Quadratic Criterion / Linear Dynamics

Consider a (finite) N-stage **non-stationary** Markov decision process with

- ***continuous*** state and decision variables,
- a convex ***quadratic*** cost at each stage (or concave return, if maximizing), and
- ***linear*** relationship between the future state and the current state & decisions.

QC/LD Model

©D.L.Bricker, 2001
Dept of Industrial Engineering
The University of Iowa

In the most general case, if x_n is the vector of *state* variables and y_n the vector of *action* or decision variables, the *cost* at stage n is the quadratic function

$$c_n\left(x_n, y_n\right) = x_n A_n x_n + x_n B_n y_n + y_n C_n y_n + d_n x_n + e_n y_n + f_n$$

and the *transition* at stage n is defined by the linear function

$$x_{n+1} = G_n x_n + H_n y_n + \xi_n$$

where $\xi_1, \xi_2, \cdots \xi_N$ are independent *random* vectors such that

$$\left|E\left(\xi_{ni}\right)\right| < \infty \quad \text{and} \quad \left|E\left(\xi_{ni}\xi_{nj}\right)\right| < \infty$$

for all i, j, and n. (Here, ξ_{ni} denotes the ith element of the vector ξ_n.)

Note: independence of the random vectors means that the process is memoryless, or Markovian.

ⓘ See **One-Dimensional QC/LD Problem**

Suppose for each $n = 1, 2, \ldots N$, the matrix A_n is **positive definite**, the matrix $B_n = 0$, and the matrix C_n is **positive semidefinite** ($\Rightarrow$ *convexity*).

Then the ***optimal value*** function is ***quadratic*** *of the form*

$$f_n\left(x\right) = xP_n x + q_n x + r_n$$

where P_n is positive semidefinite, (so $f_n\left(x\right)$ is convex) and

the optimal ***decision rule*** is ***linear***,

$$y_n = W_n x_n + v_n$$

for $n=1, 2, \ldots N$.

The closed-form expressions for P_n, q_n, r_n, W_n and v_n are quite messy!

See the section on the one-dimensional QC/LD problem for a derivation of a special case.

Backward Recursive Computation of Arrays P, q, r, W, and v

Recursive **Computation**

Let $P_{N+1} = A_{N+1}, q_{N+1} = d_{N+1}$, and $r_{N+1} = f_{N+1}$, and

for $n = N, N-1, \ldots 2, 1$: compute

$$\begin{cases} S_n = C_n + H_n^t P_{n+1} H_n \\ t_n = e_n + 2H_n^t P_{n+1} k_n + H_n^t q_{n+1} \\ U_n = B_n^t + 2H_n^t P_{n+1} G_n \end{cases}$$

and

$$\begin{cases} P_n = A_n + G_n^t P_{n+1} G_n - \frac{1}{4} U_n^t S_n^{-1} U_n \\ q_n = d_n + 2k_n^t P_{n+1} G_n + q_{n+1} G_n - \frac{1}{2} t_n^t S_n^{-1} U_n \\ r_n = f_n + k_n^t P_{n+1} k_n + q_{n+1} k_n + r_{n+1} - \frac{1}{4} t_n^t S_n^{-1} t_n \end{cases}$$

Then $W_n = -\frac{1}{2} S_n^{-1} U_n$ and $v_n = \frac{1}{2} S_n^{-1} t_n$.

Certainty Equivalence

Furthermore, the optimal decision rule for stage *n=1* depends only upon the *expected values* of the distributions of the random vectors $\xi_1, \xi_2, \ldots \xi_N$. *Thus the computations may be performed by replacing each* ξ_n *by its expected value* μ_n*, a result which is known as* ***Certainty Equivalence****!*

The major deficiency of the QC/LD model & algorithm is the lack of ability to restrict the state and decision variables (e.g., nonnegativity).

ⓘ *See the example of the* ***multi-reservoir control*** *problem for further discussion of imposing* ***constraints****.*

Example: *The following simple example suggests the validity of a certainty equivalence result for QC/LD problems.*

Consider the problem of choosing y so as to minimize

$$V(x,y) = E\left[c(gx + hy + \xi)^2\right]$$

where

- x, g, h, and c are numbers, and
- ξ is a **random** variable with **mean** μ and **variance** σ^2.

Then

$$V(x,y) = ch^2y^2 + 2ch(gx+\mu)y + c\left(g^2x^2 + \sigma^2 + \mu^2 + 2g\mu x\right)$$

We next compute the first and second derivatives of $V(x,y)$.

The first and second **derivatives** of V are

$$\frac{\partial}{\partial x} V(x,y) = 2ch^2 y + 2ch(gx+\mu)$$

and

$$\frac{\partial^2}{\partial x^2} V(x,y) = 2ch^2$$

If c>0 and h ≠ 0, V is convex and the unique minimum of V is achieved at

$$y = -\frac{gx+\mu}{h} = -\left(\frac{g}{h}x + \frac{\mu}{h}\right)$$

Notes:

- The optimal value of y depends only upon the *expected value* of ξ.
- The optimal value of y is given by a *linear decision rule*, i.e., a linear function of x.

See Daniel P. Heyman & Matthew J. Sobel, *Stochastic Models in Operations Research, Volume II: Stochastic Optimization*, McGraw-Hill Book Co., 1984, section 7-5.

HMMS

The Holt-Modigliani-Muth-Simon (HMMS) QC/LD model for production planning has

state variables $x_n = (i_n, z_n)$ where

- i_n = inventory level at stage n
- w_{n-1} = previous work force level

decision variables $y_n = (z_n , w_n)$ where

- z_n = production level at stage n
- w_n = work force level at stage n

Expected cost function is

$$\sum_n E\left[c\left(w_n - \alpha w_{n-1}\right) + h\left(i_n + z_n - D_n\right) + b\left(z_n\right) + d\left(\beta w_n - z_n\right)\right]$$

with convex (linear or quadratic) functions

- $c(\cdot)$ is cost of work force smoothing
- $h(\cdot)$ is cost of holding inventory
- $b(\cdot)$ is cost of production
- $d(\cdot)$ is cost of overtime/undertime labor

and

- α is retention rate for work force
- β is production rate of a worker

The optimal solution is, of course, a **linear decision rule**, *but one which may yield negative values of the decision variables!*

C.C. Holt, F. Modigliani, J. F. Muth, and H. A. Simon, *Planning Production, Inventories, and Work Force*, Prentice-Hall, 1960.