

REDUCTION OF STATE SPACE IN DYNAMIC PROGRAMMING
WITH INTEGRATED FORECASTING

Dieter Bartmann

Technische Universität München, München, F.R.G.

1. Introduction. In many decision problems such as inventory control, cash management or production smoothing the uncertainty of demand is a central feature. The standard assumption is that demand in different periods is independently and identically distributed. But this assumption is rarely valid in practice, in particular when forecasts of demands are being made. If planning covers more than one period then all calculations are influenced by expectations regarding future data. This in turn requires an anticipation of future movements of the underlying demand process. Commercial suppliers of software handle this problem in a three stage approach: lotsize (or cash shipment or production rate), security reserves and forecast. By this separation into three parts, the solution obtained is only suboptimal.

The optimal solution requires reformulation of the basic model. This will be performed on Beckmann's model of production smoothing (Beckmann (1961)). In general incorporating forecasts into Bellman's principle of optimality leads to an increase of the state space. If forecasts are made by exogenous random variables or by an autoregressive scheme then, however, the state space can be substantially reduced, as will be shown in this paper.

2. Beckmann's model of production smoothing. The following problem of inventory - production control is regarded. Given the actual stock level we have to control the production rate within the range of minimal unit control. The production rate will be fixed at the beginning of each planning period. The planning horizon T will cover more than one period

$(T \gg 1)$.

Let

- u demand in the present period. u is a random variable with distribution function $\Psi(u)$
- x present production rate (prior to the decision)
- y present stock level
- d change of production rate (decision variable)
- f cost function of a single period (set-up costs, inventory- and production costs)
- n number of periods up to the planning horizon
- v_n expected minimal value of all future costs n periods ahead
- β discount factor, $0 < \beta < 1$

The model is defined recursively in v_n . v_n depends on the present state (x, y) of the system (before the decision). The new production rate $x + d$ is to minimize with respect to all possible realizations of future demand. The Dynamic Program may then be formulated in terms of the two state variables

$$v_n(x, y) = \min_d \left\{ \int f(x, y, d, u) d\Psi(u) + \beta \int v_{n-1}(x + d, y + x + d - u) d\Psi(u) \right\}$$

$n = 1, 2, \dots, T \quad (1)$

$$v_0(x, y) = V_0 \dots$$

This defines a stationary decision process. We will now incorporate a forecasting procedure.

3. Forecast with exogenous variables. In practice a great variety of forecasting techniques is used. We will consider only those methods that can be formulated recursively, otherwise an integration into Bellman's functional equalities is excluded. Sometimes demand is closely correlated to exogenous influences such as gross national product, export demand, indicators of the business cycle etc. In these cases it is useful to treat the exogenous factors as random disturbances. They generate a conditional distribution of the demand

variable. We assume that all exogenous influences are aggregated to a scalar z (scenario). It is used as a forecast. Let

z actual forecast

w new forecast, concerning the next period

$\Phi(w)$ probability distribution function of w with density $\phi(w)$

From the view point of the decision maker the forecasts form a series of identically and independently distributed random variables. The principle of optimality is

$$v_n(x, y, z) = \min_{\bar{d}} \left\{ \int f(x, y, d, u) d\Psi(u|z) + \right. \\ \left. + \beta \iint v_{n-1}(x + d, y + x + d - u, w) d\Psi(u|z) d\Phi(w) \right\} \quad (2)$$

$n = 1, 2, \dots, T$

$$v_0(x, y, z) = V_0.$$

4. Reduction of state space in the exogenous case. In the sequel it will be shown how the functional equations (2) can be formulated in only two state variables. To obtain a clearer presentation we introduce an index variable t to denote the time periode the variables are referring to. For example x_t is now the production rate at the beginning of period t before decision d_t is made.

We state the following assumptions

A1. $\Psi_t(u_t | z_t) = P_t(u_t - z_t) = P_t(\epsilon_t)$ where ϵ_t is iid.

A2. The stock holding costs depend only on the stock level at the end of the period:

$$f = f(x_t, d_t, y_t + x_t + d_t - u_t) \quad (3)$$

Both assumptions are not very restrictive for real applications. For instance A1 is satisfied if the demand process is assumed to be of the

form $u_t = z_t + \varepsilon_t$, $\varepsilon_t \sim N(0, \sigma)$ for all t . Thus $z_t = E(u_t)$. Assumption A2 is standard in inventory theory.

From A2 we conclude that

$$f(x_t, y_t, d_t, u_t) = f(x_t, y_t - c, d_t, u_t - c) \quad \text{f.a. } c \in \mathbb{R} \quad (4)$$

We will now map the two-dimensional subspace $\mathbb{Y} \times \mathbb{Z}$ onto a one-dimensional space \mathbb{R} by

$$r_t = r_t(y_t, z_t) := y_t - z_t \quad (5)$$

r may be regarded as a "net inventory level" that is stock diminished by expected demand. Then the Dynamic Program (2) becomes

$$\begin{aligned} v_t(x_t, y_t, z_t) &= \min_{d_t} \left\{ \int f(x_t, y_t, d_t, u_t) d\psi_t(u_t | z_t) + \right. \\ &\quad \left. + \beta \iint v_{t+1}(x_{t+1}, y_t + x_{t+1} - u_t, w) d\psi_t(u_t | z_t) d\Phi(w) \right\} = \\ &\stackrel{(A.1)}{=} \min_{d_t} \left\{ \int f(x_t, y_t, d_t, z_t + \varepsilon_t) dP_t(\varepsilon_t) + \right. \\ &\quad \left. + \beta \iint v_{t+1}(x_{t+1}, y_t - z_t - \varepsilon_t + x_{t+1}, w) dP_t(\varepsilon_t) d\Phi(w) \right\} = \\ &\stackrel{(A.2)}{=} \min_{d_t} \left\{ \int f(x_t, \underbrace{y_t - z_t}_{r_t}, d_t, \varepsilon_t) dP_t(\varepsilon_t) + \right. \\ &\quad \left. + \beta \iint v_{t+1}(x_{t+1}, \underbrace{y_t - z_t - \varepsilon_t + x_{t+1}}_{r_{t+1}}, w) dP_t(\varepsilon_t) d\Phi(w) \right\} \end{aligned} \quad (6)$$

Thus

$$\begin{aligned} v_t(x_t, r_t) &= \min_{d_t} \left\{ \int f(x_t, r_t, d_t, \varepsilon_t) dP_t(\varepsilon_t) + \right. \\ &\quad \left. + \beta \iint v_{t+1}(x_{t+1}, r_{t+1}) dP_t(\varepsilon_t) d\Phi(w) \right\} \end{aligned} \quad (7)$$

with the transition laws

$$x_{t+1} = x_t + d_t \quad (8)$$

$$r_{t+1} = r_t + x_{t+1} - \varepsilon_t - w \quad (9)$$

and the boundary condition

$$v_T = V_t \quad (\text{finite horizon}) \quad (10)$$

In case of infinite horizon it will be more suitable to return to the notation with index n . So

$$v_0 = V_0 \quad (\text{infinite horizon}) \quad (11)$$

We summarize the results obtained so far in

Lemma 1: Given an optimization problem defined by the Dynamic Programm (2). Using the map $r_t(y_t, z_t) = y_t - z_t$ the Dynamic Programm can be reformulated (7) in only two state variables without loss of generality. Thus a reduction of the original state space $X \times Y \times Z$ to the state space $X \times R$, R real subspace, is achieved.

As it is pointed out in Bartmann (1983) the applied reduction technique (5) preserves the structure of the optimal policy. In this way structural properties of the optimal policy in the two-state Dynamic Program of the model without forecasting can easily be transferred to the model with forecasting.

5. Forecasting as an autoregressive scheme. The reduction technique developed so far can also be applied to models where the demand process is assumed to be of the type of a first order autoregressive scheme. Let for ease $E(u_t) = 0$. Then

$$u_t = \alpha u_{t-1} + \varepsilon_t \quad (12)$$

where ε_t is iid with probability distribution function $P_t(\varepsilon_t)$ and

$|\alpha| < 1$, $\alpha \in \mathbb{R}$. Using the actual observation u_t for forecasting yields a variance $\text{var}(u_{t+1}|u_t) = \sigma_\varepsilon^2$.

$$\sigma_u^2 = \frac{1}{1 - \alpha^2} \sigma_\varepsilon^2 > \sigma_\varepsilon^2, \quad \text{if } 0 < |\alpha| < 1.$$

With the assumption $|\alpha| < 1$ the demand process becomes stationary. This justifies an infinite horizon model. In the forecasting formula derived from (12)

$$\hat{u}_t = \alpha u_{t-1} \quad (13)$$

the most recent observation u_{t-1} is a sufficient statistic. We will denote it as u_1 (demand one periode ago). Then the Dynamic Program takes the form

$$\begin{aligned} v_n(x, y, u_1) = \min_d \left\{ \int f(x, y, d, u) d\Psi(u|u_1) + \right. \\ \left. + \beta \int v_{n-1}(x+d, y+x+d-u, \alpha u_1) d\Psi(u|u_1) \right\} \\ n = 1, 2, \dots \end{aligned} \quad (14)$$

$$v_0(x, y, u_1) = V_0.$$

6. Reduction of state space in the autoregressive case, when the optimal policy is myopic. Under a special demand and cost structure the optimal decision rule is myopic (cf. Edgeworth (1883), Lenz (1974), Johnson/Thompson (1975), Schneider (1979)). It therefore can be derived by minimizing the one period costs v_1 .

We will show that the application of the above reduction technique will lead to a reduction of the state space $X \times Y \times U$ by one full dimension here too. From (12) we get

$$\Psi(u|u_1) = P(u - u_1) = P(\varepsilon) \quad (15)$$

and

$$\begin{aligned}
 v_1(x, y, u_1) &= \min_d \left\{ \int f(x, y, d, u) d\Psi(u | u_1) \right\} = \\
 &= \min_d \left\{ \int f(x, y, d, \alpha u_1 + \varepsilon) dP(\varepsilon) \right\} = \\
 &\stackrel{(A.2)}{=} \min_d \left\{ \int f(x, y - \underbrace{\alpha u_1}_r, d, \varepsilon) dP(\varepsilon) \right\}
 \end{aligned}
 \tag{16}$$

thus

$$v_1(x, r) = \min_d \left\{ \int f(x, r, d, \varepsilon) dP(\varepsilon) \right\}
 \tag{17}$$

with the net stock level

$$r = y - \alpha u_1 .
 \tag{18}$$

When we construct a computer code an additional consideration is required. The space $X \times Y \times U$ is then discrete and has a finite range in each component

$$\begin{aligned}
 x^- &\leq x \leq x^+ ; \\
 y^- &\leq y \leq y^+ ; \\
 u_1^- &\leq u_1 \leq u_1^+ .
 \end{aligned}
 \tag{19}$$

We denote the ranges by

$$\begin{aligned}
 X &= x^+ - x^- + 1 ; \\
 Y &= y^+ - y^- + 1 ; \\
 U &= u_1^+ - u_1^- + 1 ; \\
 R &= r^+ - r^- + 1 .
 \end{aligned}
 \tag{20}$$

From (18) we conclude $R > Y$. The reduction will only pay if

$$R < Y \cdot U .
 \tag{21}$$

Disregarding rounding errors

$$R = Y + |\alpha|U .
 \tag{22}$$

Therefore

$$R > Y \cdot U \quad \text{for } Y, U \geq 2 .$$

The saving is the greater the weaker the autocorrelation. But this effect is partly compensated by an increasing σ_ϵ^2 when $|a|$ decreases.

7. Reduction of the state space when the optimal policy is not myopic.
 In this case we can achieve a reduction of the state space too but not at the same magnitude as before.

Lemma 2: Given an optimization problem defined by the Dynamic Program (14) with arbitrary planning horizon. The state space lattice $X \times Y \times U$ can be reduced if

$$\frac{Y}{U} > \frac{|a|^3}{1 - a^2} , \quad |a| < 1 . \tag{23}$$

Proof: Let $r := y - \alpha u_1$ and $r_{neu} = y_{neu} - \alpha u$

$$\begin{aligned} v_n(x, \underbrace{y, u_1}_r) &= \min_d \left\{ \int f(x, y, d, u) d\Psi(u|u_1) + \right. \\ &\quad \left. + \beta \int v_{n-1}(x+d, y+x+d-u, \alpha u_1) d\Psi(u|u_1) \right\} = \\ &\stackrel{(16)}{=} \min_d \left\{ \int f(x, r, d, \epsilon) dP(\epsilon) + \right. \\ &\quad \left. + \beta \int v_{n-1}(x+d, \underbrace{x+d+y-\alpha u_1}_r, \underbrace{-\epsilon, \alpha u_1 + \epsilon}_{y_{neu}}) dP(\epsilon) \right\} = \\ &= \min_d \left\{ \int f(x, r, d, \epsilon) dP(\epsilon) + \right. \\ &\quad \left. + \beta \int v_{n-1}(x+d, \underbrace{x+d+r-\epsilon-\alpha^2 u_1 - \alpha \epsilon}_{r_{neu}}) dP(\epsilon) \right\} = \\ &= v_n(x, r, \alpha^2 u_1) . \end{aligned} \tag{24}$$

Now u_1 appears only diminished by the factor $\alpha^2 < 1$. Instead of u_1 it is sufficient to keep in mind only $\alpha^2 u_1$. This reduces the range U to $\alpha^2 U$. We reformulate the Dynamic Program in the state variables x, r and $\alpha^2 u_1$

$$v_n(x, r, \alpha^2 u_1) = \min_d \left\{ \int f(x, r, d, \epsilon) dP(\epsilon) + \beta \int v_{n-1}(x+d, r_{neu}, \underbrace{\alpha^2(\alpha u_1 + \epsilon)}_{\alpha^2 u}) dP(\epsilon) \right\} \quad (25)$$

Here the state space $X \times R \times (\alpha^2 U)$ has $X \cdot R \cdot \alpha^2 \cdot U$ lattice points (rounding errors excluded). As $R = Y + |\alpha| \cdot U$ we complete the proof

$$X \cdot R \cdot \alpha^2 U < X \cdot Y \cdot U \iff \frac{Y}{U} > \frac{|\alpha|^3}{1 - \alpha^2} \quad \text{q.e.d.}$$

If the set up costs are high compared with the proportional stock-holding costs then $Y > U$. Then from (23) we immediately conclude that a reduction of the state space is achieved as soon as $|\alpha|$ falls below 0.75.

For numerical reasons (unavoidable truncation errors at the boundary of the state space) the range Y should be at least twice the range U . Thus in practice the reduction method is favourable as soon as $|\alpha| < 0.84$.

Lemma 3: In the Dynamic Programm (14) a reduction of the state space $X \times Y \times U$ to $X \times R$ is possible if the occuring single term $\alpha^2 u_1$ will be approximated by $\alpha^2 \mu$, $\mu = E(u)$.

Proof: Now we have to consider the long run expectation μ of demand explicetely. Instead of (12) we get

$$u - \mu = \alpha(u_{t-1} - \mu) + \epsilon_t \quad (26)$$

an r is defined as

$$r = y - \alpha(u_1 - \mu) - \mu \quad (27)$$

With this we can formulate the Dynamic Programm (24). If the term $\alpha^2 u_1$ at the right hand side of (24) is replaced by $\alpha^2 \mu$ than it is no longer necessary to keep it in mind. Thus

$$v_n(x, r) = \min_d \left\{ \int f(x, r, d, \epsilon) dP(\epsilon) + \right. \\ \left. + \beta \int v_{n-1}(x+d, x+d+r - \alpha^2 \mu - (1+\alpha)\epsilon) dP(\epsilon) \right\} \quad (28)$$

q.e.d.

It should be noticed that the approximation $\alpha^2 \mu \approx \alpha^2 u_1$ is only performed once. All terms appearing linearly in μ are not approximated. They are incorporated into the variable r precisely.

Last we will demonstrate the power of the presented reduction technique.

Table: number of states dependent on α (with fixed $U = Y/2$ and $Y = 100$)

α	number of states		
	original DP $U \times U$	reduced DP $R \times (\alpha^2 U)$	reduced and approximated DP R
0.2	5 000	220	110
0.4	5 000	960	120
0.6	5 000	2 340	130
0.8	5 000	4 480	140

8. Summary. In this paper we considered the production - inventory model of the AHM type. We extended it to the case where future movements of demand are being forecasted. For two concrete situations

- forecasting by means of exogenous variables
- forecasting as an autoregressive scheme

the Dynamic Program incorporating forecasting has been formulated. Since forecasts are functions of observations in the present and past an additional state variable is required to store the sufficient statistics. This immediately leads to severe numerical complications. The main topic of the paper is to show how to overcome Bellman's "curse of dimensionality" by means of a special reduction technique. When forecasts are made with exogenous random variables the state space can be reduced by one full dimension without loss of any information. The same result is derived in case of forecasting by an autoregressive scheme if the optimal policy is myopic. If not then a reduction is also possible but with a smaller saving (estimates are given). Using a partial approximation the reduction is enhanced to save one full dimension here too. As a byproduct the reduction technique reveals to be useful for some results concerning the structure of the optimal policies when forecasts are being incorporated. But this will be investigated elsewhere.

References

- Bartmann, D.: Dynamische Optimierung mit Prognose. Habilitationsschrift, TU München, 1983.
- Beckmann, M.J.: Production Smoothing and Inventory Control. Operations Research 9 (1961) pp 453 - 467.
- Bellman, R.: Dynamic Programming. Princeton University Press, Princeton N.J., 1957.
- Edgeworth, F.Y.: The Mathematical Theory of Banking. Journal of the Royal Statistic Society 51 (1888), pp. 120 ff.
- Johnson, G.D. & Thompson, H.E.: Optimality of Myopic Inventory Policies for Certain Dependent Demand Processes. Management Science 21 (1975), pp. 1303 - 1307.
- Lenz, H.-J.: Zeitreihenanalyse und Prognoseverfahren zur Optimierung der Kassenhaltung bei autokorrelierten Zahlungsprozessen. Berlin Verlag, Berlin, 1974.
- Schneider, H.: Servicegrad in Lagerhaltungsmodellen. Marchal und Matzenbacher, Berlin, 1979.