the Discovery of Solution Horizons," Technical Report 86-23, Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, Michigan 48109.

6. Bhaskaran, S. and S. Sethi [1985], "Conditions for the Existence of Decision Horizons for Discounted Problems in a Stochastic Environment: A Note," Operations Research Letters, Vol. 4, p. 61-65.

7. Denardo, E. [1982], Dynamic Programming: Models and Applications, Prentice-Hall,Inc., Englewood Cliffs, New Jersey.

8. Hopp, W., J. Bean, and R.L. Smith [1984], "Non-Homogeneous Markov Decision Processes," Technical Report 84-24, Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, Michigan 48109, October, 1984. Forthcoming in Operations Research.

9. Lee, C. and Denardo, E. [1983], "Rolling Planning Horizons: Error Bounds for the Dynamic Lot Size Model", Working Paper, Yale University. Forthcoming in Mathematics of Operations Research.

10. Powell, W. [1986] Private Communication, Department of Civil Engineering, Princeton University.

11. Seneta, E. [1981],Non-negative Matricies and Markov Chains, New York: Springer-Verlag.

Appendix

Lemma 1.1 $\hat{C} \le C_1^* + \sum_{l=1}^{\infty} E(\underline{e}_l)$ where $\underline{e}_l$ is a random vector under strategy $\hat{x}$ whose i^{th} component, $i = 1, 2, \cdots, n$, takes on the value $(e_l)_{I_l}$ when the initial state of (P) beginning period 1 is i and I_l is the state entered beginning period l.

Proof: Consider the extended strategy $\xi^{l+1,l+T}(l) = (z^{l+1}(l), z^{l+2}(l), \cdots, z^{l+T-1}(l), x^{l+T}(l))$ where $x^{l+T}(l)$ is the argmin yielding C_l. $\xi(l)$ is feasible for the $l+1^{st}$ subproblem and its cost is therefore greater than or equal to the optimal cost, i.e.,

$$C^{l+1,l+T}(\xi(l)) \ge C_{l+1}^*.$$

Premultiplying both sides by $P^l(\hat{x}^l) = P^l(z^l(l))$ we get

$$P^l(z^l(l))C^{l+1,l+T}(\xi(l)) \ge P^l(z^l(l))C_{l+1}^*.$$

Now $C^{l+1,l+T}(\xi(l)) = \sum_{j=l+1}^{l+T} \alpha^{j-1} P^{l+1,j-1}(\xi^{l+1,j-1}(l))C^j(\xi^j(l)) = \sum_{j=l+1}^{l+T-1} \alpha^{j-1} P^{l+1,j-1}(z^{l+1,j-1}(l))$
$C^j(z^j(l)) + \alpha^{l+T-1} P^{l+1,l+T-1}(z^{l+1,l+T-1}(l))C^{l+T}(x^{l+T}(l))$. Hence, $\sum_{j=l+1}^{l+T-1} \alpha^{j-1} P^{l,j-1}(z^{l,j-1}(l))C^j(z^j(l)) +$
$\alpha^{l+T-1} P^{l,l+T-1}(z(l))C^{l+T}(x^{l+T}(l)) \ge P^l(z^l(l))C_{l+1}^*$, or $\sum_{j=l}^{l+T-1} \alpha^{j-1} P^{l,j-1}(z^{l,j-1}(l))C^j(z^j(l)) -$
$\alpha^{l-1} P^{l,l-1}(z^{l,l-1}(l))C^l(z^l(l)) + \alpha^{l+T-1} P^{l,l+T-1}(z(l))C^{l+T}(x^{l+T}(l)) \ge P^l(z^l(l))C_{l+1}^*$. Since $P^{l,l-1}(z^{l,l-1}(l)) = I$, we have

$$C_l^* - \hat{C}(l,l) + e_l \ge P^l(\hat{x}^l)C_{l+1}^*, \quad l = 1, 2, \cdots.$$

Premultiplying both sides of the inequality by $P^{1,l-1}(\hat{x}^{1,l-1})$ we get $P^{1,l-1}(\hat{x}^{1,l-1})C_l^* - P^{1,l-1}(\hat{x}^{1,l-1})\hat{C}(l,l) +$
$P^{1,l-1}(\hat{x}^{1,l-1})e_l \ge P^{1,l}(\hat{x}^{1,l})C_{l+1}^*$ or

$$E(\hat{\underline{C}}(l,l)) \le E(\underline{C}_l^*) - E(\underline{C}_{l+1}^*) + E(\underline{e}_l), \quad l = 1, 2, \cdots.$$

Summing the above inequality over $l = 1, 2, \cdots$ we get

$$\sum_{l=1}^{\infty} E(\hat{\underline{C}}(l,l)) \le E(\underline{C}_1^*) + \sum_{l=1}^{\infty} E(\underline{e}_l).$$

Since

$$\sum_{l=1}^{\infty} E(\hat{\underline{C}}(l,l)) = \sum_{l=1}^{\infty} \alpha^{l-1} P^{1,l-1}(\hat{x}^{1,l-1})C^l(\hat{x}^l)$$
$$= \hat{C}$$

and $C_1^* = E(\underline{C}_1^*)$, we have

$$\hat{C} \le C_1^* + \sum_{l=1}^{\infty} E(\underline{e}_l) \quad \blacksquare$$

Lemma 2.1 Every one step transition probability matrix in (P) can be expressed as a convex combination of a stochastic matrix and a stable matrix (i.e., a stochastic matrix of equal rows) dependent only on period where the multiplier of the former, β is given by

$$\beta = \max_k (1 - \min_{x^k \in D^k} \sum_j \min_i P_{ij}^k(x^k)).$$

That is, for all k and all $x^k \in D^k$,

$$P^k(x^k) = \beta \tilde{P}^k(x^k) + (1 - \beta)\tilde{L}^k$$

where $\tilde{P}^k(x^k)$ is a stochastic matrix, $\tilde{L}^k$ is a stable stochastic matrix independent of x^k, and $0 \leq \beta \leq 1$ is given as above.

Proof: We say that a matrix X is dominated by a comparable matrix Y if $x_{ij} \leq y_{ij}$ for all i, j. Let L^k be the stable matrix of greatest row sum dominated by every matrix $P^k(x^k)$, $x^k \in D^k$ and let $\beta_k = 1 - \sum_j L^k_{ij}$. By construction,

$$\beta_k = 1 - \min_{x^k \in D^k} \sum_j \min_i P^k_{ij}(x^k)$$

and hence $\beta = \max_k \beta_k$. Let

$$\tilde{L}^k = \frac{L^k}{1 - \beta_k}$$

if $\beta_k < 1$ and otherwise set $\tilde{L}^k$ to an arbitrary stable stochastic matrix. In either case $\tilde{L}^k$ is a stable stochastic matrix. Also let

$$\tilde{P}^k(x^k) = \frac{1}{\beta}(P^k(x^k) - (1 - \beta)\tilde{L}^k)$$

which is a stochastic matrix since

$$\begin{aligned}
\sum_j \tilde{P}^k_{ij}(x^k) &= \tfrac{1}{\beta}(\sum_j P^k_{ij}(x^k) - (1 - \beta)\sum_j \tilde{L}^k_{ij}) \\
&= \tfrac{1}{\beta}(1 - (1 - \beta)) \\
&= 1
\end{aligned}$$

and

$$\begin{aligned}
\tilde{P}^k_{ij} &= \tfrac{1}{\beta}(P^k_{ij}(x^k) - (1 - \beta)\tilde{L}^k_{ij}) \\
&\geq \tfrac{1}{\beta}(P^k_{ij}(x^k) - (1 - \beta_k)\tilde{L}^k_{ij}) \\
&\geq \tfrac{1}{\beta}(P^k_{ij}(x^k) - L^k_{ij}) \\
&\geq 0.
\end{aligned}$$

Then $P^k(x^k) = \beta \tilde{P}^k(x^k) + (1 - \beta)\tilde{L}^k$. $\blacksquare$

Lemma 2.2 (P) and $(\tilde{P})$ are equivalent Markov Decision Processes.

Proof: Let $P(l, k)$ and $\tilde{P}(l, k)$ be the subproblems of P and $\tilde{P}$ respectively covering periods l through k. We fix k and begin with the inductive hypothesis $H(l, k)$ that for $l = 1, 2, \cdots, k$ states

1. The optimal substrategies $x^*(l, k)$ and $\tilde{x}^*(l, k)$ are identical for $P(l, k)$ and $\tilde{P}(l, k)$, respectively,

2. The optimal costs $C^*(l, k)$ and $\tilde{C}^*(l, k)$ discounted to the beginning of period l are related by $C^*(l, k) = \tilde{C}^*(l, k) + \theta_{lk}$ where $\theta_{lk} = \sum_{j=0}^{k-l-1}(\alpha\beta)^j \gamma_{l+j}$ and $\gamma_j = \alpha(1 - \beta)\tilde{L}^j C^*(j + 1, k)$ is $\alpha(1 - \beta)$ times the optimal cost for problem (P) beginning period $j + 1$ with the initial distribution $\tilde{L}^j$.

It is clear that $H(k, k)$ is true with $\theta_{kk} = 0$ since

$$C^*(k, k) = \tilde{C}^*(k, k) = \min_{x^k \in D^k} C^k(x^k).$$

Now assume that $H(l, k)$ is true for $l = j, j+1, \cdots, k$. We now show that $H(l, k)$ is true for $l = j - 1$. Subproblem $P(j - 1, k)$ phrased in terms of C becomes

$$\min_{x^{j-1,k} \in D^{j-1,k}} C^{j-1,k}(x^{j-1,k})$$

or

$$\min_{x^{j-1,k} \in D^{j-1,k}} \sum_{i=j-1}^{k} \alpha^{i-j+1} P^{j-1,i-1}(x^{j-1,i-1}) C^i(x^i).$$

Rewriting this gives

$$\min_{x^{j-1,k} \in D^{j-1,k}} \quad \{ P^{j-1,j-2}(x^{j-1,j-2}) C^{j-1}(x^j - 1)$$

$$+ \quad \alpha \sum_{i=j}^{k} \alpha^{i-j} P^{j-1,i-1}(x^{j-1,i-1}) C^i(x^i)\}.$$

or since $P^{j-1,j-2}(x^{j-1,j-2}) = I$,

$$\min_{x^{j-1} \in D^{j-1}} \quad \{ C^{j-1}(x^{j-1})$$

$$+ \quad \alpha P^{j-1}(x^{j-1}) \min_{x^{j,k} \in D^{j,k}} \sum_{i=j}^{k} \alpha^{i-j} P^{j,i-1}(x^{j,i-1}) C^i(x^i)\}.$$

Making the substitution $P^{j-1}(x^{j-1}) = \beta \tilde{P}^{j-1}(x^{j-1}) + (1 - \beta) \tilde{L}^{j-1}$ from Lemma 2.1 and noting that the resulting term involving $\tilde{L}^{j-1}$ is independent of x^{j-1}, we get

$$\min_{x^{j-1} \in D^{j-1}} \quad \{ C^{j-1}(x^{j-1})$$

$$+ \quad \alpha \beta \tilde{P}^{j-1}(x^{j-1}) \min_{x^{j,k} \in D^{j,k}} \sum_{i=j}^{k} \alpha^{i-j} P^{j,i-1}(x^{j,i-1}) C^i(x^i)\}$$

$$+ \quad \alpha(1 - \beta) \tilde{L}^{j-1} \min_{x^{j,k} \in D^{j,k}} \sum_{i=j}^{k} \alpha^{i-j} P^{j,i-1}(x^{j,i-1}) C^i(x^i)$$

or

$$\min_{x^{j-1} \in D^{j-1}} \{ C^{j-1}(x^{j-1}) + \alpha \beta \tilde{P}^{j-1}(x^{j-1}) C^*(j, k)\} + \gamma_{j-1}.$$

Invoking the inductive hypothesis $H(l, k)$ for $l = j$ and substituting $C^*(j, k) = \tilde{C}^*(j, k) + \theta_{jk}$, we get

$$\min_{x^{j-1} \in D^{j-1}} \{ C^{j-1}(x^{j-1}) + \alpha \beta \tilde{P}^{j-1}(x^{j-1}) \tilde{C}^*(j, k)\} + \alpha \beta \theta_{jk} + \gamma_{j-1}.$$

or, by the principle of optimality,

$$\min_{x^{j-1} \in D^{j-1}} \{ \tilde{C}^{j-1,k}(x^{j-1,k})\} + \alpha \beta \theta_{jk} + \gamma_{j-1}.$$

We conclude first of all that the argument x^{j-1} solving subproblem $P(j - 1, k)$ also solves $\tilde{P}(j - 1, k)$. Secondly, that

$$C^*(j - 1, k) = \tilde{C}^*(j - 1, k) + \alpha \beta \theta_{jk} + \gamma_{j-1}$$

16

where $\theta_{jk} = \sum_{i=0}^{k-j-1}(\alpha\beta)^i\gamma_{j+1}$. Hence

$$C^*(j-1,k) = \tilde{C}^*(j-1,k) + \sum_{i'=0}^{k-j}(\alpha\beta)^{i'}\gamma_{j+i'-1}$$

where $i' = i + 1$ thus restoring the inductive hypothesis for $l = j - 1$. $\blacksquare$

Lemma 2.3 The error in cost between a rolling horizon strategy $\hat{x}$ and an optimal strategy x^* for problem (P) is given in terms of problem $(\tilde{P})$ by

$$\hat{C} - C^* = \tilde{\hat{C}} - \tilde{C}^* + (1-\beta)\sum_{l=1}^{\infty}\beta^{-l}\tilde{L}^l\left(\tilde{\hat{C}}(l+1) - \tilde{C}^*(l+1)\right).$$

Proof: We first establish by induction that, for all $x \in D$,

$$P^{1k}(x^{1k}) = \beta^k\tilde{P}^{1k}(x^{1k}) + (1-\beta)\sum_{i=1}^{k}\beta^{k-i}\tilde{L}^i\tilde{P}^{i+1,k}(x^{i+1,k}).$$

The relationship holds for $k = 1$ since $P^1(x^1) = \beta\tilde{P}^1(x^1) + (1-\beta)\tilde{L}^1$. Suppose now it holds for $k = 1, 2, \cdots, j-1$. We show it holds for $k = j$. By definition $P^{1j}(x^{1j}) = P^{1,j-1}(x^{1,j-1})P^j(x^j) = \beta^{j-1}\tilde{P}^{1,j-1}(x^{1,j-1}) + (1-\beta)\sum_{i=1}^{j-1}\beta^{j-i-1}\tilde{L}^i\tilde{P}^{i+1,j-1}(x^{i+1,j-1}))(\beta\tilde{P}^j(x^j) + (1-\beta)\tilde{L}^j)$. Since $PL = L$ for any stochastic matrix P and stable matrix L, $P^{1j}(x^{1j}) = \beta^j\tilde{P}^{1j}(x^{1j}) + (1-\beta)\sum_{i=1}^{j-1}\beta^{j-i}\tilde{L}^i\tilde{P}^{i+1,j}(x^{i+1,j}) + (1-\beta)\beta^{j-1}\tilde{L}^j + (1-\beta)^2\sum_{i=1}^{j-1}\beta^{j-i-1}\tilde{L}^j = \beta^j\tilde{P}^{1j}(x^{1j}) + (1-\beta)\sum_{i=1}^{j-1}\beta^{j-i}\tilde{L}^i\tilde{P}^{i+1,j}(x^{i+1,j}) + ((1-\beta)\beta^{j-1} + (1-\beta)^2\sum_{i=1}^{j-1}\beta^{j-i-1})\tilde{L}^j = \beta^j\tilde{P}^{1j}(x^{1j}) + (1-\beta)\sum_{i=1}^{j}\beta^{j-i}\tilde{L}^i\tilde{P}^{i+1,j}(x^{i+1,j})$ since $(1-\beta)\beta^{j-1} + (1-\beta)^2\sum_{i=1}^{j-1}\beta^{j-i-1} = 1 - \beta$. Hence the relationship holds for $k = j$.
Substituting this relationship, we obtain, for all $x \in D$, the identity

$$\sum_{k=1}^{\infty}\alpha^{k-1}P^{1,k-1}(x^{1,k-1})C^k(x^k) = \sum_{k=1}^{\infty}(\alpha\beta)^{k-1}\tilde{P}^{1,k-1}(x^{1,k-1})C^k(x^k)$$

$$+(1-\beta)\sum_{k=2}^{\infty}\alpha^{k-1}\sum_{i=1}^{k-1}\beta^{k-i-1}\tilde{L}^i\tilde{P}^{i+1,k-1}(x^{i+1,k-1})C^k(x^k).$$

Since all terms are nonnegative in the double sum above, we can reverse the order of summation by Fubinis' Theorem to obtain for all $x \in D$

$$\sum_{k-1}^{\infty}\alpha^{k-1}P^{1,k-1}(x^{1,k-1})C^k(x^k) = \sum_{k=1}^{\infty}(\alpha\beta)^{k-1}\tilde{P}^{1,k-1}(x^{1,k-1})C^k(x^k)$$

$$+(1-\beta)\sum_{i=1}^{\infty}\beta^{-i}\tilde{L}^i\sum_{k=i+1}^{\infty}(\alpha\beta)^{k-1}\tilde{P}^{i+1,k-1}(x^{i+1,k-1})C^k(x^k).$$

Substituting $x = \hat{x}$ and $x = x^*$ we get, respectively

$$\hat{C} = \tilde{\hat{C}} + (1-\beta)\sum_{i=1}^{\infty}\beta^{-i}\tilde{L}^i\tilde{\hat{C}}(i+1)$$

17

and

$$C^* = \tilde{C}^* + (1 - \beta) \sum_{i=1}^{\infty} \beta^{-i} \tilde{L}^i \tilde{C}^*(i + 1)$$

since $\hat{\tilde{x}} = \hat{x}$ and $\tilde{x}^* = x^*$. Subtracting the above equalities, we get

$$\hat{C} - C^* = \hat{\tilde{C}} - \tilde{C}^* + (1 - \beta) \sum_{i=1}^{\infty} \beta^{-i} \tilde{L}^i (\hat{\tilde{C}}(i + 1) - \tilde{C}^*(i + 1)) \quad \blacksquare$$

Table 1: Expected Daily Arrivals by Station and Day of Week

Day	λ_{1d}	λ_{1d}	λ_{1d}
1	1	1	1
2	2	1	1
3	1	2	1
4	1	1	2
5	2	2	1
6	2	1	2
7	1	2	2

Table 2: Transition Vectors for each Assignment of Day 1

Assignment	003	012	021	030	102	111	120	201	210	300
003	.3678	.1839	.0459	.0100	.1839	.0919	.0301	.0459	.0301	.0100
012	.1162	.1934	.0885	.0243	.1744	.1675	.0694	.0790	.0660	.0208
021	.0243	.0885	.1934	.1162	.0694	.1675	.1744	.0660	.0790	.0208
030	.0100	.0459	.1839	.3678	.0301	.0919	.1839	.0301	.0459	.0100
102	.1162	.1744	.0790	.0208	.1934	.1675	.0660	.0885	.0694	.0243
111	.0367	.1110	.1110	.0367	.1110	.2231	.1110	.1110	.1110	.0367
120	.0208	.0790	.1744	.1162	.0660	.1675	.1934	.0694	.0885	.0243
201	.0243	.0694	.0660	.0208	.0885	.1675	.0790	.1934	.1744	.1162
210	.0208	.0660	.0694	.0243	.0790	.1675	.0885	.1744	.1934	.1162
300	.0100	.0301	.0301	.0100	.0459	.0919	.0459	.1839	.1839	.3678
Minimums:	.0100	.0301	.0301	.0100	.0301	.0919	.0301	.0301	.0301	.0100

Table 3: Coupling Coefficients for each Day of the Week

Day of Week	Limit on Number of vehicles Assigned to a Station			
	5	4	3	2
1	.9606	.8746	.7106	.3909
2	.9296	.8127	.5893	.2986
3	.9296	.8127	.5893	.2986
4	.9296	.8127	.5893	.2986
5	.8513	.7267	.4668	.1969
6	.8513	.7267	.4668	.1969
7	.8513	.7267	.4668	.1969
Maximums:	.9606	.8746	.7106	.3909

Table 4: Horizons for the Coupling Coefficients of Table 5 with $f = .01$

Assignment Limit	coupling Coefficient	Horizon in Days
5	.9606	114.56
4	.8746	34.37
3	.7106	13.48
2	.3909	5.01