ARKIV FÖR MATEMATIK Band 3 nr 32

Communicated 12 September 1956 by HARALD CRAMÉR and OTTO FROSTMAN

A prediction problem in game theory

BY ULF GRENANDER

1 Diagram

1. Introduction

Two persons, X and P, play the following game. X controls a random mechanism producing a stationary stochastic process $x=\{x_t\}$. The mean value $E\,x(t)=0$ and the variance $E\,x^2(t)=1$ are given, but the form of the spectrum of the process can be determined completely by X. The player P observes the values $x(t)$ for $t\leq 0$ and wants to predict the value of

$$z=\int_0^1 a(t)\,x_t\,dt=A\,x,$$

where $a(t)$ is a real and continuous function of t in $(0, 1)$. For this purpose P has available all the linear predictors formed from the stochastic variables x_t, $t\leq 0$. Let us denote the chosen predictor by p and the predicted value by $p x$.

The payoff function of the game is $(A x-p x)^2$. X wishes to maximize the quantity $\|A x-p x\|^2=E\,[A x-p x]^2$ and P tries to minimize the same expression.

In the next section we study the value of $\max\min\|A x-p x\|^2$ and how this value is attained. In section 3 we show that this value coincides with $\min\max\|A x-p x\|^2$, the game is definite, and the pair of pure strategies is obtained that forms a solution of the game. It may be a matter of convention whether the strategy of the player X is considered as pure (a single choice of $g(t)$) or randomized (a choice of the stochastic process x). Sections 4 and 5 are devoted to some applications of the theorem in section 3.

The author was led to this result when working on an applied problem. The reader may be interested in consulting the paper of M. C. Yovits and J. L. Jackson [1] for a different approach to a similar problem.

2. Derivation of max min

When P first chooses his strategy p to minimize $\|A x-p x\|^2$ and then X chooses x to maximize the minimum value, it is clear that we can immediately limit the set of strategies of X to processes that are completely non deterministic.

To see this we note first that for given x we minimize $\|Ax-px\|^2$ by choosing

$$px=\int_0^1 a(t)\,x_t^*\,dt,$$

where x_t^* is the best prediction of the value x_t, when x_τ, $\tau\leq 0$, have been observed. Second, if x_t is an arbitrary stationary process it is well known that it can be decomposed $x_t=y_t+z_t$ into two terms, one, y_t, completely non deterministic and the other, z_t, deterministic and such that $y_s\perp z_t$ for all s, t. But then for the optimal predictor corresponding to x_t we have $\|Ax-px\|^2=\|Ay-py\|^2$ and $\|y\|\leq\|x\|=1$ with equality only if $z=0$. If $z\neq 0$ we would have reached a larger value be choosing the process $\|y\|^{-1}y_t$ instead of x_t, which proves the statement.

But if x_t is completely non-deterministic it can be represented as

$$x_t=\int_{-\infty}^{t} g(t-u)\,d\xi(u),$$

where $\xi(u)$ is a process with orthogonal increments and normed $\|\Delta\xi(u)\|^2=\Delta u$, and

$$\int_0^\infty g^2(t)\,dt=1. \tag{1}$$

There are in general many different representations of this form, but we are especially interested in the one (which always exists) such that the linear manifolds spanned by x_s, $s\leq t$, and $\xi(s)$, $s\leq t$, coincide for all values of t. Then the optimal predictor is simply

$$x_t^*=\int_{-\infty}^{0} g(t-u)\,d\xi(u),$$

and we have

$$\min_p \|Ax-px\|^2=\left\|\int_0^1 a(t)\,(x_t-x_t^*)\,dt\right\|^2=\int_0^1\int_0^1 a(s)\,a(t)\,r(s,t)\,ds\,dt,$$

where

$$r(s,t)=E\,(x_s-x_s^*)\,(x_t-x_t^*)=\int_0^{\min(s,t)} g(s-u)\,g(t-u)\,du.$$

Hence the minimum value is given by

$$\min_p \|A x - p x\|^2 = \int_0^1 \int_0^1 a(s)\, a(t) \int_0^{\min(s,t)} g(s-u)\, g(t-u)\, du\, ds\, dt$$

$$= \int_0^1 K(x, y)\, g(x)\, g(y)\, dx\, dy$$

with
$$K(x, y) = \int_0^{\min(1-x,\,1-y)} a(x+u)\, a(y+u)\, du. \tag{2}$$

Now the player X tries to make this minimum value as large as possible by choosing x in an appropriate way. Because of the side condition (1) it is clear that

$$\max_x \min_p \|A x - p x\|^2 \le v$$

if v is the largest eigenvalue of the continuous and symmetric kernel $K(x, y)$ in (2). On the other hand if $\gamma(t)$ is the normed eigenfunction corresponding to the eigenvalue v (or one of the functions if v is a multiple value) then

$$\min_p \|A x - p x\|^2 \ge \int_0^1 K(x, y)\, \gamma(x)\, \gamma(y)\, dx\, dy = v.$$

For this inequality and other results on stationary processes used in this section we refer the reader to J. L. Doob. [1].

We have now proved that max min = the largest eigenvalue v of the kernel $K(x, y)$. We now proceed to the derivation of the value of min max.

3. Derivation of min max

Let us approximate $a(t)$ uniformly with a step-function

$$a_n(t) = a\left(\frac{\nu+1}{n}\right) = a_\nu \quad \text{for} \quad \frac{\nu}{n} \le t < \frac{\nu+1}{n}, \qquad \nu = 0, 1, \ldots n-1.$$

As

$$\left\| \int_0^1 a(t)\, x_t\, dt - \int_0^1 a_n(t)\, x_t\, dt \right\| \le \|x_t\| \int_0^1 |a(t) - a_n(t)|\, dt$$

we see that we commit an arbitrarily small error if we deal with $\int_0^1 a_n(t)\, x_t\, dt$ instead of $A x$. Let us denote by C the class of all predictors formed from the stochastic variables

$$x_\nu = n \int_{\frac{\nu}{n}}^{\frac{\nu+1}{n}} x_t \, dt, \qquad \nu = 0 \pm 1, \ \pm 2, \ \dots .$$

Then it is clear that

$$\min_p \max_x \| A_n x - p x \| \le \min_{p \in C} \max_x \| A_n x - p x \|.$$

The x_ν form a stationary process with a variance less than but close to one when n is large. Let us denote by Γ the class of all discrete stochastic processes that can be represented as x_ν above in terms of a stationary stochastic process with a continuous time parameter. Then

$$\min_{p \in C} \max_x \| A_n x - p x \| = \min_{p \in C} \max_{x \in \Gamma} \| A_n x - p x \| \le \min_{p \in C} \max_{x \in D} \| A_n x - p x \|,$$

where D stands for the class of all discrete stationary processes with variance one.

Now the maximum can be expressed conveniently in another way. Using the well-known spectral representation we can write

$$x_\nu = \int_{-\pi}^{\pi} e^{i \nu \lambda} \, d Z(\lambda)$$

with $\| \Delta Z(\lambda) \|^2 = \Delta F(\lambda)$ and $F(\pi) - F(-\pi) = 1$. We also have

$$\frac{1}{n} \sum_0^{n-1} a_\nu x_\nu - \sum_{-\infty}^{-1} c_\nu x_\nu = \int_{-\pi}^{\pi} [\alpha(\lambda) - \gamma(\lambda)] \, d Z(\lambda),$$

where

$$\left. \begin{aligned} a(\lambda) &= \frac{1}{n} \sum_0^{n-1} a_\nu e^{i \nu \lambda} \\ \varphi(\lambda) &= \sum_{-\infty}^{-1} c_\nu e^{i \nu \lambda} \end{aligned} \right\}$$

and

$$\left\| \frac{1}{n} \sum_0^{n-1} a_\nu x_\nu - \sum_{-\infty}^{-1} c_\nu x_\nu \right\|^2 = \int_{-\pi}^{\pi} | \alpha(\lambda) - \gamma(\lambda) |^2 \, d F(\lambda).$$

Then

$$\max_{x \in D} \| A_n x - p x \|^2 = \max_\lambda | \alpha(\lambda) - \gamma(\lambda) |^2.$$

We can now apply the following theorem: Consider the class of all power series $f(\zeta)$ which are regular in $|\zeta| < 1$ and begin with the given terms $\sum_{k=0}^{n} \alpha_k \zeta^k$. We denote by μ_n^2 the highest eigenvalue of the matrix with elements

$$\left.\begin{aligned}&h_{pq}=\alpha_p\,\bar{\alpha}_q+\alpha_{p-1}\,\bar{\alpha}_{q-1}+\cdots+\alpha_0\,\bar{\alpha}_{q-p}\quad\text{for}\quad p\leq q,\\&h_{pq}=\bar{h}_{qp},\\&p,\,q=0,\,1,\,\ldots\,m.\end{aligned}\right\}$$

Then if $\mu_n>0$,

$$\max_{|\zeta|=1}|f(\zeta)|\geq\mu_n$$

and equality holds for a function of the form

$$f(\zeta)=\mu_n\,e^{i\gamma}\prod_{k=1}^{m}\frac{\zeta+w_k}{1+\bar{w}_k\,\zeta},\qquad\gamma\text{ real},\quad|w_k|<1.$$

For a fuller discussion and proof of this theorem see G. Szegö and U. Grenander [1].

In the present case we have to derive the largest eigenvalue of the symmetric matrix with the elements

$$h_{pq}=\frac{1}{n^2}\left[a(1-p/n)\,a(1-q/n)+a\left(1-\frac{p-1}{n}\right)a\left(1-\frac{q-1}{n}\right)+\cdots+\right.$$
$$\left.+a(1)\,a\left(1-\frac{q-p}{n}\right)\right]\quad\text{for}\quad p\leq q,\ p,\ q=0,\,1,\,\ldots,\,n-1.$$

As $a(t)$ is a continuous function of t it follows that for $p/n\to x$, $q/n\to y$, $x\leq y$ we have

$$\lim_{n\to\infty}n\,h_{pq}=\int_0^{\min(x,\,y)}a(1-x+u)\,a(1-y+u)\,du=K(1-x,\,1-y).$$

It now follows that $\lim_{n\to\infty}\mu_n^2=v$, the largest eigenvalue of the kernel $K(x,\,y)$. Hence we have shown that

$$\min_p\max_x\|A\,x-p\,x\|^2\leq v=\max_x\min_p\|A\,x-p\,x\|^2,$$

where only the equality sign is possible. This completes the proof of the

Theorem: The zero sum two person game defined above with the payoff function $(A\,x-p\,x)^2$ is definite. A solution of the game is given by the pure strategies

$$\left.\begin{aligned}&x_t=\int_{-\infty}^{t}\gamma(t-u)\,d\,\xi(u)\\&px=\int_0^1 a(t)\,x_t^*\,d\,t\end{aligned}\right\},$$

where $\gamma(t)$ is an eigenfunction corresponding to the largest eigenvalue v of the symmetric kernel

$$K(x, y) = \int_0^{\min(1-x, 1-y)} a(x+u)\, a(y+u)\, du.$$

The value of the game is v. When v is *a* simple eigenvalue the solution of the game is unique.

4. First application

When $a(x) = 1 - x$ we have to solve the integral equation

$$\int_0^1 K(x, y)\, g(x)\, dx = \lambda\, g(y)$$

with $K(x, y) = K(y, x)$ and

$$K(x, y) = \int_0^{1-x} (1 - u - y)(1 - u - x)\, du \quad \text{for } y < x.$$

This elementary problem will be solved by an argument of the type used in connection with Green's functions.

We have the partial derivatives with respect to y

$$\left.\begin{aligned} K'_y &= -\int_0^{1-x} (1 - x - u)\, du \\ K''_y &= K'''_y = 0 \end{aligned}\right\} \text{for } y < x$$

and

$$\left.\begin{aligned} K'_y &= -\int_0^{1-y} (1 - x - y)\, du \\ K''_y &= y - x \\ K'''_y &= 1 \end{aligned}\right\} \text{for } y > x.$$

Except for the point $y = x$ the function $K(x, y)$ has continuous derivatives of all orders with respect to y. In $y = x$ the first and second derivatives are continuous

$$\left.\begin{aligned} (K'_y)_{y=x-} = (K'_y)_{y=x+} &= -\int_0^{1-x} (1 - x - u)\, du \\ (K''_y)_{y=x-} = (K''_y)_{y=x+1} &= 0 \end{aligned}\right\}$$

but the third derivative has a discontinuity. Let us differentiate the integral equation twice with respect to y,

$$\lambda g''(y) = \int_0^1 K_y''(x, y)\, g(x)\, dx = \int_0^y (y-x)\, g(x)\, dx.$$

Differentiating this equation twice more we get

$$\lambda g^{\mathrm{IV}}(y) = g(y).$$

The solutions of this differential equation are of the form

$$g(t) = A e^{i\varkappa t} + B e^{-i\varkappa t} + C e^{\varkappa t} + D e^{-\varkappa t}$$

with $\varkappa = \lambda^{-\frac{1}{4}}$. To determine A, B, C and D we have to observe the boundary conditions that $g(t)$ has to satisfy. Indeed, as $K(x, 1) = K_y'(x, 1) = 0$ and because of $g''(0) = g'''(0) = 0$ we get

$$\left.\begin{aligned} A e^{i\varkappa} + B e^{-i\varkappa} + C e^{\varkappa} + D e^{-\varkappa} &= 0 \\ \varkappa\,(A i e^{i\varkappa} - B e^{-i\varkappa} + C e^{\varkappa} - D e^{-\varkappa}) &= 0 \\ \varkappa^2\,(-A - B + C + D) &= 0 \\ \varkappa^3\,(-i A + i B + C - D) &= 0 \end{aligned}\right\}.$$

In order that this system of linear equations have non-trivial solutions it is necessary that its determinant is zero

$$D(\varkappa) = \begin{vmatrix} e^{i\varkappa}, & e^{-i\varkappa}, & e^{\varkappa}, & e^{-\varkappa} \\ i e^{i\varkappa}, & -i e^{-i\varkappa}, & e^{\varkappa}, & -e^{-\varkappa} \\ -1, & -1, & 1, & 1 \\ -i, & i, & 1, & -1 \end{vmatrix} = 0.$$

Expanding this determinant after its first row we have

$$\begin{aligned} D(\varkappa) = {} & [2 i e^{-i\varkappa} + (i-1) e^{\varkappa} + (i+1) e^{-\varkappa}]\, e^{i\varkappa} - \\ & -[-2 i e^{i\varkappa} + (-1-i) e^{\varkappa} + (1-i) e^{-\varkappa}]\, e^{-i\varkappa} + \\ & +[(1+i) e^{i\varkappa} + (-1+i) e^{-i\varkappa} + 2 i e^{-\varkappa}]\, e^{\varkappa} - \\ & -[(1-i) e^{i\varkappa} + (-1-i) e^{-i\varkappa} - 2 i e^{\varkappa}]\, e^{-\varkappa} \\ = {} & 8 i\,[1 + \cos\varkappa \cosh \varkappa]. \end{aligned}$$

The largest value of λ corresponds to the smallest $\varkappa$ satisfying

$$\cos\varkappa \cosh \varkappa = -1.$$

We find approximately $\varkappa_{\min} \cong 1.88$, so that the value of the game is

$$v = \lambda_{\max} = \varkappa_{\min}^{-\frac{1}{2}} \cong 0.08.$$

To determine the corresponding eigenfunction we start from

$$\left.\begin{array}{llllll} A e^{i\varkappa} & + B e^{-i\varkappa} & + C e^{\varkappa} & + D e^{-\varkappa} & = 0 \\ -A & -B & +C & +D & =0 \\ -iA & +iB & +C & -D & =0 \end{array}\right\}$$

Putting

$$\left.\begin{array}{l} \alpha = A + B = C + D \\ \dfrac{\beta}{i} = A - B = \dfrac{C-D}{i} \end{array}\right\}$$

or

$$\left.\begin{array}{ll} A = \frac{1}{2}(\alpha + \beta/i), & B = \frac{1}{2}(\alpha - \beta/i) \\ C = \frac{1}{2}(\alpha + \beta), & D = \frac{1}{2}(\alpha - \beta) \end{array}\right\}$$

we get

$$\gamma(t) = \alpha\,(\cos \varkappa t + \cos h \varkappa t) + \beta\,(\sin \varkappa t + \sin h \varkappa t),$$

where α and β can be determined from

$$\left.\begin{array}{l} \gamma(1) = 0 \\ \int_0^1 \gamma^2(t)\,dt = 1 \end{array}\right\}$$

The form of $\gamma(t)$ is shown in the diagram.

5. Second application

If $a(x) \equiv 1$ we have to study instead the simple integral equation

$$\int_0^1 K(x, y)\, g(x)\, dx = \lambda g(y)$$

with $K(x, y) = \min(1-x, 1-y)$,

$$\lambda g(y) = (1-y) \int_0^y g(x)\,dx + \int_y^1 (1-x)\, g(x)\, dx.$$

With a familiar argument we get $\lambda g''(y) + g(y) = 0$ with $g(1) = g'(0) = 0$. Hence

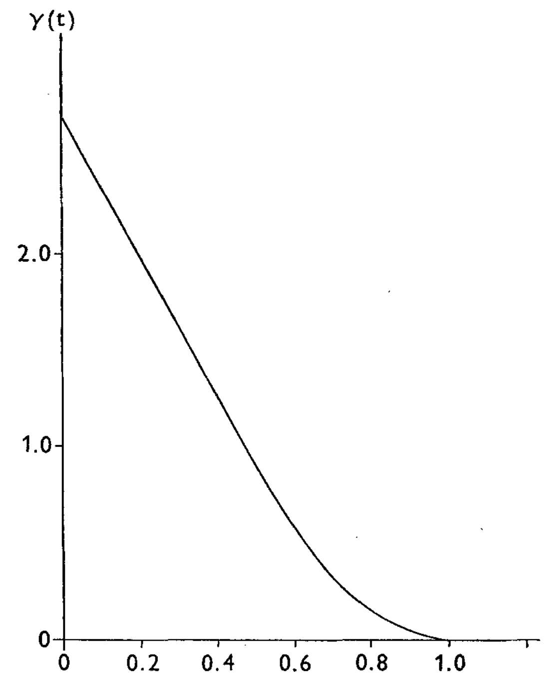

Diagram 1

$$g(t) = A e^{i\varkappa t} + B e^{-i\varkappa t}$$

with $\varkappa = \lambda^{-\frac{1}{2}}$. As

$$A e^{i\varkappa} + B e^{-i\varkappa} = 0,$$
$$A \quad - B \qquad = 0,$$

we should have $\cos \varkappa = 0$, so that $v = \lambda_{\max} \doteq \sqrt{2/\pi}$ and

$$\varkappa(t) = \sqrt{2} \cos \frac{\pi}{2} t.$$

REFERENCES

DOOB, J. L. [1], Stochastic Processes. New York, 1953.
SZEGÖ, G. and GRENANDER, U. [1], Toeplitz Forms and Their Applications. (To appear.)
YOVITS, M. C. and JACKSON, J. L. [1], Linear Filter Optimization with Game Theory Considerations. Inst. Radio Engin. Convention Record, 1955, Part 4.

Department of Mathematical Statistics, Stockholm, Sweden.

Tryckt den 27 december 1956

Uppsala 1956. Almqvist & Wiksells Boktryckeri AB