The Hermite polynomials in problem of the searching for position of the global maximum.

Yu.A.Yatsunenko
e-mail: yuyatsu@fnal.gov

Abstract

The symmetrization of the multi-extreme function around the objective point "u" of position of the global maximum simplifies the view of their expansion in the Hermite's polynomials-series. The analysis of the forth coefficient ($kurtosis, excess$) of the Hermite-series allows one to get the approach to this point "u".

The Hermite expansion of symmetrized function.

The developed methods of the searching for position (u) of the global maximum (the abbreviation "GM" will be used in the following) are mainly based on the Taylor-series expansion[1], [2] of the analyzed function $f(x)$

$$f(x) = f(u) + f^{'}(u)(x-u) + \frac{1}{2}f^{''}(u)(x-u)^2 + ... \quad (1)$$

However the Taylor-series (1) is defined at small x-vicinity of $f(u)$ ("locally") whereas the searching for the GM position foresees a large x-region where the multi-extreme character of the $f(x)$ is evinced. Thus the expansion of a $f(x)$ into the functional series should have a "non-local" character or, in other words, the coefficients of that should be determined by an integration over a x-region.

The functional series based on the Hermite polynomials [3] can be considered as the well suitable candidates for such expansion. Let us suppose for the following that the $f(x)$ is of the positive-definite character at $A < x < B$ and $f(x) \equiv 0$ for $-\infty < x < A$ and $B < x < \infty$. The direct expansion of the $f(x)$ in this Hermite-series does not give an idea for the searching of the GM position, however for the symmetrized construction (around unknown point u which is about of the GM position)

$$F(x,u) = \frac{1}{2}[f(x) + f(2u-x)] \quad (2)$$

this expansion is simplified significantly

$$F(x,u) = \frac{e^{-t^2}}{\sqrt{2\pi\mu_2(F)}}[1 + 0 \cdot H_1(t) + 0 \cdot H_2(t) + 0 \cdot H_3(t) + k_4 \cdot H_4(t) + ...] \quad (3)$$

where

$$t \equiv \frac{x-u}{\sqrt{2\mu_2(F)}}. \tag{4}$$

and $\mu_n(F)$ are the integral momenta of the $F(x,u)$ which determine the coefficients (k_n) at the $H_n(t)$ of (3)

$$\mu_n(F) = \mu_o^{-1}(F)\int F(x,u)(x-u)^n dx \tag{5}$$

Thus the $k_1 = 0, k_3 = 0$ due to the symmetrical character (around u) of the $F(x,u)$, k_2 equals to zero due to the chosen type of variable $t(4)$ and the forth coefficient k_4 becomes the function of point u

$$k_4 = \frac{1}{24}[\frac{\mu_4(F)}{\mu_2^2(F)} - 3] \equiv -\frac{E_F(u)}{12} \tag{6}$$

where the E_F is the *kurtosis* or *excess* which characterizes a sharpness of a function F.

If one can analytically calculate the momenta $\mu_n(f)$ of the original function $f(x)$

$$\mu_n(f) = \mu_o^{-1}(f)\int f(x)(x-c)^n dx,\ \ c \equiv \mu_1(f) = \mu_o^{-1}(f)\int f(x)x dx, \tag{7}$$

then all the momenta of the symmetrized $F(x,u)$ are easy determined by these $\mu_n(f)$ and therefore $E_F(u)$ has the exact analytical view

$$E_F(u) = \frac{(c-u)^4 - 2(c-u)\gamma + \sigma^2 \cdot R_o^2}{[\sigma^2 + (c-u)^2]^2} \tag{8}$$

where $\sigma^2 \equiv \mu_2(f), \gamma \equiv \mu_3(f)$ and the parameter

$$2R_o^2 = 3\sigma^2 - \frac{\mu_4(f)}{\sigma^2} \tag{9}$$

characterizes the distance between two peaks of $f(x)$ of the equal amplitude.

The extreme characteristic of excess.

Thus the symmetrization gives a very simple view of the Hermite-series (3): there is one maximum (on the variable t^2) the position and the amplitude of that is determined by the value of the $E_F(u)$. Therefore the extreme values of the $E_F(u)$ can determine the process of the symmetrization: if u is chosen near the GM position then function (2) has a narrow and high peak at x in the region of this maximum otherwise $F(x,u)$ is blurred. The requirement for the extreme value of the $E_F(u)$ ($dE_F(u)/du = 0$) gives the cubic equation for u

$$2(c-u)^3\sigma^2 + 3(c-u)^2\gamma - 2(c-u)R_o^2\sigma^2 - \gamma\sigma^2 = 0 \qquad (10)$$

It is the solution of the problem.

The iterative procedure of the convergence.

One of three possible real roots of this equation corresponding to the minimal value of the excess (6) places u most near to the GM position of $f(x)$. Then an iterative procedure of the successive reduction of the integration region of $f(x)$ can be organized: each new x-region can be determined by the chosen u and, for instance, by the current value of σ^2. A more cautious attitude to the approximate expansion (3) requires a necessity of the analysis of all the roots of equation (10) which correspond to the minimal value of the excess(6) (there are either 2 or 1). Undoubtedly, this is a very complicated analysis, however there is the well defined rule of the termination of the iterations: if $R_o^2(9) < 0$ or the cubic equation has one root only, then one can suppose that $f(x)$ is of unimodal character in the last integration region and the *local* methods can be used for the exact determination of the GM position.

The multidimensional analog of the excess.

It should be also noted that the Hermite-series expansion for multidimensional functions could give a similar way for the problem of the searching for the GM position. However the expansion which should be a suitable for a practical application has not been obtained yet. Nevertheless the analogous ideas about the symmetrization of the objective multidimensional function $f(\vec{r})$ (defined in a V-region)

$$F(\vec{r}, \vec{u}) = \frac{1}{2}[f(\vec{r}) + f(2\vec{u} - \vec{r})] \tag{11}$$

and the minimization of the excess can be simply used at this case. One can give the final form of the excess of the $F(\vec{r}, \vec{u})$ as the function of variable $\vec{R}(\vec{u}) = \vec{c} - \vec{u}$:

$$E_F(\vec{R}) = \frac{R^4 - 2\vec{R}\vec{\gamma} + 2\vec{R}^T \hat{K} \vec{R} + \sigma^2 R_o^2}{(\sigma^2 + R^2)^2} \tag{12}$$

where $\vec{c}$ is the center of the $f(\vec{r})$ (suppose in following that $\int\limits_V f(\vec{r}) dv = 1$)

and

$$\sigma^2 \equiv \int\limits_V f(\vec{r})(\vec{r}-\vec{c})^2 dv \tag{13}$$

$$\vec{\gamma} \equiv \int\limits_V f(\vec{r})(\vec{r}-\vec{c})^3 dv \tag{14}$$

$$2R_o^2 \equiv 3\sigma^2 - \sigma^{-2} \int\limits_V f(\vec{r})(\vec{r}-\vec{c})^4 dv \tag{15}$$

$$K_{mn} \equiv \sigma^2 \cdot \delta_{mn} - \int\limits_V f(\vec{r})(x_n - c_n)(x_m - c_m) dv \tag{16}$$

($\delta_{mn} = 1$ if $m = n$ otherwise $\delta_{mn} = 0$ and $m, n = 1, ..., N$; N is the dimension of vector $\vec{r}$).

Minimization of the excess (12) on the vector $\vec{R}$ leads to the equation

$$\frac{2\sigma^2 R^2 + 4\vec{R}\vec{\gamma} - 4\vec{R}^T \hat{K}\vec{R} - 2\sigma^2 R_o^2}{\sigma^2 + R^2}\vec{R} = \vec{\gamma} - 2\hat{K}\vec{R} \tag{17}$$

An attempt to find the solution based on the eigen vectors of matrix $\hat{K}$ (16) does not look as the practical one (it is under consideration, nevertheless).

A simplest solution of this equation is based on the assumption that the direction of the GM position is determined by the asymmetry ($\vec{\gamma}$)

$$\vec{R}(z) = z \cdot \frac{\vec{\gamma}}{\Gamma}, \quad (\Gamma \equiv \mid \vec{\gamma} \mid) \tag{18}$$

leads to the cubic equation for parameter z

$$2(\sigma^2 - \lambda_n)z^3 + 3\Gamma z^2 - 2\sigma^2(R_o^2 - \lambda_n)z - \sigma^2\Gamma = 0 \tag{19}$$

where λ_n are the eigenvalues of matrix $\hat{K}$ (16).

However the excess(12) has the finite number of extremal values (there are 3 ones in a common case) therefore the traditional local methods of the finding of the minimal value of the excess can be effectively applied and the above given solutions can be used as the start points.

A remark about the seraching for global minimum.

The naive substitution $\phi(x) = -f(x)$ is not suitable one whereas the construction $\phi(x) = C - f(x)$ leads out this problem to the above described method of the searching for the position of the global maximum. This constant C is determined, again, by the minimization of the excess $E_\phi(C)$, $(dE_\phi(C)/dC = 0)$ in the each x-region (A, B) of the integration (these calculations are simple enough if one can analytically calculate the momenta $\mu_n(f)$).

Problems to be solved.

Undoubtedly there are some open questions (convergence, optimal iterative procedure) which determine the efficiency of the described method. One can hope that a development of the expansion (3) (an analysis of the coefficients at H_6, H_8) can give new ideas in the analysis of multi-extrema functions.

Searching for the global maximum position of the two dimensional vertex function.

Despite to very approximate assumptions (12,18) the problem the searching for the GM position had been successfully solved for the two-dimensional function of the enough complicated view [5] as well as the effective realization of the unidimensional analysis (8) were made for the multi-extreme functions "vertex functions" (defined in previous talk and in [6]).

This"vertex function" for the straight tracks on the $XZ - plane$ has the view:

$$D(z,x) = \sum_{k=1}^{M_N} \sum_{n=1}^{N} \sum_{m=1}^{M_n} G[\frac{x - a_{kN}}{z_N - z}(z_N - z_n) + a_{kN} - a_{mn}; s_{mn}] \tag{20}$$

where N is the number of detectors and each of them has $M_n(n = 1, 2, ..., N)$ hits: $a_{mn}; (m = 1, 2, ..., M_n)$. The big maxima of the vertex function correspond to the possible vertices (x, z) $G[y; s]$ is the function of the accuracy (s) of the detection

$$G(y; s) \equiv exp(-\frac{y^2}{2s^2}) \tag{21}$$

The example of such vertex functions given on the fig.1 (upper plot), the same function in the contour view is shown on fig.2(upper plot too). The global maximum $(x = -10.0mm, z = 170.0mm)$ of this corresponds to the point of the most intensive interaction (18 trajectories) the second significant maximum $(x = 0.0mm, z = 140.0mm)$ reflects any "noise" interaction (there are 12 "noise" trajectories). There are a lot of small maxima too which have a combinatorial nature.

The discrete form the function of accuracy (21):

$$G[y; s] = \begin{cases} 1 & \text{if } \mid y \mid \leq \text{s} \\ 0 & \text{otherwise} \end{cases} \tag{22}$$

simplifies the analytical calculation of the integral momenta (13 - 16). The view of the excess (12) at these momenta is shown on fig.1 (lower plot) and on fig.2 (lower plot) in the contour representation. The roots of the cubic equation (13) are given on fig.2 (lower plot): there were two eigenvalues and 4 roots (x, z) totally (1 root for the first eigenvalue and 3 roots for the second one), the root minimizing the excess is shown as the "dark circle" others three ones are shown as the "dark squares" on this picture. The root "dark circle" was used as the start point for determination (by *local* methods) of the point of the minimum of the excess that is shown as the "dark star". This last point well coincides with the position of the main maximum of the vertex function and the following sharpening of that position by the *local* methods is performed quickly. These roots are shown on the contour view of the vertex function (upper plot of fig.2).

References

[1] Ya.D.Sergeev, Journal of Comput.Math. and Math.Phys. (Russia) 1995, V35, No 5, p707.

A.G.Sukharev, A.B.Timokhov and V.V.Fedorov. "Manual of the methods of the optimization", "Nauka", Moskow, 1986, p49.

[2] A.A.Goldstein and J.F.Price, Math.Comput.25,1971,p569.

[3] A.Korn and M.Korn. "Mathematical Handbook" (Russian edition) "Nauka",Moskow,1977,p.775.

[4] Yu.A.Yatsunenko,JINR Communication,P5-88-74,1988,Dubna(Russia).

[5] Yu.A.Yatsunenko,JINR Rapid Communication, No 5,1989,p26, Dubna(Russia).

[6] Yu.A.Yatsunenko, Nucl.Instr.and Methods(NIM), A287(1990), p422.

Figure captions

Fig.1 The view of the vertex function (20) is shown on upper plot. The excess (12) of this function is shown on the lower plot.

Fig.2 The contour view of the vertex function (20) (upper plot) and of the excess (12) (lower plot) presented on fig.1. The "dark squares" , "dark circle" are the roots of the cubic equation (13) for two eigenvalues. The root "dark circle" was used for the numerical determination of the point ("dark star") of the minimum of the excess.