## APPROXIMATION OF A STRAIGHT LINE IN SOFTWARE FOR CO-ORDINATE MEASURING MACHINE


#### Abstract

The software for co-ordinate measuring machine (CMM) is used for approximating geometric elements in plane (2d) and in space (3d) by approximation algorithms when number of measuring points obtained is larger than the minimum number of points necessary for element definition. Orthogonal regression methods are discussed in this paper. For 2d and 3d cases and they are compared with methods for defining straight line using linear regression. The advantage of orthogonal regression over linear regression is shown by comparing variance of measuring point displacement from both approximation line types.Algorithms for orthogonal regression make it possible to determine optimum position of a straight line for which value of variance of point distance from the line is the lowest possible. The tests have been performed on the Carl Zeiss CMM.


## 1. Introduction

A set of points obtained in the machine head movement along the plane perpendicular to the CMM table is approximated by a 2 d straight line in the Cartesian system of co-ordinates XY. A set of points positioned along axes of successive cylinders in cylinder measuring procedure is approximated by a 3d straight line in the Cartesian space XYZ. The straight lines in question are called regression lines. The least squares method by Gauss is used for determining position of the regression line. Linear regression and orthogonal regression method has been described in available mathematical references [2], [3], [4].

Linear regression method is based on calculating the least squares of point displacement from theoretical line as measured parallel to the OX or

[^0]OY axis of the Cartesian system. The regression lines obtained are at angle $\varphi$ named correlation scissors.

Orthogonal regression method is based on calculating the least squares of point displacement from theoretical line taken in perpendicular direction relative to theoretical line.

Both cases account in determining the minimum function value using methods of differential calculus. The methods presented below have been applied for the 2 d and 3 d case. The 3 d measurement of the straight line is practically impossible on the CMM and it can be determined only by a calculation procedure, such as the one called CALYPSO, provided by the Zeiss company, where straight line passes by the consecutive cylinder bases [1]. The orthogonal regression methods are usually used by CMM software.

## 2. Linear regression

A sequence of points $P_{i}\left(x_{i}, y_{i}\right) i=1, \ldots, n$ is usually obtained as a result of CMM measurements along a straight line (Fig. 1) and this sequence is recorded as a PC computer file.


Fig. 1. Diagrams for calculating deviations $\eta_{i}$ of measuring point positions from ideal line passing through mass center $\bar{x}, \bar{y}$ of a set of measuring points $\mathrm{P}_{\mathrm{i}}$ : a) deviation $\eta_{i}$ of measuring point positions from the PL1 line measured as parallel to the OY axis, b) deviation $\eta_{i}$ of measuring point positions from the PL1 line measured as parallel to the OY axis

Measurements of points in the XY plane can be carried out using a couple of different methods: a) measuring head is driven as parallel to OY axis until the spherical terminal contacts the straight line, b) measuring head is driven as parallel to OX axis until the spherical terminal contacts the straight line
(linear regression). The points measured are used for determining line PL1 (Fig. 1a) or line PL2 (Fig. 1b).

### 2.1. Linear regression in use for $2 d$ line definition by the Linnik method

Equation of 2d straight line in its directional form goes as follows:

$$
\begin{equation*}
y=b_{0}+b_{1} x \tag{1}
\end{equation*}
$$

There are at least two measuring points $P_{i}\left(x_{i}, y_{i}\right) i=1,2$ needed for determining the equation (1) where $b_{0}$ and $b$ are to be calculated. The calculated coefficients will be dependent on which two points have been selected from point set $P_{i}\left(x_{i}, y_{i}\right) i=1, \ldots, n$.

Least squares method (Gauss method) is used for determining $b_{0}$ and $b_{1}$ if number of points $P_{i}\left(x_{i}, y_{i}\right)$ is greater than two. In this method, quadratic values $\eta_{i}$ for deflections of measuring points $P_{i}\left(x_{i}, y_{i}\right)$ from theoretical straight line along OY axis (2) or OX axis (3) are defined, then a sum of $\eta_{i}^{2}$ squares is computed and minimum from the expression (2) or (3) is determined:

$$
\begin{align*}
& F\left(b_{0}, b_{1}\right)=\sum_{i=1}^{i=n} \eta_{i}^{2}=\sum_{i=1}^{i=n}\left(y_{i}-\left(b_{0}+b_{1} x_{i}\right)\right)^{2}=\min  \tag{2}\\
& F\left(b_{0}, b_{1}\right)=\sum_{i=1}^{i=n} \eta_{i}^{2}=\sum_{i=1}^{i=n}\left(x_{i}-\left(b_{0}+b_{1} y_{i}\right)\right)^{2}=\min \tag{3}
\end{align*}
$$

where $n$ is the measurement number.
Regression lines pass through mass center of a set of measuring points $P_{i}\left(x_{i}, y_{i}\right)$ (Fig. 2). Both regression lines $\mathrm{PL}_{\mathrm{y} / \mathrm{x}}$ and $\mathrm{PL}_{\mathrm{x} / \mathrm{y}}$ develop the so called correlation scissors of width $\varphi$. Tangent of $\varphi$ is calculated according to formula [4]:

$$
\begin{equation*}
\operatorname{tg} \varphi=\frac{\sigma_{x} \sigma_{y}}{\sigma_{x}^{2}+\sigma_{y}^{2}} \frac{1-\rho^{2}}{|\rho|} \tag{4}
\end{equation*}
$$

where $\sigma_{x}, \sigma_{y}$ - standard deviations for variables $X$ and $Y, \rho$ - correlation coefficient for variables X and Y .

Only when measuring points are located exactly on the straight line, i.e. functional relationship exists, correlation coefficient $\rho=1$, hence $\tan (\varphi)=0$ and correlation scissors close. If $\rho=0$, then $\tan (\varphi)=\infty$ and correlation scissors open up to $90^{\circ}$. Points can not be approximated by a straight line equation this time.


Fig. 2. Regression lines PL1 and PL2 cross at mass center ( $\bar{x}, \bar{y}$ ) and they develop correlation scissors

Because of high CMM accuracy, measuring points are located very close to regression line so that correlation coefficient is close to 1 . Value of angle $\varphi$ is close to zero and correlation scissors are about to close.

Function $\mathrm{F}\left(\mathrm{b}_{0}, \mathrm{~b}_{1}\right)$ ((2),(3)) has been minimized according to total differential principle, i.e. partial derivatives with respect to parameters $b_{0}$ and $\mathrm{b}_{1}$ have been calculated and then zeroed. After a few transformations, relationship describing $b_{1}$ coefficient can be obtained as tangent of angle for the slope of the regression line passing through the mass center of point set $P_{i}\left(x_{i}, y_{i}\right)$ :

$$
\begin{equation*}
b_{1}=\operatorname{tg} \alpha=\frac{\operatorname{cov}(X, Y)}{D^{2}(X)}=\frac{\sum_{i=1}^{n}\left(x_{i}-\bar{x}\right)\left(y_{i}-\bar{y}\right)}{\sum_{i=1}^{n}\left(x_{i}-\bar{x}\right)^{2}} \tag{5}
\end{equation*}
$$

Another constant $b_{0}$ can be calculated using the formula:

$$
\begin{equation*}
b_{0}=\left(\bar{y}-\frac{\operatorname{cov}(X, Y)}{D^{2}(X)} \cdot \bar{x}\right) \tag{6}
\end{equation*}
$$

Constants $b_{0}$ and $b_{1}$ are used to define linear function (1), i.e. function $y=f(x)$. The regression line with respect to independent variable $y$ can be defined in a similar way as an inverse function $x=f(y)$ with respect to regression function (1).

## 3. Orthogonal regression for 2 d line

### 3.1. Orthogonal regression for 2d line definition by Lagrange

In order to determine line of orthogonal regression, the distances $\eta_{\mathrm{i}}$ of points $P_{i}\left(x_{i}, y_{i}\right)$ from theoretical line PL (Fig. 3) have been calculated. The line defined by orthogonal regression passes through the mass center of measuring point set. A normal line defined in the local co-ordinate system ( $X^{\prime}, Y^{\prime}$ ), with the origin at the mass center, can be described as follows:

$$
\begin{equation*}
b_{1} x^{\prime}+b_{2} y^{\prime}=0 \tag{7}
\end{equation*}
$$

with additional condition:

$$
b_{1}^{2}+b_{2}^{2}-1=0
$$

where local coordinates ( $x_{i}^{\prime}, y_{i}^{\prime}$ ) are connected with global coordinates $\left(x_{i}, y_{i}\right)$ in the following way:

$$
\left.\begin{array}{l}
x_{i}^{\prime}=x_{i}-\bar{x} \\
y_{i}^{\prime}=y_{i}-\bar{y} \tag{8}
\end{array}\right\}
$$



Fig. 3. Diagram for calculating distances $\eta_{i}$ of measuring points from theoretical line PL passing through mass center ( $\bar{x}, \bar{y}$ ) of measuring point set $\mathrm{P}_{\mathrm{i}}$. Distances $\eta_{\mathrm{i}}$ are measured along normal direction to the PL line; local co-ordinate system $\left(X^{\prime}, Y^{\prime}\right)$ is shifted with respect to global co-ordinate system $(X, Y)$ by vector $[\bar{x}, \bar{y}]$

Further analysis of the line slope will be carried out in the local coordinates $\left(x_{i}^{\prime}, y_{i}^{\prime}\right.$ ) but the 'denotation will be skipped. The quoted technique
is called the condition extreme Lagrange method [5]. The sum of squares of the measuring point distances from PL line (Fig. 3) is calculated according to the formula:

$$
\begin{equation*}
f\left(b_{1}, b_{2}\right)=\sum_{i=1}^{n} \eta_{i}^{2}=\sum_{i=1}^{n}\left(b_{1} x_{i}+b_{2} y_{i}\right)^{2}=\min \tag{9}
\end{equation*}
$$

with extra condition: $\quad b_{1}^{2}+b_{2}^{2}-1=0$.
The Lagrange function takes the following shape

$$
\begin{equation*}
F\left(b_{1}, b_{2}, \lambda\right)=f\left(b_{1}, b_{2}\right)+\lambda\left(b_{1}^{2}+b_{2}^{2}-1\right) \tag{10}
\end{equation*}
$$

The conditional extreme of the Lagrange function can happen only when the partial derivatives of function (10) with respect to parameters $b_{1}, b_{2}$, and $\lambda$ equal zero.

After some necessary transformations, the following system of linear equations is obtained.

$$
\left.\begin{array}{l}
b_{1} \sum_{i=1}^{n} x_{i}^{2}+b_{2} \sum_{i=1}^{n} x_{i} y_{i}=\lambda b_{1}  \tag{11}\\
b_{1} \sum_{i=1}^{n} x_{i} y_{i}+b_{2} \sum_{i=1}^{n} x_{i}^{2}=\lambda b_{1}
\end{array}\right\}
$$

The introduction of designations $A=\left[\begin{array}{cc}D^{2}(X) & \operatorname{cov}(X, Y) \\ \operatorname{cov}(X, Y) & D^{2}(Y)\end{array}\right], b=\left[\begin{array}{l}b_{1} \\ b_{2}\end{array}\right]$ simplifies the description of the equation (11) to:

$$
\begin{equation*}
A \cdot b=\lambda \cdot b \tag{12}
\end{equation*}
$$

The system of equations (12) will be homogeneous after some transformations. Non-zero solution of homogeneous system, eigen solution for symmetric matrix, can be found using NROOT and EIGEN subroutines [9]. These subroutines provide normalized values of vector $b$, components of which denote cosines for directional angles of regression line with respect to OX and OY axes. Moreover, NROOT subroutine provides two eigenvalues $\lambda_{1}$ and $\lambda_{2}$ determining components of two eigenvectors. Vector defining the searched line is connected with the smaller value of $\lambda$. The other eigenvector defines perpendicular line.

### 3.2. Orthogonal regression for determining $2 d$ line by the Linnik method

The essence of 2 d orthogonal regression application was presented in Fig. 4. The local co- ordinate system $(U, V)$ was set with origin located in mass center of point set $P_{i}\left(x_{i}, y_{i}\right)$.

The $(U, V)$ system was rotated with respect to $\left(X^{\prime}, Y^{\prime}\right)$ system which in turn was translated with respect to $(X, Y)$ system. Co-ordinates of $P_{i}$ point in ( $U, V$ ) system is bound by the following relationships:

$$
\left.\begin{array}{l}
u_{i}=x_{i}^{\prime} \cos \alpha+y_{i}^{\prime} \sin \alpha  \tag{13}\\
v_{i}=-x_{i}^{\prime} \sin \alpha+y_{i}^{\prime} \cos \alpha
\end{array}\right\}
$$



Fig. 4. Location of measuring point $\mathrm{P}_{\mathrm{n}}\left(\mathrm{x}_{\mathrm{n}}, \mathrm{y}_{\mathrm{n}}\right)$ in co-ordinate system ( $U, V$ ) rotated with respect to ( $X^{\prime}, Y^{\prime}$ ) by an angle $\alpha$. The ( $X^{\prime}, Y^{\prime}$ ) system is located at mass center $(\bar{x}, \bar{y})$ for point set $\mathrm{P}_{\mathrm{i}}$

The distance of point from local axis $U$ is denoted by $v_{i}$. According to the Gauss rule, the minimizing procedure is applied:

$$
\begin{equation*}
\Theta=\sum_{i=1}^{n} v_{i}^{2}=\min \tag{14}
\end{equation*}
$$

After differentiation of function (14) with respect to parameter $\alpha$ and equaling the derivative to zero, relationship is obtained for tangent of double $\alpha$ :

$$
\begin{equation*}
\operatorname{tg} 2 \alpha=\frac{2 \sum_{i=1}^{n} x_{i}^{\prime} y_{i}^{\prime}}{\sum_{i=1}^{n} x_{i}^{\prime 2}-\sum_{i=1}^{n} y_{i}^{\prime 2}} \tag{15}
\end{equation*}
$$

The angle of the line slope to OX axis in global co-ordinate system is calculated according to another relationship

$$
\begin{equation*}
\alpha=\frac{1}{2} \operatorname{arctg} \frac{\sum_{i=1}^{n}\left(x_{i}-\bar{x}\right)\left(y_{i}-\bar{y}\right)}{\sum_{i=1}^{n}\left(x_{i}-\bar{x}\right)^{2}-\left(y_{i}-\bar{y}\right)^{2}} \tag{16}
\end{equation*}
$$

## 4. Orthogonal regression for 3d line

3d line in Cartesian system can be expressed in many ways. The most important ones go as follows:
direction form of line L equations: $\mathrm{L}: \frac{x-x_{0}}{v_{x}}=\frac{y-y_{0}}{v_{y}}=\frac{z-z_{0}}{v_{z}}$
parametric form of line L equations: $\mathrm{L}:\left\{\begin{array}{l}x=x_{0}+v_{x} \cdot t \\ y=y_{0}+v_{y} \cdot t \\ z=z_{0}+v_{z} \cdot t\end{array}\right.$
where: $\left(\mathrm{x}_{0}, \mathrm{y}_{0}, \mathrm{z}_{0}\right)$ - co-ordinates of point placed on line, $\left[v_{x}, v_{y}, v_{z}\right]-$ directional vector components, $t$ - parameter.

The components of vector $\vec{v}$ are direction cosines of the line 3 d assuming positive sense of Cartesian co-ordinate system axes. PL1 line in $\mathrm{R}^{3}$ plane and distribution of measuring points $P_{i}$ in its neighborhood are presented in Fig. 5.


Fig. 5. Diagram for calculating distances $\eta_{i}$ of measuring points from theoretical line PL passing through mass center of measuring points set $\mathrm{P}_{\mathrm{i}}\left(\mathrm{x}_{\mathrm{i}}, \mathrm{y}_{\mathrm{i}}, \mathrm{z}_{\mathrm{i}}\right) ;\left(\mathrm{X}_{\mathrm{i}^{\prime}}, \mathrm{y}_{\mathrm{i}^{\prime}}, \mathrm{z}_{\mathrm{i}^{\prime}}\right)$ are coordinates of $\mathrm{P}_{\mathrm{i}}$ in $\left(X^{\prime}, Y^{\prime}, Z^{\prime}\right)$ local co-ordinate system shifted with respect to global co-ordinate system ( $X, Y, Z$ ) by vector $[\bar{x}, \bar{y}, \bar{z}]$

### 4.1. Orthogonal regression for 3d line according to Lagrange method

The line defined by orthogonal regression passes through the mass center of measuring points registered on the CMM. The distances of measuring points $P_{i}$ from line PL are denoted as $\eta_{\mathrm{i}}$ (fig. 6).

The distance $\eta_{\mathrm{i}}$ of measuring point $P_{i}\left(x_{i}^{\prime}, y_{i}^{\prime}, z_{i}^{\prime}\right)$ from line PL in local co-ordinate system ( $X^{\prime}, Y^{\prime}$ ) can be calculated from the following formula [5]:

$$
\begin{equation*}
\eta_{i}=\frac{\left|\vec{b} \times \vec{P}_{i}^{\prime}\right|}{|\vec{b}|} \tag{19}
\end{equation*}
$$

where $\vec{b}=\left[b_{1}, b_{2}, b_{3}\right], \vec{P}_{i}^{\prime}=\left[x_{i}^{\prime}, y_{i}^{\prime}, z_{i}^{\prime}\right]$ and there is additional requirement:

$$
\begin{equation*}
|\vec{b}|=\sqrt{b_{1}^{2}+b_{2}^{2}+b_{3}^{2}}=1 \tag{20}
\end{equation*}
$$

Further notation will not include the ' sign. It should be remembered that coordinates $\left(x_{i}, y_{i}\right)$ are connected with the local system ( $X^{\prime}, Y^{\prime}$ ). The absolute value of a vector $\left|\vec{b} \times \overrightarrow{P_{i}}\right|$ can be computed from the co-ordinates of vectors $\vec{b}$ and $\vec{P}_{i}$ [5]:

$$
\left|\vec{b} \times \vec{P}_{i}\right|=\sqrt{\left|\begin{array}{cc}
b_{2} & b_{3}  \tag{21}\\
y_{i} & z_{i}
\end{array}\right|^{2}+\left|\begin{array}{cc}
b_{1} & b_{3} \\
x_{i} & z_{i}
\end{array}\right|^{2}+\left|\begin{array}{cc}
b_{1} & b_{2} \\
x_{i} & y_{i}
\end{array}\right|^{2}}
$$



Fig. 6. Line PL passes through mass center $(\bar{x}, \bar{y}, \bar{z})$ of a set of measuring points $P_{i}\left(x_{i}, y_{i}, z_{i}\right)$. The distance $\eta_{\mathrm{i}}$ of point $\mathrm{P}_{\mathrm{i}}$ from the line is computed basing on absolute value of a vector $\left[\vec{b} \times \vec{P}_{i}\right]$ and of directional vector $[\vec{b}]$

A function $f\left(b_{1}, b_{2}, b_{3}\right)$ being a sum of distance $\eta_{\mathrm{i}}$ squares was subjected to minimizing procedure which can be expressed as follows:

$$
f\left(b_{1}, b_{2}, b_{3}\right)=\sum_{i=1}^{n} \eta_{i}^{2}=\sum_{i=1}^{n}\left(\left|\begin{array}{ll}
b_{2} & b_{3}  \tag{22}\\
y_{i} & z_{i}
\end{array}\right|^{2}+\left|\begin{array}{ll}
b_{1} & b_{3} \\
x_{i} & z_{i}
\end{array}\right|^{2}+\left|\begin{array}{ll}
b_{1} & b_{2} \\
x_{i} & y_{i}
\end{array}\right|^{2}\right)=\min
$$

The Lagrange function is created in order to find conditional extreme for three variables:

$$
\begin{equation*}
F\left(b_{1}, b_{2}, b_{3}, \lambda\right)=f\left(b_{1}, b_{2}, b_{3}\right)-\lambda\left(b_{1}^{2}+b_{2}^{2}+b_{3}^{2}-1\right) \tag{23}
\end{equation*}
$$

for additional condition: $b_{1}^{2}+b_{2}^{2}+b_{3}^{2}-1=0$
Partial derivatives with respect to four parameters are formulated and they are made equal to zero. Then, a system of liner equations is created as follows:

$$
\left.\begin{array}{l}
b_{1} \sum_{i=1}^{n}\left(y_{i}^{2}+z_{i}^{2}\right)+b_{2} \sum_{i=1}^{n}-x_{i} y_{i}+b_{3} \sum_{i=1}^{n}-x_{i} z_{i}=\lambda \cdot b_{1} \\
b_{1} \sum_{i=1}^{n}-x_{i} y_{i}+b_{2} \sum_{i=1}^{n}\left(x_{i}^{2}+z_{i}^{2}\right)+b_{3} \sum_{i=1}^{n}-y_{i} z_{i}=\lambda \cdot b_{2}  \tag{25}\\
b_{1} \sum_{i=1}^{n}-x_{i} z_{i}+b_{2} \sum_{i=1}^{n}-y_{i} z_{i}+b_{3} \sum_{i=1}^{n}\left(x_{i}^{2}+y_{i}^{2}\right)=\lambda \cdot b_{3}
\end{array}\right\}
$$

After adopting the matrix-type notations, it can be rewritten as follows:

$$
\mathrm{A}=\left[\begin{array}{lll}
\sum_{i=1}^{n}\left(y_{i}^{2}+z_{i}^{2}\right) & \sum_{i=1}^{n}-x_{i} y_{i} & \sum_{i=1}^{n}-x_{i} z_{i} \\
\sum_{i=1}^{n}-x_{i} y_{i} & \sum_{i=1}^{n}\left(x_{i}^{2}+z_{i}^{2}\right) & \sum_{i=1}^{n}-y_{i} z_{i} \\
\sum_{i=1}^{n}-x_{i} z_{i} & \sum_{i=1}^{n}-y_{i} z_{i} & \sum_{i=1}^{n}\left(x_{i}^{2}+y_{i}^{2}\right)
\end{array}\right], \quad b=\left[\begin{array}{l}
b_{1} \\
b_{2} \\
b_{3}
\end{array}\right]
$$

or shortly put as

$$
\begin{equation*}
A \cdot b=\lambda \cdot b \tag{26}
\end{equation*}
$$

it can also be expressed as:

$$
\begin{equation*}
(A-\lambda \cdot E) \cdot b=0 \tag{27}
\end{equation*}
$$

where E is unit matrix.
Homogeneous equation (27) will have a solution different from zero $(b \neq 0)$ when the matrix $(A-\lambda \cdot E)$ is a singular i.e. when its determinant

$$
\begin{equation*}
\operatorname{det}(A-\lambda \cdot E)=0 \tag{28}
\end{equation*}
$$

Condition (28) leads to the characteristic equation of a matrix solutions of which are eigenvalues $\lambda_{i}, i=1,2,3$. Eigenvalues $\lambda_{i}$ and eigenvectors $b_{i}$ are computed by the NROOT procedure within which an EIGEN subroutine is called [9]. In this procedure, an iterative method of Jacobi is applied. The computations carried out with help of the subroutines yielded three eigenvalues $\lambda_{i}, i=1,2,3$ and corresponding eigenvectors of components equal to direction cosines of vectors parallel to the lines. The vector corresponding to the smallest eigenvalue is the object to be found. Summing up, 3d regression
line passing through mass center $(\bar{x}, \bar{y}, \bar{z})$ of the set of points $P_{i}\left(x_{i}, y_{i}, z_{i}\right)$ and components of this line direction vector are determined by $\vec{b}$ for the smallest eigenvalue $\lambda_{i}$.

## 5. Variance for deviations of measuring points from $2 d$ line of linear and orthogonal regression

Lines of linear and orthogonal regression pass through the mass center of point set $P_{i}\left(x_{i}, y_{i}\right)$ obtained during CMM measurements. The slope values of regression lines are much different. Optimum angle for regression line slope will be researched according to the square root from the variance of distances of measuring points from theoretical line. The 2 d line passing through mass center $(\bar{x}, \bar{y})$ is controlled by the following relationship:

$$
\begin{equation*}
y-\bar{y}=\operatorname{tg} \alpha \cdot(x-\bar{x}) \tag{29}
\end{equation*}
$$

The distance $\eta_{i}$ from a point $P_{i}\left(x_{i}, y_{i}\right)$ from theoretical line can be described as follows:

$$
\begin{equation*}
\eta_{i}=\frac{\left|\operatorname{tg} \alpha \cdot x_{i}-y_{i}+(\bar{y}-\bar{x} \cdot \operatorname{tg} \alpha)\right|}{\sqrt{\operatorname{tg}^{2} \alpha+1}} \tag{30}
\end{equation*}
$$

The following substitutions facilitating mathematical description were used:

$$
\begin{equation*}
c f x=\operatorname{tg} \alpha, \quad c f=(\bar{y}-\bar{x} \cdot \operatorname{tg} \alpha) \tag{31}
\end{equation*}
$$

A variance of points $P_{i}\left(x_{i}, y_{i}\right)$ with respect to a line of linear regression can be expressed as follows:

$$
\begin{equation*}
D^{2}(\eta)=\frac{\sum_{i=1}^{n}\left(c f x \cdot x_{i}-y_{i}+c f\right)^{2} /\left(c f x^{2}+1\right)}{n} \tag{32}
\end{equation*}
$$

Standard deviation $\sigma$ for measuring points $P_{i}\left(x_{i}, y_{i}\right)$ can be computed from:

$$
\begin{equation*}
\sigma=\sqrt{D^{2}(\eta)} \tag{33}
\end{equation*}
$$

Variance of measuring points deviation from the 2d linear regression line
In 2d linear regression, variance according to formula (32) and standard deviation for distances of points $P_{i}\left(x_{i}, y_{i}\right)$ according to formula (33) are calculated for the regression slope angle $\alpha$ computed according to formula
(5). The formulae describing coefficients $c f x$ and $c f$ should use angle $\alpha$ computed as recommended above.

Variance of measuring points deviation from the $2 d$ orthogonal regression line by Lagrange

For the 2 d orthogonal regression line by Lagrange, the variance by formula (32) and the standard deviation by formula (33) for points $P_{i}\left(x_{i}, y_{i}\right)$ can be obtained from the matrix R2 $(2 \times 2)$, which results from subroutines NROOT and EIGEN. One of the column vectors of matrix R2 describes direction cosines of orthogonal regression line.

Let it be a vector located in the second column of matrix R2. Then, the regression line slope will be available from the following formula:

$$
\begin{equation*}
\alpha=\operatorname{arcos}(\mathrm{R} 2(1,2)) \tag{34}
\end{equation*}
$$

The formulae describing coefficients $c f x$ and $c f$ should use angle $\alpha$ computed according to expression (34).

Variance of measuring points deviation from the $2 d$ orthogonal regression line by Linnik

For the 2d orthogonal regression line by Linnik, the variance given by formula (32) and the standard deviation given by formula (33) for points $P_{i}\left(x_{i}\right.$, $y_{i}$ ) can be obtained for the regression slope angle $\alpha$ computed according to formula (16). The formulae describing coefficients $c f x$ and $c f$ should use angle $\alpha$ computed as recommended above.

## 6. Variance for deviations of measuring points from 3d line for orthogonal regression by Lagrange

The components of the direction vector for orthogonal regression line 3d can be determined from three columns of matrix $R 3(3 x 3)$ including components of direction vectors for lines corresponding to eigenvalue $\lambda_{i}$, $i=1,2,3$ (subroutine NROOT [9]). The sought vector of orthogonal regression line corresponds to the smallest eigenvalue $\lambda_{i}$. Let us denote this vector as follows: $\vec{b}=[R 3(1, i), R 3(2, i), R 3(3, i)]$, where $i$ is column number. Components of vector $\vec{b}$ are direction cosines calculated with respect to axes $\mathrm{OX}, \mathrm{OY}$ and OZ .

As the 3d line passes through mass center ( $\bar{x}, \bar{y}, \bar{z}$ ) of the set of points $P_{i}\left(x_{i}, y_{i}, z_{i}\right)$ [12] located in the neighbourhood of theoretical line, the dis-
tance of this point from this line can be determined form the following expression [5]:

$$
\begin{equation*}
\eta_{i}=\left|\left[\left(x_{i}-\bar{x}\right),\left(y_{i}-\bar{y}\right),\left(z_{i}-\bar{z}\right)\right] \times[R 3(1, i), R 3(2, i), R 3(3, i)]\right| \tag{35}
\end{equation*}
$$

provided that $|[R 3(1, i), R 3(2, i), R 3(3, i)]|=1$.
After having extended the cross product (35), one can present the variance $D_{i}^{2}(\eta)$ of the distance of measuring points $P_{i}\left(x_{i}, y_{i}, z_{i}\right)$ from theoretical 3d line as follows:
$D_{i}^{2}(\eta)=\frac{\sum_{i=1}^{n}\left|\begin{array}{c}\left(y_{i}-\bar{y}\right),\left(z_{i}-\bar{z}\right) \\ R 3(2, i), R 3(3, i)\end{array}\right|^{2}+\left|\begin{array}{c}\left(x_{i}-\bar{x}\right),\left(z_{i}-\bar{z}\right) \\ R 3(1, i), R 3(3, i)\end{array}\right|^{2}+\left|\begin{array}{c}\left(x_{i}-\bar{x}\right),\left(y_{i}-\bar{y}\right) \\ R 3(1, i), R 3(2, i)\end{array}\right|^{2}}{n}$
where $n$ - number of measuring points $P_{i}\left(x_{i}, y_{i}, z_{i}\right)$.
Standard deviation $\sigma$ of measuring points $P_{i}\left(x_{i}, y_{i}, z_{i}\right)$ can be calculated as previously shown:

$$
\begin{equation*}
\sigma=\sqrt{D^{2}(\eta)} \tag{37}
\end{equation*}
$$

## 7. Summary of numerical results

The presented algorithms were tested for data acquired from measurements carried out in the CNC operating cycle of the co-ordinate measuring machine ZEISS Vista equipped with the trigger probe TP2 manufactured by Renishaw company. Basic software supplied for Vista machine, named as Calypso v. 3.5.04 was used.

Table 1.
Results of regression line parameter estimation

| Calculations according to: | CALYPSO Carl Zeiss | Elaborated software |
| :--- | :--- | :--- |
| Line passing through point $[\mathrm{mm}]:$ | $\mathrm{x}=6.0902, \mathrm{y}=3.6760$ | $\mathrm{x}=22.0681, \mathrm{y}=10.4633$ |
| Line slope [deg]: | $\mathrm{K} 1=\mathrm{Y} / \mathrm{X}=26.7387$ | $\alpha=26.7387$ |
| Standard deviation $[\mathrm{mm}]:$ | $\mathrm{s}=0.0023$ | $\sigma=0.000098$ |

The results obtained using Calypso were compared with the results obtained from computations basing on the presented relationships. The comparative specification is given in Table 1. In the course of computations made
basing on the elaborated software, compatibility of the line slope angle $\alpha$ results with the results basing on the Calypso Carl Zeiss was confirmed up to .0001 . There are differences as to location of passing point. According to a standard [12], this point should coincide with mass centre, and the Calypso software does not meet this standard. It should be improved in the subsequent versions of the Calypso.

## 8. Summary

CMM software has been elaborated by teams of mathematicians, metrology specialists and programmers for various assumptions using more or less sophisticated software. Therefore, measuring capabilities and accuracy of the CMM software vary from product to product. Inaccuracy of CMM computations is an important component of measuring uncertainty. CMM manufacturers usually do not give any details concerning the type of software, particularly approximation technique used or its accuracy. CMM users are left with two possibilities of verifying accuracy of calculations performed using the standard CMM software.

1) Method consisting in checking the standard software by a master, dimensions of which (linear, angular, shape or positional error levels) are entered from the keyboard and CMM measurements are bypassed
2) Method consisting in inspection of a geometric element on the CMM, and then processing the obtained data by self-made verification software.
The parameter of standard deviation $\sigma$ entered in calculations for describing the deviation of points from line is a good measure for accuracy of approximation software. The value of this parameter should be around $10 \%$ of measuring uncertainty value. The software which does not meet this estimation should be replaced by an improved software.

Presented algorithms for approximation of 2d and 3d lines were coded and compiled in FORTRAN 77 computer language and were named the sigmroot. The software has been tested at the Warsaw University of Technology and was based on data acquired from measurements carried out in the CNC operating cycle of the co-ordinate measuring machine Vista manufactured by Carl Zeiss. The presented software can be used for extending measuring capabilities of imperfect CMM programming.

## REFERENCES

[1] CALYPSO, Software für Koordinatenmesstechnik, Benutzerhandbuch. Version 2/3, Carl Zeiss, 2000.
[2] Demidowicz B. P.: Osnovy Wyczislitielnoj Matematiki. Izdatielstvo Nauka, Moskva, 1970.
[3] Gussak A. A.: Posobie k reszeniu zadacz po vysszej matematikie. Izdatielstvo B.G.U., Minsk, 1973.
[4] Linnik J. W.: Metoda najmniejszych kwadratów i teoria opracowania obserwacji, [The least squares method and theory of observation processing]. Warszawa, 1962, PWN.
[5] Łubowicz H.,Wieprzkowicz B.: Matematyka [Mathematics]. OWPW, Warszawa, 1999.
[6] Pfeifer T., Hemdt A.: Berechnung der Basiselemente und die Tasterkompensation in der Koordinatenmesstechnik. Teil 1: Basiselemente, Technisches Messen tm 57 (1990) 3, pp. 114 $\div 123$.
[7] Ralston A., Wilf H. S.: Mathematical Methods for Digital Computers. John Wiley \& Sons, Inc., Publishers, Copyright 1960, Part II, Chapter 7.
[8] Ratajczyk E.: Współrzędnościowa Technika pomiarowa, Maszyny i Roboty Pomiarowe [Modern technique of co-ordinate measurement, measuring machines and robots]. OWPW, Warszawa 1994.
[9] System/360 Scientific Subroutine Package (360A-CM-03X) Version III, Programmer's Manual, Fourth Edition IBM, Technical Publications Department, New York 10604, 1969.
[10] Swornowski P.: Określenie dokładności oprogramowań wspórrzędnościowych maszyn pomiarowych [Determining accuracy of co-ordinate measuring machines. Proc. of II Int'l Conf. Manufacturing 04], Zeszyty Naukowe Politechniki Poznańskiej, Budowa Maszyn i Zarządzanie Produkcja, 1, 2004, II Międzynarodowa Konferencja ‘MANUFACTURING’ 04, Współczesne problemy projektowania i wytwarzania. Poznań, 2004, Wydawnictwo Politechniki Poznańskiej, pp. 353 $\div 359$.
[11] Trajdos T.: Matematyka [Mathematics]; cz. III, WNT, Warszawa, 1981.
[12] Standards PN - EN ISO 10360 - 6: Specyfikacje geometrii wyrobów (GPS) [Geometric Part Specification]. Badania odbiorcze i okresowe współrzędnościowych maszyn pomiarowych (CMM). Część 6: Szacowanie błędów przy wyznaczaniu elementów skojarzonych metodą najmniejszych kwadratów (Gaussa). PKN Polska Norma, wrzesień 2003.

## Aproksymacja prostą w oprogramowaniu współrzędnościowej maszyny pomiarowej

## Streszczenie

Oprogramowanie wspórrzędnościowych maszyn pomiarowych (CMM) przybliża twory geometryczne na płaszczyźnie (2D) i w przestrzeni (3D) algorytmami aproksymacyjnymi, gdy liczba punktów pomiarowych jest większa od matematycznie najmniejszej liczby punktów niezbędnych do ich zdefiniowania. W artykule omówiono metody regresji ortogonalnej definiowania prostych 2D i 3D i porównano je z metodami definiowania prostych metodami regresji prostoliniowej. Przewage prostych regresji ortogonalnej nad prostymi regresji prostoliniowej wykazano poprzez porównanie wariancji odchyleń położenia punktów pomiarowych względem obu typów prostych. Algorytmy regresji ortogonalnej pozwalają ustalić optymalne położenie prostej, przy którym wariancja odległości punktów pomiarowych względem tej prostej przyjmuje wartość najmniejszą. Testowanie przeprowadzono na CMM VISTA firmy Carl Zeiss.


[^0]:    - Warsaw University of Technology, al. Niepodleglości 222, 00-663 Warsaw; E-mail:ld@meil.pw.edu.pl

