Journal of the Nigerian Association of Mathematical Physics
Volume **15 (November, 2009), pp 253 - 264**

On a dynamic costs optimality for a production company

[1]Iwebuke Charles Nkeki and [2]Chukwuma Raphael Nwozo
[1]*Department of Mathematics, University of Benin,*
Benin City, Nigeria.
[2]*Department of Mathematics, University of Ibadan,*
Ibadan, Nigeria.

Abstract

The allocation of human and physical resources over time is a fundamental problem that is central to science, management science and engineering. In this paper, we consider the dynamics of allocation of resources at a minimum cost in a production company in Nigeria. The company is assumed to be made up of different departments. Our aim is to consider problem arising from freight department. We consider a process where by n-dimensional vector functions F with error sequence $\| \Gamma^k F - F^* \| \leq \beta^k N$, ***for all*** $N \in \Re, 0 \leq \beta < 1$ ***is use to determined the minimum costs of distributing products from production centres to the markets. We found that the minimum costs of the operations converges to minimum costs when error bounds are included. In other word, minimum costs of the operation with error bounds and without error bounds assumed the same vector values only at infinite stage. We also found out that the first production centre has the minimum costs.***

Keywords
Value Iteration; Error Bounds; Dynamic Costs Optimality; Transition Cost; Transition matrices.

1.0 Introduction

The allocation of human and physical resources over time is a fundamental problem in science and engineering. In a production company, for instant, must manage personnel and equipment as well as shipments in a timely manner in the presence of a variety of dynamic information process such as customer demands, equipment failures, weather delay, accident on the way and failures of execution. This is a high-dimensional problem since it comprises a large number of resources, each of which must be considered as it is affected by decisions and uncertainties, (see [4]). The problem of dynamic resource allocation can be treated as Markov decision processes and solved using value iteration. Markov decision processes provide a unified framework for the treatment of problems of sequential decision-making under uncertainty. For a variety of optimality criteria, these problems can be solved by dynamic programming via value iteration. Value iteration play tremendous role in determining the value

[1]**Corresponding author:**
[1]**e-mail address: nkekicharles2003@yahoo.com;**
[1]**Telephone: +234-08038667530**
[2]**e-mail address: crnwozo@yahoo.com**

[2]Telephone: +234-08056028461

of cost or returns in organisation settings. It is used to successively approximate costs value function(s) as well as return value function(s).

In this paper, our approach builds on previous research on dynamic programming principles, (DPP). Many authors have used this principles in solving problems arising from various settings. [3], presented computationally efficient approximate dynamic programming algorithms for application to problems in freight transportation. In their work, they considered problems arising from shippment in a sea port. Mulvey and Vladimirou [7], used the stochastic programming technique of dynamic programming in financial asset allocation problems for designing low-risk portfolios. They found out that the use of DPP brings about computational efficiency even under uncertainties. Van Roy et al [8], proposed the idea of using a parsimonious sufficient static in an application of approximate dynamic programming to inventory management. Powell [5], used dynamic programming for large-scale asset management problems for both single and multiple assets. Topaloglu and Kunnumkal [7], extended an approximate dynamic programming method to optimize the distribution operations of a company manufacturing certain products at multiple production plants and shipping to different customer locations for sales. But they did not considered the possibility of uncertainty that may arise on the process of distribution of the products. In this paper, we intend to consider that as well. Nkeki [1], considered the allocation of buses from a single station to different routes for profit maximization. They developed efficient algorithm of DPP for the allocation of buses putting into consideration the possibility of break down as a result of bad roads and depreciation. Nkeki and Nwozo [2], considered the use of value iteration to minimize the costs of shipping different goods without error bounds. In this paper, we consider the dynamic of distribution of products from production centre to the markets at minimum costs with the possibility of uncertainty that may arise on the process of operation. We also considered the problem using error bounds.

1.1 Basic definitions and assumptions

β: discount factor, $0 < \beta < 1$.

S: the state space i.e. the set of all buses

T: set of time periods in the planning horizon.

$\pi : S \rightarrow X$: is a rule which chooses an action $x \in X$ based on current state of the system

s_t: number of products at period t, $s_t \in S$.

φ_t^n : expected return of products from centre n at period, t.

Π: set of all admissible policy; $\pi \in \Pi$

m: number of market under consideration

n : number of production centre

s_0 : number of products at period $t = 0$.

We make the following assumptions:

(1) The amount spent for distributing products from production centres to the markets depend on the distance covered and the nature of the routes.

(2) There are n production centres and n markets.

(3) Each production centre produce a unique product.

(4) The production centres distribute their product to all the markets independently.

1.2 One-period expected costs function

Suppose that the costs of distributing the products from the production centre S_i to the market S_k is $\varphi_t(i,k)$ at period t, the number of products in production centre is given as S_t^{pc} at period t and the number of products in the markets is given as S_t^{mc} at period *t*, then the costs over *T*-horizon is $\sum_{t=0}^{T}\varphi_t(i,k)(S_t,\pi(S_t))$, (see [9]).

Let $c_h, h=1,2,...,n$ be different kinds of products to be distributed from the production centre to the markets. Then, $\sum_{h=1}^{m} c_{h,t}^{pc} \geq \sum_{h=1}^{m} c_{h,t}^{mc}$, $c^{pc}, c^{mc} \in C, t \in T$, where $c_{h,t}^{pc}$ is the number of products before distribution at period t and $c_{h,t}^{mc}$ is the number of products that is already in the markets at period *t*.

The expected minimum costs function obtained under control policies π, at period t is given as follows:

$$\mathrm{X}_t^{\pi} = E_t[\min_{c^{pc}, c^{mc} \in C} \sum_{t=0}^{T} \beta^t \varphi_t^{\pi}(i,k)(c_t^{\pi^n}(S_{t-1}))], \tag{1.1}$$

subject to $\sum_{h=1}^{m} c_{h,t}^{pc} \geq \sum_{h=1}^{m} c_{h,t}^{mc}, c^{pc}, c^{mc} \in C, t \in T, \; c_{h,t}^{pc}, \; x_{h,t}^{mc} \geq 0, h=1,...,m$ (1.2)

where $c_{h,t}^{pc}, c_{h,t}^{mc} \in C$ is the set of feasible solutions of problem (1.1). We can express (1.1) as the expected minimal cost from period t onward as an optimization over $\{c_1, c_2, ..., c_m\}$ condition on $S_t = s_t$ as follows:

$$\mathrm{X}_t^{\pi}(S_t) = E_t[\min_{c_t,...,c_{T-1}} \{\sum_{t'=t}^{T} \beta^{t'} \varphi_{t'}^{\pi}(i,k)(c_{t'}^{\pi^n}(S_{t'-1})) \mid S_t = s_t\}], S_t \in S, c_t \in C. \tag{1.3}$$

$i = 1,2,...,n; k = 1,2,...,m$.

subject to $\sum_{h=1}^{m} c_{h,t}^{pc} \geq \sum_{h=1}^{m} c_{h,t}^{mc}, c^{pc}, c^{mc} \in C, t \in T, c_{h,t}^{pc}, x_{h,t}^{mc} \geq 0, h=1,...,m$. For a function $\varphi : S \to \Re^n$, we accumulate the cost of the first *T*-stage and add to it the terminal costs $\varphi_T(S_T) = \sum_{i=1}^{n}\sum_{k=1}^{m} \varphi(i,k)(s_T)$, then (1.3) becomes

$$\mathrm{X}_t^{\pi}(S_t) = E_t[\min_{c_t}\{\sum_{t'=t}^{T-1}\sum_{i=1}^{n}\sum_{k=1}^{m} \beta^{t'} \varphi(i,k)_{t'}^{\pi}(c_{t'}^{\pi^n}(S_{t'-1}) + \beta^T \varphi_T(S_T)\} \mid S_t = s_t\}], S_t \in S, c_t \in C \tag{1.4}$$

subject to $\sum_{h=1}^{m} c_{h,t}^{pc} \geq \sum_{h=1}^{m} c_{h,t}^{mc}, c^{pc}, c^{mc} \in C, t \in T, c_{h,t}^{pc}, x_{h,t}^{mc} \geq 0, h=1,...,m$.

1.3 Dynamic programming formulation

The company decides to minimize the costs of distributing m kinds of products from n numbers of production centres to m numbers of markets. The company ships single product to all the markets over time. It is also expected that the products that leave the centres to the markets will not come back (in the case of defective, damage, e.t.c, items). The company considered n numbers of control policies, $\pi = \{\pi^1, \pi^2, ..., \pi^n\}$ to determine which of them will yield optimum control policy. They also estimated that certain percentage of the products are to reach their final destination successfully at a minimum costs. We are going to adopt the monotonic error bounds which will serve as a by product in our computational work.

Let S_t be the state variable at period *t* and *S* the state space, we formulate the problem as a dynamic program. The number of products i that leave the production centre to the market

k at period t is given by $P_{i,k}c_{h,t}^{pc}$, where $P_{i,k}$ is the transition percentage from production centre i to market k. Hence, the total expected cost of goods lost in the process of distribution is given by $\beta\sum_{h=1}^{m}\sum_{k=1}^{m}P_{i,k}c_{h,t}^{pc}$, $t = 1,\Lambda\ ,T$; $i = 1,\Lambda\ ,n.$

Let S_{t-1} be the number of products to be distributed in period *t* - 1, then S_t is the expected number of products that will get to the markets and let γ be the percentage of the products that is recovered from the lost ones which are expected to go into the markets at period t, then we have that

$$S_t = S_{t-1} - (1-\gamma)\sum_{h=1}^{m}\sum_{k=1}^{m}P_{i,k}c_{h,t}^{\pi^n},\ \ i = 1,\Lambda\ ,n,\ \ t = 1,2,\Lambda\ ,T, \tag{1.5}$$

where S_t is the products that are successfully in the markets and S_{t-1} is the goods that are in the production centre before distribution takes place . Equation (1.5) is the transformation equation and S_t is a random variable. We can express (1.5) as follows:

$$S_t^i = S_{t-1}^i - \beta\sum_{h=1}^{m}\sum_{k=1}^{m}P_{t,k}c_{h,t}^{\pi^n},\ i = 1,\Lambda\ ,n.,\ t = 1,2,,T. \tag{1.6}$$

The optimal policy can be found by computing the value functions through the optimality equation

$$F_t^{\pi}(S_t) = \min_{c_t}\varphi_t\ (i,k)(c_t^{\pi^n}(S_{t-1})) + \beta E\{F_{t+1}^{i,k}(S_{t+1})\} \mid S_t = s_t\}$$

$$= \min_{c_t}\varphi_t\ (i,k)(c_t^{\pi^n}(S_{t-1})) + \beta\sum_{k=1}^{m}P_{i,k}(\pi^n)\{F_{t-1}^{i,k}(S_t)\} \tag{1.7}$$

subject to $\sum_{h=1}^{m}c_{h,t}^{pc} \geq \sum_{h=1}^{m}c_{h,t}^{mc},\ c^{pc},c^{mc} \in C,\ t \in T$. Equivalently,

$$\sum_{h=1}^{m}c_{h,t}^{pc} = \sum_{h=1}^{m}c_{h,t}^{mc} + \eta_t,\ c^{pc},c^{mc} \in C,\ t \in T, \tag{1.8}$$

$c_{h,t}^{pc},\ c_{h,t}^{mc} \geq 0,\ h = 1,\Lambda\ ,m,\ \eta_t \geq 0$ where η_t, is the number of products that is lost in transit at period *t*. If $\eta_t = 0$, it implies that all the products that left the production centres get to the markets succussfully without damages or lost. It can be shown that (1.4) is equal to (1.7), (see [8, 12 and 14]). We may use (1.4) and (1.7) interchargeably. We now find the best control policy, π, that minimize our problem. We do that by solving the optimality equation

$$F_t^{\pi}(S_t) = \min_{c^{pc},c^{mc}\in C}\varphi_t\ (i,k)(c_t^{\pi^n}(S_{t-1})) + \beta E\{F_{t+1}^{i,k}(S_{t+1})\} \mid S_t = s_t\}. \tag{1.9}$$

If $\tilde{\varphi}_t^i(S_t^f,\pi^n,S_t^s)$ is the costs of using policy π^n at state $S_t^{pc} = i$ and moving to state $S_t^{mc} = k$ at period t, we use as costs per stage the expected costs $\varphi(S_t^{pc},\pi^n,S_t^{mc})$ given by

$$\varphi(S_t^{pc},\pi^n,S_t^{mc}) = \sum_{k=1}^{m}P_{i,k}(\pi)\tilde{\varphi}(S_t^{pc},\pi^n,S_t^{mc}),\ t \in T,\ = \sum_{k=1}^{m}P_{i,k}(\pi)\tilde{\varphi}(i,\pi,k),\ i = 1,\Lambda\ ,n.$$

Let Γ and Γ_π be mapping, such that $\Gamma : F \to \Re \; and \; \Gamma_\pi : F_\pi \to \Re$,then, we express $(\Gamma F)(S_t^i) = \min_{\pi\in\Pi(S_t)} [\varphi(i,k)(S_t^{pc}, \pi_t^n, S_t^{mc}) + \beta \sum_{k=1}^{m} P_{i,k}(\pi^n)\{F \; (S_t^j)\}, \; S_t^i = \{S_t^1,..., S_t^n\}$ and $(\Gamma_\pi F)(S_t^i) = \min_{\pi\in\Pi(S_t)} [\varphi(i,k)(S_t^{pc}, \pi_t^n, S_t^{mc}) + \beta \sum_{k=1}^{m} P_{i,k}(\pi^n)\{F_\pi \; (S_t^k)\}, S_t^i = \{S_t^1,..., S_t^n\}$.

The Lemma 1.1 below is the monotonicity property. It enables us to analysis the error bounds associated with our problem. It also enhance the computational aspect of our work.

***Lemma* 1.1**

1. The operator Γ has a unique fixed point (given by F*).
2. For any F, $\Gamma_\infty F = F^*$
3. For any F, if $\Gamma F \geq F$, then $F^* \geq \Gamma^j F$ for all k

***Proof*:** (see [14]).

***Theorem* 1.1**

Let the bounded optimal cost function $F : S \to \Re^n$ be n-dimensional vectors. Then F satisfies $F^*(S) = \lim_{j\to\infty} (\Gamma^j F)(S_t), \forall S_t \in S.$

Proof

Since F is a bounded function, then $\| F - F^* \|_\infty \leq \psi.$ It implies that

$$F(S_t) - \psi \leq F^*(S_t) \leq F(S_t) + \psi, S_t \in S, t \in T.$$

Using Lemma 1.1, we have

$$(\Gamma^k F)(S_t) - \beta^k \psi \leq F^*(S_t) \leq (\Gamma^k F)(S_t) + \beta^k \psi, \; S_t \in S, \; t \in T.$$

This shows that $\| \Gamma^k F - F^* \|_\infty$ is bounded by a constant multiple of β^k. When $k \to \infty$, we have $F^*(S) = \lim_{k\to\infty} (\Gamma^k F)(S_t), \; \forall S_t \in S$, as required.

The Theorem1.3 below characterized the optimal costs function F*, as well as optimal stationary policies. The result also give conditions under which value iteration converges to the optimal costs function F*. In the proof, we will often need to interchange expectation and limit in various relations. This interchange is valid under the assumption of the following theorem:

***Theorem* 1.2**

Let $P = \{p_1, p_2, ...\}$ be a probability distribution over $S = \{s_1, s_2, ...\}$. Let $\{h_T\}$ be a sequence of extended real-value functions on S such that for all $s_i \in S \; and \; T = 1,2,\Lambda$ $0 \leq h_T(s_i) \leq h_{T+1}(s_i).$ Let $h : S \to [0, \infty]$ be the limit function such that $h(s_i) = \lim_{T\to\infty} h_T(s_i).$

Then $\lim_{T\to\infty} \sum_{i=1}^{\infty} p_i h_T(s_i) = \sum_{i=1}^{\infty} p_i \lim_{T\to\infty} h_T(s_i) = \sum_{i=1}^{\infty} p_i h(s_i).$

Proof (see [2]).

***Theorem* 1.3**:

Suppose that the costs per stage $\sum_{i=1}^{n}\sum_{k=1}^{m} \varphi(i,k)$ satisfies

$$\sum_{i=1}^{n}\sum_{k=1}^{m} \varphi(i,k)(c(s)) \geq 0 \;\; for \; all \; (s,c) \in S \times C.$$

Then the following holds:

(i) Let F_T be the optimal costs function for the corresponding T-stage problem, then $F^* \leq (\Gamma F^*)(s)$ *for all* $s \in S$.

(ii) If $F:S\to(-\infty,\infty]$ satisfies $F \geq \Gamma F$ and either F is bounded below and $\beta < 1$ or $F \geq 0$, then $F \geq F^*$.

Proof

For any admissible policy $\pi = \{\pi_0, \pi_1,\ldots\}$, we consider the costs function $F^\pi(s_0)$ corresponding to π when the initial state is s_0, we have

$$F^\pi(s_0) = E\left[\sum_{i=1}^{n}\sum_{k=1}^{m}\varphi(i,k)(c(s)) + G^\pi(\pi_0(s))\right] \tag{1,10}$$

where, for all $s_1 \in S$,

$$G^\pi(s_1) = \lim_{T\to\infty} E\left[\sum_{t=1}^{T-1}\sum_{i=1}^{n}\sum_{k=1}^{m}\beta^t\varphi_t(i,k)(\pi_t(s_t))\right], t = 0,1,\Lambda$$

Thus, G'(s_1) is the costs from stage 1 to infinity using π when the initial state is s_1 We clearly have

$$G^\pi(s_1) \geq \beta F^*(s_1),\ \textit{for all}\ s_1 \in S.$$

Hence, from (1.10),

$$F^\pi(s_0) \geq E\left[\sum_{i=1}^{n}\sum_{k=1}^{m}\varphi(i,k)(\pi_0(s)) + \beta F^*(\pi_0(s))\right] \geq \min_\pi E\left[\sum_{i=1}^{n}\sum_{k=1}^{m}\varphi(i,k)(\pi(s)) + \beta F^*(\pi(s))\right]$$

Taking the minimum over all admissible policies, we obtain

$$\min_\pi F^\pi(s_0) = F^*(s_0) \geq \min_{c\in C(s)} E\left[\sum_{i=1}^{n}\sum_{k=1}^{m}\varphi(i,k)(c(s)) + \beta F^*(c(s))\right] = (\Gamma F^*)(s_0). \tag{1.11}$$

Next, we prove that the reverse inequality also holds. We are to proof that there exists π such that $\Gamma_\pi F^* = \Gamma F^*$. Since in general π need not exist, we introduce a positive sequence $\{\epsilon_i\}$ and we choose an admissible policy $\pi = \{\pi_0, \pi_1,\ldots\}$ such that

$$(\Gamma_{\pi_t} F^*)(s_0) \leq (\Gamma F^*)(s_0) + \in_t,\ s_0 \in S,\ t = 0,1,\Lambda$$

$\textit{provided}\ F^*(s_0) \geq 0\ \textit{for all}\ s_0$.

By using the inequality $\Gamma F^* \leq \Gamma F^*$. as obtain in (1.11), we obtain

$$(\Gamma_{\pi_t} F^*)(s_0) \leq F^*(s_0) + \in_t,\ s_0 \in S,\ t = 0,1,\ldots$$

Applying $\Gamma_{\pi_{t-1}}$ to both sides of this relation, we have

$$\begin{aligned}(\Gamma_{\pi_{t-1}}\Gamma_{\pi_t} F^*)(s_0) &\leq (\Gamma_{\pi_{t-1}} F^*)(s_0) + \beta\in_t,\ s_0 \in S,\ t = 0,1,\ldots\\ &\leq (\Gamma F^*)(s_0) + \in_{t-1} + \beta\in_t\\ &\leq F^*(s_0) + \in_{t-1} + \in_t\ .\end{aligned}$$

Continuing this process, we obtain

$$(\Gamma_{\pi_0}\Gamma_{\pi_1}\ldots\Gamma_{\pi_t})(s_0) \leq (\Gamma F^*)(s_0) + \sum_{i=1}^{t}\beta^i \in_i\ .$$

By taking the limit as $t \to \infty$ and noting that

$$F^*(s_0) \leq F^\pi(s_0) = \lim_{t\to\infty}(\Gamma_{\pi_0}\Gamma_{\pi_1}\ldots\Gamma_{\pi_t} F_0) \leq \lim_{t\to\infty}(\Gamma_{\pi_0}\Gamma_{\pi_1}\ldots\Gamma_{\pi_t} F^*)(s_0),$$

where F_0 is the zero function. It follows that

$$F^*(s_0) \leq F^\pi(s_0) = (\Gamma F^*)(s_0) + \sum_{i=0}^{\infty}\beta^i \in_i,\ s_0 \in S.$$

Since the sequence $\{\epsilon_i\}$ is arbitrary, we can take $\sum_{i=0}^{\infty} \beta^i \in_i,\ s_0 \in S$ as small as desired and we obtain

$$F^*(s_0) \le (\Gamma F^*)(s_0),\ \text{for all } s_0 \in S.$$

Combinning this with (1.11), we have

$$F^*(s_0) = \min_c E\left[\sum_{i=1}^{n}\sum_{k=1}^{m} \varphi(i,k)(c(s)) + \beta F^*(c(s))\right],\ s \in S.$$

or equivalently, $F^* = \Gamma F^*$.

(iv) Let z be a scalar such that $F(s) + z$ for all $s \epsilon S$ and $\beta \ge 1$, let $z = 0$.

For any sequence $\{\epsilon_t\}$ with $\epsilon_i > 0$, let $\pi = \{\pi_0, \pi_1, \ldots\}$, be an admissible policy such that, for every $s \epsilon S$ and t, we obtain

$$E\left[\sum_{i=1}^{n}\sum_{k=1}^{m} \varphi(i,k)(\pi_t(s)) + \beta F(\pi_t(s))\right] \le (\Gamma F)(s) + \in_t,\ t = 0,1,\ldots,\ s \in S, \tag{1.12}$$

such a policy exists since (TF) > $-\infty$ for all $s \epsilon S$. We have for any initial state $s_0 \epsilon S$,

$$F^*(s_0) = \max_{\pi} \lim_{T\to\infty} E\left[\sum_{t=0}^{T-1}\sum_{i=1}^{n}\sum_{k=1}^{m} \beta^t \varphi(i,k)(\pi_t(s_t))\right]$$

$$\le \max_{\pi} \lim_{T\to\infty} \inf E\left[\beta^T (F(s_T) + z) + \sum_{t=0}^{T-1}\sum_{i=1}^{n}\sum_{k=1}^{m} \beta^t \varphi_t(i,k)(\pi_t^i(s_t))\right]$$

$$\le \lim_{T\to\infty} \max_{\pi} \inf E\left[\beta^T (F(s_T) + z) + \sum_{t=0}^{T-1}\sum_{i=1}^{n}\sum_{k=1}^{m} \beta^t \varphi_t(i,k)(\pi_t^i(s_t))\right]$$

Using (1.12) and assumption $F \ge \Gamma F$, we obtain

$$E\left[\beta^T F(s_T) + \sum_{t=0}^{T-1}\sum_{i=1}^{n}\sum_{k=1}^{m} \beta^t \varphi_t(i,k)(\pi_t^i(s_t))\right] = E\left[\beta^T F(s_{T-1}, \pi_{T-1}(s_{T-1})) + \sum_{t=0}^{T-1}\sum_{i=1}^{n}\sum_{k=1}^{m} \beta^t \varphi_t(i,k)(s_t, \pi_t^i(s_t))\right]$$

$$\le E\left[\beta^{T-1} F(s_{T-1}) + \sum_{t=0}^{T-2}\sum_{i=1}^{n}\sum_{k=1}^{m} \beta^t \varphi_t(i,k)(s_t, \pi_t^i(s_t))\right] + \beta^{T-1} \in_{T-1}$$

$$\le E\left[\beta^{T-2} F(s_{T-2}) + \sum_{t=0}^{T-3}\sum_{i=1}^{n}\sum_{k=1}^{m} \beta^t \varphi_t(i,k)(s_t, \pi_t^i(s_t))\right] + \beta^{T-2} \in_{T-2} + \beta^{T-1} \in_{T-1} \le F(s_0) + \sum_{t=0}^{T-1} \beta^t \in_t .$$

Combinning these inequalities, we obtain

$$F^*(s_0) \le F(s_0) + \lim_{T\to\infty}\left(\beta^T z + \sum_{t=0}^{T-1} \beta^t \in_t\right).$$

Since the sequence $\{\epsilon_t\}$ is arbitrary (except for $\epsilon_i > 0$), we may select $\{\epsilon_t\}$ so that $\lim_{T\to\infty} \sum_{t=0}^{T-1} \beta^t \in_t$ is arbitrary close to zero, and the result follows.

The Theorem 1.4 below provides a measure of quality of policy π, the expected increased in the infinite-horizon discounted costs, conditioned on the initial state of the system being distributed according to a probability distribution w; i.e.,

$E_X[F^\pi(X) - F^*(X)] = \| F^\pi - F^* \|_{1,w}$.

We defined a measure $\vartheta_{\pi,w}$ over the state space associated with each policy π and probability distribution, *w* as

$$\vartheta^{Tr}{}_{\pi,w} = (1-\beta)w^{Tr}\sum_{t=0}^{\infty}\beta^t P_\pi^t,$$

Since

$\sum_{t=0}^{\infty}\beta^t P_\pi^t = (I-\beta P_\pi)^{-1}$, we have that $\vartheta^{Tr}{}_{\pi,w} = (1-\beta)w^{Tr}(I-\beta P_\pi)^{-1}$.

The measure $\vartheta^{Tr}{}_{\pi,w}$ captures the expected frequency of visit to each state when the system runs under policy π, conditioned on the initial state being distributed according to *w*.

***Theorem* 1.4**

Suppose F satisfied $F \le \Gamma F$. If ϑ is the policy defined as

$$\vartheta = \arg\max_{S_t} F(S_t),$$

then the following bound holds:

$$\left\|F_\vartheta - F^*\right\|_{1,v} \le \frac{\beta}{1-\beta}\left\|F-F^*\right\|_{1,w} \tag{1.13}$$

where *v* is an arbitrary probability on *S* and *w* is the probability distribution defined as $w = \vartheta_{\pi,v} = (1-\beta)(I-\beta P_\pi)^{-1}v$.

Proof

We show first that w is a probability distribution. Let *e* denote the unit vector for each entry. Since $e^{Tr}v = 1 \quad and\ e^{Tr}P_t^\pi = e^{Tr}$ for all *t*, (where *Tr* denotes transpose) we have

$$e^{Tr}w = e^{Tr}\vartheta_{\pi,v} = \sum_{s_t\in S}\vartheta_{\pi,v}(s_t) = (1-\beta)e^{Tr}(I-\beta P_\pi)^{-1}v = (1-\beta)\sum_{t=0}^{\infty}\beta^t e^{Tr}P_t^\pi v = (1-\beta)\sum_{t=0}^{\infty}\beta^t e^{Tr}v$$

$$= \sum_{t=0}^{\infty}\beta^t e^{Tr}v - \beta\sum_{t=0}^{\infty}\beta^t e^{Tr}P_t^\pi v = \sum_{t=0}^{\infty}\beta^t - \beta\sum_{t=0}^{\infty}\beta^t = 1.$$

Also, since $P_\pi \ge 0$ and $v \ge 0$ componentwise, $w \ge 0$. We begin by establishing that (1.13) holds by first showing that.

$$\| F_\pi - F^*\|_{1,w} \le \| F_\pi - \Gamma_\pi F\|_{1,w} + \|\Gamma_\pi F - F^*\|_{1,w} \le \|\Gamma_\pi F_\pi - \Gamma_\pi F\|_{1,w} + \|\Gamma_\pi F - F^*\|_{1,w} \tag{1.14}$$

$$\le \beta\| P_\pi(F_\pi - F)\|_{1,w} + \beta\| F - F^*\|_{1,w}$$

Since $F \le \Gamma F$ we have $F \le F^* \le F_\pi$. It then follows that

$$\| P_\pi(F_\pi - F)\|_{1,w} = \vartheta_{\pi,v}^{Tr}P_\pi(F_\pi - F) = w^{Tr}P_\pi(F_\pi - F) \tag{1.15}$$

Combining (1.12) and (1.15), we have

$$\| F_\pi - F^*\|_{1,w} \le \beta\vartheta_{\pi,v}^{Tr}P_\pi(F_\pi - F) + \beta\| F - F^*\|_{1,w} \le \beta w^{Tr}P_\pi(F_\pi - F) + \beta\| F - F^*\|_{1,w} \tag{1.16}$$

But,

$$\vartheta_{\pi,v}^{Tr}(F_\pi - F) \le \beta\vartheta_{\pi,v}^{Tr}P_\pi(F_\pi - F) + \beta\| F - F^*\|_{1,w} \tag{1.17}$$

It implies that $\vartheta_{\pi,v}^{Tr}(F_\pi - F) - \beta\vartheta_{\pi,v}^{Tr}P_\pi(F_\pi - F) \le \beta\| F - F^*\|_{1,w}$

$$\vartheta_{\pi,v}^{Tr}(I-\beta P_\pi)(F_\pi - F) \le \beta\| F - F^*\|_{1,w} \tag{1.18}$$

Hence, $(1-\beta)v^{Tr}(I-\beta P_\pi)^{-1}(I-\beta P_\pi)(F_\pi - F) \le \| \Gamma_\pi F - F \|_{1,w} \le \beta \| F - F^* \|_{1,w}$.

$$(1-\beta)v^{Tr}(F_\pi - F) \le \| \Gamma_\pi F - F \|_{1,w} \tag{1.19}$$

$(1-\beta)\| F_\pi - F \|_{1,v} \le \| \Gamma_\pi F - F \|_{1,w}$

Since we require that $F \le \Gamma F$, then by monotonicity of Γ, this implies $F^* \ge \Gamma^k F \ge F$, for all k.
By Lemma 1.1 and Theorem 1.1, we have that $\Gamma^k F \to F^*$ for any F. Also, the policy π is greedy with respect to F, we have that $F \le \Gamma F = \Gamma_\pi \le F^*$ so that

$$\| \Gamma_\pi F - F \|_{1,w} \le \beta \| F - F^* \|_{1,w}, \text{ and } \| F_\pi - F^* \|_{1,v} \le \| F_\pi - F \|_{1,v} \tag{1.20}$$

Combining (1.18), (1.19) and (1.20), we have

$$\| F_\pi - F^* \|_{1,v} \le \| F_\pi - F \|_{1,v} \le \frac{1}{1-\beta} \| \Gamma_\pi F - F \|_{1,w} \le \frac{\beta}{1-\beta} \| F - F^* \|_{1,w} .$$

$$\therefore \qquad \frac{\beta\eta'}{1-\beta} \le \| F_\pi - F^* \|_{1,w} \le \frac{\beta\eta''}{1-\beta},$$

where $\eta' = \min_{S_t}[(\Gamma_\pi F)(S_t) - F(S_t)]$ and $\eta'' = \max_{S_t}[(\Gamma_\pi F)(S_t) - F(S_t)]$.

We now set

$$\psi' = \frac{\beta\eta'}{1-\beta} \quad and \quad \psi'' = \frac{\beta\eta''}{1-\beta},$$

which are the error bounds associated with our problem.

2.1 Computational work

The transition matrices corresponding to the control policies π^1, π^2, π^3, π^4, π^5, π^6, π^7 are respectively presented below: (For the transition diagram, [8]

$$P(\pi^1) = \begin{bmatrix} 1/4 & 9/40 & 1/5 & 3/20 & 1/10 & 1/20 & 1/40 \\ 1/4 & 9/40 & 1/5 & 3/20 & 1/10 & 1/20 & 1/40 \\ 1/4 & 9/40 & 1/5 & 3/20 & 1/10 & 1/20 & 1/40 \\ 1/4 & 9/40 & 1/5 & 3/20 & 1/10 & 1/20 & 1/40 \\ 1/4 & 9/40 & 1/5 & 3/20 & 1/10 & 1/20 & 1/40 \\ 1/4 & 9/40 & 1/5 & 3/20 & 1/10 & 1/20 & 1/40 \\ 1/4 & 9/40 & 1/5 & 3/20 & 1/10 & 1/20 & 1/40 \end{bmatrix} \quad P(\pi^4) = \begin{bmatrix} 1/10 & 1/4 & 9/40 & 1/20 & 1/40 & 1/5 & 3/20 \\ 1/10 & 1/4 & 9/40 & 1/20 & 1/40 & 1/5 & 3/20 \\ 1/10 & 1/4 & 9/40 & 1/20 & 1/40 & 1/5 & 3/20 \\ 1/10 & 1/4 & 9/40 & 1/20 & 1/40 & 1/5 & 3/20 \\ 1/10 & 1/4 & 9/40 & 1/20 & 1/40 & 1/5 & 3/20 \\ 1/10 & 1/4 & 9/40 & 1/20 & 1/40 & 1/5 & 3/20 \\ 1/10 & 1/4 & 9/40 & 1/20 & 1/40 & 1/5 & 3/20 \end{bmatrix} \quad P(\pi^7) = \begin{bmatrix} 1/5 & 1/40 & 3/20 & 1/10 & 1/20 & 1/4 & 9/40 \\ 1/5 & 1/40 & 3/20 & 1/10 & 1/20 & 1/4 & 9/40 \\ 1/5 & 1/40 & 3/20 & 1/10 & 1/20 & 1/4 & 9/40 \\ 1/5 & 1/40 & 3/20 & 1/10 & 1/20 & 1/4 & 9/40 \\ 1/5 & 1/40 & 3/20 & 1/10 & 1/20 & 1/4 & 9/40 \\ 1/5 & 1/40 & 3/20 & 1/10 & 1/20 & 1/4 & 9/40 \\ 1/5 & 1/40 & 3/20 & 1/10 & 1/20 & 1/4 & 9/40 \end{bmatrix}$$

$$, P(\pi^2) = \begin{bmatrix} 1/40 & 1/20 & 1/10 & 1/5 & 3/20 & 9/40 & 1/4 \\ 1/40 & 1/20 & 1/10 & 1/5 & 3/20 & 9/40 & 1/4 \\ 1/40 & 1/20 & 1/10 & 1/5 & 3/20 & 9/40 & 1/4 \\ 1/40 & 1/20 & 1/10 & 1/5 & 3/20 & 9/40 & 1/4 \\ 1/40 & 1/20 & 1/10 & 1/5 & 3/20 & 9/40 & 1/4 \\ 1/40 & 1/20 & 1/10 & 1/5 & 3/20 & 9/40 & 1/4 \\ 1/40 & 1/20 & 1/10 & 1/5 & 3/20 & 9/40 & 1/4 \end{bmatrix} \quad P(\pi^5) = \begin{bmatrix} 9/40 & 3/20 & 1/4 & 1/40 & 1/5 & 1/10 & 1/20 \\ 9/40 & 3/20 & 1/4 & 1/40 & 1/5 & 1/10 & 1/20 \\ 9/40 & 3/20 & 1/4 & 1/40 & 1/5 & 1/10 & 1/20 \\ 9/40 & 3/20 & 1/4 & 1/40 & 1/5 & 1/10 & 1/20 \\ 9/40 & 3/20 & 1/4 & 1/40 & 1/5 & 1/10 & 1/20 \\ 9/40 & 3/20 & 1/4 & 1/40 & 1/5 & 1/10 & 1/20 \\ 9/40 & 3/20 & 1/4 & 1/20 & 1/5 & 1/10 & 1/20 \end{bmatrix}$$

$$P(\pi^3) = \begin{bmatrix} 1/20 & 1/10 & 1/40 & 1/4 & 9/40 & 3/20 & 1/5 \\ 1/20 & 1/10 & 1/40 & 1/4 & 9/40 & 3/20 & 1/5 \\ 1/20 & 1/10 & 1/40 & 1/4 & 9/40 & 3/20 & 1/5 \\ 1/20 & 1/10 & 1/40 & 1/4 & 9/40 & 3/20 & 1/5 \\ 1/20 & 1/10 & 1/40 & 1/4 & 9/40 & 3/20 & 1/5 \\ 1/20 & 1/10 & 1/40 & 1/4 & 9/40 & 3/20 & 1/5 \\ 1/20 & 1/10 & 1/40 & 1/4 & 9/40 & 3/20 & 1/5 \end{bmatrix} \quad P(\pi^6) = \begin{bmatrix} 3/20 & 1/5 & 1/20 & 9/40 & 1/4 & 1/40 & 1/10 \\ 3/20 & 1/5 & 1/20 & 9/40 & 1/4 & 1/40 & 1/10 \\ 3/20 & 1/5 & 1/20 & 9/40 & 1/4 & 1/40 & 1/10 \\ 3/20 & 1/5 & 1/20 & 9/40 & 1/4 & 1/40 & 1/10 \\ 3/20 & 1/5 & 1/20 & 9/40 & 1/4 & 1/40 & 1/10 \\ 3/20 & 1/5 & 1/20 & 9/40 & 1/4 & 1/40 & 1/10 \\ 3/20 & 1/5 & 1/20 & 9/40 & 1/4 & 1/40 & 1/10 \end{bmatrix}$$

Table 2.1 shows the transition costs of the operation.

Table 2.1: Transition Costs

r	$\varphi(S_1,\pi^r)$ $(Naira)$ $\times 100$	$\varphi(S_2,\pi^r)$ $(Naira)$ $\times 100$	$\varphi(S_3,\pi^r)$ $(Naira)$ $\times 100$	$\varphi(S_4,\pi^r)$ $(Naira)$ $\times 100$	$\varphi(S_5,\pi^r)$ $(Naira)$ $\times 100$	$\varphi(S_6,\pi^r)$ $(Naira)$ $\times 100$	$\varphi(S_7,\pi^r)$ $(Naira)$ $\times 100$
1	15000	12000	7900	6900	12820	10550	11200
2	5000	8700	10000	18000	10120	11250	10720
3	11000	13500	6900	9800	7550	9610	8005
4	8500	6600	12200	10200	7200	8840	8665
5	8000	6900	8020	8200	7050	7950	7700
6	9500	7500	10300	7900	6600	6750	7780
7	9100	8000	9900	7630	6500	5980	10990

We take 5% to be the value of our discount factor, represents the percentage amount of products that are damaged in transit, MATLAB was used to solve the problem. The results are shown in the Tables below:

Table 2.2: Value iteration without error bounds

k	$(\Gamma^k F)(S_1)$ (Naira) ×100	$(\Gamma^k F)(S_2)$ (Naira) ×100	$(\Gamma^k F)(S_2)$ (Naira)×100	$(\Gamma^k F)(S_4)$ (Naira)×100	$(\Gamma^k F)(S_5)$ (Naira)×100	$(\Gamma^k F)(S_6)$ (Naira)×100	$(\Gamma^k F)(S_7)$ (Naira) ×100
0	0	0	0	0	0	0	0
1	5000.00	6600.00	6900.00	6900.00	6500.00	5980.00	7700.00
2	5338.50	6928.00	7235.40	7214.60	6822.10	6302.10	8014.80
3	5354.60	6944.40	7251.40	7231.00	6838.40	6318.40	8031.20
4	5355.40	6945.20	7252.20	7231.80	6839.20	6319.20	8032.10
∞	5355.40	6945.20	7252.20	7239.00	6839.20	6319.20	8032.10

Table2.3: Value iteration with minimum error bounds

k	$(\Gamma^k F)(S_1)$ ψ' (Naira) ×100	$(\Gamma^k F)(S_2)$ ψ' (Naira) ×100	$(\Gamma^k F)(S_3)$ ψ' (Naira) ×100	$(\Gamma^k F)(S_4)$ ψ' (Naira) ×100	$(\Gamma^k F)(S_5)$ ψ' (Naira) ×100	$(\Gamma^k F)(S_6)$ ψ' (Naira) ×100	$(\Gamma^k F)(S_7)$ ψ' (Naira) ×100
0	_	_	_	_	_	_	_
1	5263.20	6863.20	7163.20	7163.20	6763.20	6243.20	7963.20
2	5355.10	6944.60	7251.90	7231.10	6838.70	6318.70	8031.30
∞	5355.40	6945.20	7252.20	7231.90	6839.20	6319.20	8032.10

Table 2.4: Value iteration with maximum error bounds

k	$(\Gamma^k F)(S_1)$ ψ' ' (Naira) × 100	$(\Gamma^k F)(S_2)$ ψ'' (Naira) ×100	$(\Gamma^k F)(S_3)$ ψ'' (Naira) ×100	$(\Gamma^k F)(S_4)$ ψ'' (Naira) ×100	$(\Gamma^k F)(S_5)$ ψ'' (Naira) ×100	$(\Gamma^k F)(S_6)$ ψ'' (Naira) ×100	$(\Gamma^k F)(S_7)$ ψ'' (Naira) ×100
0	_	_	_	_	_	_	_
1	5405.30	7005.30	7305.30	7305.30	6905.30	6385.30	8105.30
2	5356.30	6945.90	7253.20	7232.40	6839.90	6319.90	8032.60
3	5355.40	6945.30	7252.20	7231.90	6839.20	6319.20	8032.10
∞	5355.40	6945.20	7252.20	7231.90	6839.20	6319.20	8032.10

3.0 Discussion

Table 2.1 contains the transition cost of distributing the products from the production centres to the markets. The values in Table 2.1 and the transition probabilities above are obtained from past records of the company. These values are multiples of 100, i.e.,

$$\varphi(S_1,\pi^1) = 15000\times 100,\ \varphi(S_1,\pi^2) = 12000\times 100,$$
$$\varphi(S_1,\pi^3) = 7900\times 100, \ldots\ ,\varphi(S_7,\pi^7) = 10990\times 100.$$

When we solve our problem using the information in Table 2.1 and the transition probabilties of the distribution, we obtained the results in Table 2.2 to 2.4. The results in Table 2.2 are obtained by using value iteration without error bounds. As the period, k become larger and larger we obtained the

following results:

$$F^*(S) = \lim_{k\to\infty} \Gamma^k F(S) = (5355.40,\ 6945.20,\ 7252.20,\ 7231.90,\ 6839.20,\ 6319.20,\ 8032.10),\tag{3.1}$$

which is the optimal solutions of the system when error bounds are not included. Note that the vector in (2.1), the actual values are obtained by multipling F*(S) by 100. Hence, minimum costs of distributing the products from the production centre 1 to 7 to the markets are 535,540; 694,520; 725,220; 723,190; 683,920; 631,920; 803,210 in naira respectively.

Again, Table 2.3 contained the results of the operation when minimum error bounds are included. As the period, *k* become larger and larger we obtained the following results:

$$\lim_{k\to\infty} \Gamma^k F(S) + \psi'_k = (5355.40, 6945.20, 7252.20, 7231.90, 6839.20, 6319.20, 8032.10) \tag{3.2}$$

Also, the actual values are obtained by multipling $\lim_{k\to\infty} \Gamma^k F(S) + \psi'_k$ by 100. Table 2.4 contained the results of the operation when maximum error bounds are included. As the period, k become larger and larger we obtained the following results:

$$\lim_{k\to\infty} (\Gamma^k F)(S) + \psi''_k = (5355.40, 6945.20, 7252.20, 7231.90, 6839.20, 6319.20, 8032.10) \tag{3.3}$$

Also, the actual values are obtained by multiplying $\lim_{k\to\infty} \Gamma^k F(S) + \psi''_k$ by 100.

We observed that the vector values in (2.1), (2.2) and (2.3) are the same. Hence, we write

$$F^*(S) = \lim_{k\to\infty} \Gamma^k F(S) + \psi'_k = \lim_{k\to\infty} (\Gamma^k F)(S) + \psi''_k = \lim_{k\to\infty} (\Gamma^k F)(S) = .$$
$$(5355.40,\ 6945.20,\ 7252.20,\ 7231.90,\ 6839.20,\ 6319.20,\ 8032.10) \tag{2.4}$$

Also, the actual values are obtained by multipling $F^*(S)$ by 100. This is only true at k equals infinity.

4.0 Conclusion

The results show that the minimum costs of distributing the products from production centre 1 to 7 to the markets be 535,540; 694,520; 725,220; 723,190; 683,920; 631,920; 803,210 in naira, respectively. We discovered that the value iteration without and with error bounds converges to the same vector values as presented in equation (2.4) above. This is true only at infinite stage. We also observed that the first production centre of the production centres has the minimum costs.

Reference

[1] Nkeki, C. I. (2006). On a dynamic programming algorithm for resource allocation problems. M.Sc. thesis, Department of Mathematics, University of Ibadan, Ibadan, Nigeria.

[2] Nkeki, C.I. and Nwozo, C.R. (2009). The use of value iteration to minimize the costs of shipping different goods. Accepted by International Journal of Natural and Applied Sciences.

[3] Powell, W.B. and Topaloglu, H. (2003). Stochastic programming in transportation and logistics, in A. Ruszczynski and A. Shapiro, eds. Handbook in Operation Research and Management Science, Volume on Stochastic Programming; Elsevier, Amsterdam, pp. 555-635, 8.

[4] Powell, B.W. and Van Roy, B. (2003). "Approximate dynamic programming for high-dimensional resource allocation problems", Operations Research, pp. 1-27.

[5] Powell, W.B. (2004). "Approximate Dynamic programming for asset management", Princeton University, Princeton.

[6] Rust, J. (1994). Numerical dynamic programming in Economics, Yale University.

[7] Topaloglu, H. & Kunnumkal, S. (2006). Approximate dynamic programming methods for an inventory allocation problem under uncertainty. Cornell University, Ithaca, U.S.A.

[8] Van Roy, B., Bertsekas, D. P., Lee, Y. and Tsitsiklis, J. N. (1997). A neuro-dynamic programming approach to retailer inventory management, in 'Proceedings of the IEEE Conference on Decision and control', 15.