COWLES COMMISSION DISCUSSION PAPER: MATHEMATICS NO. 421

Some General Mean Value Operations, and Matrice-Operations Related to "Min" and "Max" Operations.

By Leo Törnquist

April 9, 1953

In the following we will make an attempt to study the relations between some often used mathematical operators, of which we pick out the "mean value" operators related to the problem of calculating "cumulants" to a probability distribution, determinant-calculation and some other new matrice operators nearly related to determinant operators, with special attention to their relations to each other and the "min" and "max" operators especially useful in the theory of rational decision making.

We first study the "cumulant operator:"

(1) $$\mathcal{C}^{(t)} = t^{-1} \log \; E \; \log^{-1} t = t^{-1} \, k^{(t)} \quad ; \quad \log^{-1} x = e^{x}.$$

which applied to a stochastic variable $P(\xi - x) = F(x)$ gives the function:

$$\mathcal{E}^{(t)} \xi = t^{-1} \log \int_{-\infty}^{\infty} e^{tx}\, dF(x) = t^{-1} \log \left(\sum_n v_n \frac{t^v}{n!}\right)$$

$$= E \xi + \sigma^2 \cdot \frac{t}{2} + E(\xi - E\xi)^3 \cdot \frac{t^2}{2!} + \dots \frac{k_n}{n} t^{n-1}$$

where $v_n = E \xi^n$ are the moments, and k_n the cumulants to $F(x)$. The function $\mathcal{E}^{(t)}(\xi)$ is a monotonically increasing function of (t) increasing from

(2) $$\mathcal{E}^{(-\infty)}(\xi) = \text{"Min"}\ \xi ;\quad dF(\xi) \neq 0$$

to

(3) $$\mathcal{E}^{(+\infty)}(\xi) = \text{"Max"}\ \xi ;\quad dF(\xi) \neq 0.$$

The operator $\mathcal{E}^{(t)}(\xi)$ is thus a mean value operator, connecting "Min" and "Max" operators by a set of continuously changing operators. We say that the operators $\mathcal{E}^{(t)}$ itself have an ordering relation

(4) $$\mathcal{E}^{(t)} > \mathcal{E}^{(t')} ;\quad \text{if } t > t'.$$

The derivates of the operator $k^{(t)} = t\,\mathcal{E}^{(t)}$ are also interesting. The first derivative

$$\frac{\partial k^{(t)}}{\partial t} \xi = \frac{\partial}{\partial t} \log E \log_t^{-1} \xi = \frac{E \xi \log^{-1} t \xi}{E \log^{-1} t \xi} = \xi_t$$

$$\xi_t = E \xi + \dots + \frac{k_n}{n!} t^{n-1} + \dots$$

is also a mean value operator. The mean value ξ_t increases monotonically from:

(5) $$\frac{\partial k^{(-\infty)}}{\partial t} \xi = \text{Min}\ \xi \qquad \text{to} \qquad \frac{\partial k^{(\infty)}}{\partial t} = \text{Max}\ \xi .$$

The second derivative

$$(6) \qquad \frac{\partial^2 k^{(t)}}{\partial t^2}\xi = \frac{E\,\xi^2 \log^{-1} t\,\xi}{E \log^{-1} t} - \left(\frac{\partial k^{(t)}}{\partial t}\xi\right)^2 = \sigma_t^2\ \xi.$$

is a measure of variance for ξ if the variance is calculated by means of the modified distribution assigning the probability

$$(7) \qquad dF_t(x) = \frac{(\log^{-1} tx)\ dF(x)}{\int \log^{-1} tx\ dF(x)} \ .$$

to the value $\xi = x.$

The derivates of the operator $k^{(t)}$ defines thus the semi invariants to a stochastic variable with the distribution

$$(8) \qquad F_t(x) = \int_{-\infty}^{x} dF_t(x) \ .$$

giving by increasing t more and more weight to the larger values of ξ.

If simple expressions for ${}^t\mathcal{E}\,\xi$ or $k^t\,\xi$ are known it is thus an easily solved problem to determine "$\underset{dF(\xi)\neq 0}{\text{Min}}\ \xi$ and $\underset{F(\xi)\neq 0}{\text{Max}}\ \xi$," and the statistically meaningful parameters $k_{tn} = \frac{\partial^n}{\partial t}\ k^{(t)}\xi$ = semi invariants to the distribution $F_t(x)$.

If the stochastical variable ξ is a table of numbers

$$(9) \qquad \xi = \left\{\left\{_{\substack{i=1\\ j=1\ \ldots\ n}}^{n} \ \xi_{ij}\right\}\right.$$

and $\vartheta(\xi)$ is a stochastical variable defined when ξ is known, the problem to maximize or minimize $\vartheta(\xi)$ can be solved if $k^{(t)}\,\vartheta(\xi)$ can be calculated.

A class of interesting problems refers to the case when every ξ obtainable by permutations of a basic table

$$(10) \qquad \xi^{o} = \left\{ \prod_{\substack{i=1\\ j=1,\,..\,n}}^{n} \vartheta_{ij} \right\}$$

have the same probability $P(\xi) = P(\xi_o)$. Basic operators applicable to a matrice ξ are determinant operators:

$$(11) \qquad D(\xi) = \sum_{p=P_n} D(p) \prod_{i=1} \vartheta_{ipi} \; ; \; P_n = \text{ the set of all } n!$$

permutations P. Where $D(p) = \pm 1$ is the determinant of the permutation matrice containing the elements

$$\xi_{ipi} = 1 \;;\; i = 1\ldots,n, \quad \xi_{ij} = 0 \quad j \neq p_i$$

and $p = \left\{ p_i \;;\; i = 1 \ldots n \right\}$ is a permutation of $(1, \ldots, n)$.

The properties of $D(\xi)$ operators are assumed to be known. We shall therefore study a broad class of matrice operators; with many properties similar to the determinant operators.

We define the matrice operator $G_e^{(t)}$ called "genna."

$$G_e^{(t)}(\xi) = G\left\{ e^{t\,\xi_{ij}} \right\} \sum_{p \in P} e^{\sum_i \xi_{ip(i)}} = \sum_p G_e^{(o)}(p)\, e^{t\, \xi_p}.$$

where

$$G_e^{(o)}(p) = 1 \quad \text{and} \quad \xi_p = \sum_i \xi_{ip(i)}$$

The cumulant function for the distribution of ξ_p is in the case when every p has the same probability $\frac{1}{n!}$

$$k^{(t)}(\xi_p) = \log (G_e^{(t)}(\xi) \quad ; \quad G_e^{(o)}(\xi) = \log \frac{1}{n!} G_e^{(t)}(\xi) .$$

making a study of the distribution of (ξ_p) easy if we can find a simple method for calculating $G_e^{(t)}(\xi)$. The similarity between $G_e^{(t)}(\xi)$ and the determinant operator $D(\xi)$ is obvious; a little change in the rules for calculation of determinants gives the corresponding rules for calculating $G_e^{(t)}(\xi)$.

In the case when n = 2 we have

$$G_e^{(t)} \left\{ \begin{matrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{matrix} \right| = e^{t(\xi_{11} + \xi_{22})} + e^{(\xi_{12} + \xi_{21}) t}$$

$$D_e^{(t)} \left\{ \begin{matrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{matrix} \right| = e^{t(\xi_{11} + \xi_{22})} - e^{(\xi_{12} + \xi_{21}) t}$$

The only difference is a difference in the sign rule for connecting the terms in the development of $G_e^{(t)}(\xi)$ respective $D_e^{(t)}(\xi)$. We draw the conclusion that $G_e^{(t)}(\xi)$ can be calculated according to the same rules as $D_e^{(t)}(\xi)$ if we always use <u>addition operator</u> by connecting corresponding "minors" of the expression $G_e^{(t)}(\xi)$.

The basic rule is thus:

$$G_e^{(t)}(\xi) = \sum_{i=1}^{n} e^{t\xi_{i1}} G_{e,i1}^{(t)}(\xi) \quad ; \text{ where } G_{e,i1}^{(t)}(\xi) \text{ is}$$

the genna for the minor matrice corresponding to the element ν_{i1}.

A considerable savings in the labor to find max ξ_p and min ξ_p can be obtained by taking already from the beginning $t \approx \pm \infty$ and modifying the rules for calculating $G_e^{(t)}$ in a suitable way to exclude terms which have no influence on $\text{Max}\, \nu_{pt} = \lim_{t \to \infty} \frac{\partial}{\partial t} G_e^{(t)}(\xi)$.

Such an operator is the matrice operator we will call "Maxma."

$\underline{\text{Maxma}}\left\{\xi\right\} = \hat{G}\xi$. We define $\hat{G}$ by the rule:

$$\hat{G}\begin{Bmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{Bmatrix} = \text{Max}\left\{(\xi_{11} + \xi_{22}), (\xi_{12} + \xi_{21})\right\}$$

and correspondingly for $m > 2$. Instead of addition of terms of the form $e^{t(\xi_{11} + \xi_{21})}, e^{t(\xi_{21} + \xi_{12})}$ in the "genna" we thus by calculating maxma only notice the factor f for t in the largest term e^{tf}. In other respects all rules for calculating determinants are valid. We get the interesting result:

$$\text{Maxma}\;(\xi) = \text{Max}_{p \in P}\; \xi_p = \xi_{\hat{p}} = \Big|_t^{\infty} \frac{\partial}{\partial t} \log G_e^{(t)}(\xi) = \text{Max}\,(\xi_{11} + \hat{G}_{11}\xi).$$

To the "Maxma" operator corresponds a similar "Minma" operator where we always notice only the smallest factor f in terms of the form e^{tf} in the corresponding "genna" operations.

The "Maxma" operator seems to me to be one of the most economical ways to calculate $\underset{p \in P}{\operatorname{Max}} \left(\nu_{p\zeta} = \sum_{i=1}^{n} \nu_{ipi} \right)$

Example 1.

$$\operatorname{Maxma} \begin{Bmatrix} \nu_{ij} \end{Bmatrix} = \operatorname{Maxma} \begin{Bmatrix} 5, & 3, & 1 \\ 6, & 7, & 0 \\ 1 & 5 & 9 \end{Bmatrix} = \operatorname{Max}\ (5+7+9),\ (6+3+9),\ (1+7+1) = 21$$

Example 2.

$$\operatorname{Maxma} \zeta = \operatorname{Maxma} \begin{Bmatrix} 1 & 5 & 4 & 8 \\ 3 & 5 & 3 & 1 \\ 2 & 6 & 7 & 0 \\ 1 & 1 & 5 & 9 \end{Bmatrix} = \operatorname{Max}\ (1+21,\ 5+19,\ 4+18,\ 8+14) = 24$$

Utilizing the following theorem 1 we have:

$$\text{If} \qquad \zeta'_{ij} = \left(\underset{i}{\operatorname{Max}}\ \zeta_{ij} \right) - \zeta_{ij}\ , \qquad \hat{\zeta}_{ij} = \underset{i}{\operatorname{Max}}\ \zeta_{ij}$$

$$\zeta''_{ij} = \zeta'_{ij} - \underset{j}{\operatorname{Min}}\ \zeta'_{ij} = \zeta'_{ij} - \check{\zeta}'_{ij}$$

then the permutation p giving

$$\operatorname{Minma} \zeta' = \operatorname{Minma} \zeta''$$

is the same as the permutation giving maxma $\zeta \circ \mathcal{E}x$.

$$\check{G}\xi' = \begin{vmatrix} 7 & 3 & 4 & 0 \\ 2 & 0 & 2 & 4 \\ 5 & 1 & 0 & 7 \\ 8 & 8 & 3 & 0 \end{vmatrix} = 5\,, \quad \check{G}\xi'' = \begin{vmatrix} 5 & 3 & 4 & 0 \\ 0 & 0 & 2 & 4 \\ 3 & 1 & 0 & 7 \\ 6 & 8 & 3 & 0 \end{vmatrix} = 3 = 5 - 2$$

$$5 - 3 = \check{G}(\xi'_{ij})\,, \qquad \hat{G}(\xi) = \hat{G}(\hat{\xi}) - \check{G}(\xi'_{ij}) - \check{G}(\xi'')$$

$$= 26 - 2 - 3 = 21.$$

$$\hat{G}(\hat{\xi}) = \check{G}(\hat{\xi}) = 26, \quad \hat{G}(\xi'_{ij}) = \check{G}(\xi'_{ij}) = 2, \quad \check{G}(\xi'') = 3.$$

"A Useful relations between "Maxma" and "Minma""

Theorem 1. If $\hat{\xi}$ denotes a matrice such that $\sum_i \hat{\xi}_{ipi} = \hat{\nu}_p$ = Maxma $\{\hat{\xi}_{ij}\}$ for every p for instance a matrice with the property $\hat{\xi}_{ij} = \underset{j}{\mathrm{Max}}\ \xi_{ij}$; $i = 1, \ldots, n,\ j = 1, \ldots, n$ or $\hat{\xi}_{ij} = \underset{i}{\mathrm{Max}}\ \xi_{ij}$, or some weighted mean of corresponding elements in these two matrices: Then

$$\text{Maxma } \xi = \text{Maxma } \hat{\xi} - \text{Minma } (\hat{\xi} - \xi) = \hat{G}\hat{\xi} - \check{G}(\xi'); \quad \xi' = \hat{\xi} - \xi .$$

The theorem is proved by making use of the fact that all $\nu_{p\hat{\xi}} = \nu_{\hat{p}\hat{\xi}}$ independently of the permutation used.

This theorem enables us to eliminate Maxma operators and make all necessary calculations of Minma $(\hat{\xi} - \xi)$, with comparitively small non negative numbers. It is often very easy to see by inspection of the matrice $(\hat{\xi} - \hat{\xi})$ which permutation $p = \hat{p}$ gives Minma $(\hat{\xi} - \xi)$. If $\hat{\xi}_{ij} = \underset{j}{\mathrm{Max}}\ \hat{\xi}_{ij}$, the matrice $(\hat{\xi} - \xi)$ has at least one element = 0 in every column. If these elements are located also in different rows Minma $(\hat{\xi} - \xi) = 0$, and Maxma ξ = Maxma $\hat{\xi}$.

The minma of the matrice in example 2 is $\mathrm{Min}(1 + 4,\ 5 + x,\ 4 + x',\ 8 + x'') = 3$. Where x, x', x'' is readily seen to be larger than 1 and can be left undetermined.

Problems of maximizing or minimizing "linear" expression of the form $\sum_i \nu_{ipi}$ over all permutations $p \in P$ are thus very simply solved by help of the rules for the minma $\left\{ \upsilon_{ij} \right\}$ operator.

A large class of optimalisation problems can be solved by means of sequences of the operators Max, Min, Maxma and Minma. Quantities such as $\underset{\rho}{\text{Min}}(\underset{ij}{\text{minma}}\ \zeta_{ij\rho})$, $\underset{ij}{\text{Minma}}\ (\underset{\lambda\rho}{\text{minma}}\ \zeta_{ij\lambda\rho})$ etc. are in many cases optimal values to econometrically significant optimalisations problems. We shall however, not in this introductory paper discuss these problems.