# EXTENDING THE LINEAR LEAST SQUARES PROBLEM FOR ORBIT CORRECTION IN CIRCULAR ACCELERATORS

C. Scafuri, Sincrotrone Trieste S.C.p.A.


*Abstract*

A method for extending the linear least squares problem applicable for correcting the orbit of circular accelerators is proposed. The method is based on the definition of a suitable cost function which weighs both orbit deviations and the correction effort, that is steerer kicks. The paper presents the full derivation of the formulas and the results of simulations. The application of this method for the Global Orbit Feedback system of the Elettra storage ring is being evaluated.


## ORBIT CORRECTION AT ELETTRA

Orbit correction in circular accelerators is a well understood since it can be formulated in terms of linear algebra with very good adherence to experimental results. Two main correction strategies are routinely used at Elettra: local closed bumps and global corrections by means of Singular Values Decomposition (SVD). Both are well known and need no further explanation.

The studies of the new Global Orbit Feedback system[1] led us to investigate other orbit correction techniques; in particular we want to test a class of algorithms that can offer a balance of orbit control of the and correction effort.

## EXTENDED LEAST SQUARES ALGORITHM

### *Definition of terms and hypothesis*

Let $\boldsymbol{\epsilon}=(\epsilon_1,\ldots,\epsilon_n)$ be the vector of readings of the distorted orbit, where $n$ is the number of beam position monitors.

Let $\boldsymbol{x}=(x_1,\ldots,x_m)$ be the vector of corrector kicks, where $m$ is the number of correctors.

We suppose that the response of the beam orbit to the corrector kicks is linear; then it can be described with the machine *response matrix* $\boldsymbol{R}$.

Let $\boldsymbol{y}=(y_1,\ldots,y_n)$ be the vector of orbit displacements due to a kick $\boldsymbol{x}$: $\boldsymbol{y}=\boldsymbol{R}\,\boldsymbol{x}$ with $\boldsymbol{R}:[n\times m]$

### *Correction strategy*

In order to find an optimal correction kick we form an *objective* or *cost function*. We choose to form a cost function in the form of a quadratic function. We weight the cost of two contributions:

a) the residual orbit distortion:

$$\boldsymbol{r}=\boldsymbol{\epsilon}+\boldsymbol{y}=\boldsymbol{\epsilon}+\boldsymbol{R}\,\boldsymbol{x} \tag{1}$$

b) the correction effort : $\boldsymbol{x}$

Point b), coming from ideas used in management science and optimal control theory, is the novel point. Using a quadratic cost function helps us in the calculation of differentials; on the other hand it also means that we try to find a compromise between the residual *orbit rms* and the *kicks rms*.

The cost of the residual orbit distortion is defined as:

$$J_1=\boldsymbol{r}^T\boldsymbol{W}\,\boldsymbol{r} \tag{2'}$$

where $\boldsymbol{W}$ is a positive definite n by $n$ matrix. $J_1$ is a real number. The hypothesis on $\boldsymbol{W}$ insures that $J_1>0$ .

The cost of the correction effort is defined as:

$$J_2=\boldsymbol{x}^T\boldsymbol{K}\,\boldsymbol{x} \tag{2''}$$

where $\boldsymbol{K}$ is a semi positive definite $m$ by $m$ matrix. $J_2$ is a real number. The hypothesis on $\boldsymbol{K}$ insures that $J_2\geq 0$.

The total cost is thus:

$$J=J_1+J_2=\boldsymbol{r}^T\boldsymbol{W}\,\boldsymbol{r}+\boldsymbol{x}^T\boldsymbol{K}\,\boldsymbol{x} \tag{3}$$

The $\boldsymbol{W}$ and $\boldsymbol{K}$ matrices allow us to assign different weights to the various components of the residual distortion and correction effort.

### *Derivation of correction formula*

Formula (3), from (1), can be written as:

$$J=(\boldsymbol{\epsilon}+\boldsymbol{R}\,\boldsymbol{x})^T\boldsymbol{W}\,(\boldsymbol{\epsilon}+\boldsymbol{R}\,\boldsymbol{x})+\boldsymbol{x}^T\boldsymbol{K}\,\boldsymbol{x}$$

$$J=(\boldsymbol{\epsilon}^T\boldsymbol{W}+\boldsymbol{x}^T\boldsymbol{R}^T\boldsymbol{W}\,)(\boldsymbol{\epsilon}+\boldsymbol{R}\,\boldsymbol{x})+\boldsymbol{x}^T\boldsymbol{K}\,\boldsymbol{x}$$

$$J=\boldsymbol{\epsilon}^T\boldsymbol{W}\,\boldsymbol{\epsilon}+\boldsymbol{x}^T\boldsymbol{R}^T\boldsymbol{W}\,\boldsymbol{\epsilon}+\boldsymbol{\epsilon}^T\boldsymbol{W}\,\boldsymbol{R}\boldsymbol{x}+\boldsymbol{x}^T\boldsymbol{R}^T\boldsymbol{W}\,\boldsymbol{R}\,\boldsymbol{x}+\boldsymbol{x}^T\boldsymbol{K}\,\boldsymbol{x} \tag{4}$$

Now we must find the $\boldsymbol{x}$ that minimizes the objective function. In order to perform the differentiation, we should remember some results about the differentiation of quadratic forms (matrix $\boldsymbol{A}$ is square):

$$\frac{\partial}{\partial\boldsymbol{x}}(\boldsymbol{x}^T\boldsymbol{A}\,\boldsymbol{x})=(\boldsymbol{A}+\boldsymbol{A}^T)\,\boldsymbol{x} \tag{5'}$$

$$\frac{\partial}{\partial\boldsymbol{x}}(\boldsymbol{x}^T\boldsymbol{A}\,\boldsymbol{y})=\boldsymbol{A}\,\boldsymbol{y} \tag{5''}$$

$$\frac{\partial}{\partial\boldsymbol{y}}(\boldsymbol{x}^T\boldsymbol{A}\,\boldsymbol{y})=\boldsymbol{A}^T\boldsymbol{x} \tag{5'''}$$

From (4) and (5), we get:

$$\frac{\partial J}{\partial\boldsymbol{x}}=\boldsymbol{R}^T\boldsymbol{W}\,\boldsymbol{\epsilon}+\boldsymbol{R}^T\boldsymbol{W}^T\boldsymbol{\epsilon}+(\boldsymbol{R}^T\boldsymbol{W}\,\boldsymbol{R}+\boldsymbol{R}^T\boldsymbol{W}^T\boldsymbol{R})\boldsymbol{x}+(\boldsymbol{K}+\boldsymbol{K}^T)\boldsymbol{x} \tag{6}$$

In order to have a minimum, the Hessian matrix of the objective function must be a definite positive matrix:

$$H = \frac{\partial^2 J}{\partial \boldsymbol{x}^2} = \boldsymbol{R}^T \boldsymbol{W} \boldsymbol{R} + \boldsymbol{R}^T \boldsymbol{W}^T \boldsymbol{R} + \boldsymbol{K} + \boldsymbol{K}^T > 0 \qquad (7)$$

The optimal solution is given by the $\boldsymbol{x}$ that zeroes (6), that is, by solving:

$$\frac{\partial J}{\partial \boldsymbol{x}} = 0 \qquad (8)$$

If we can find the two weight matrices $\boldsymbol{W}$ and $\boldsymbol{K}$ satisfying (7), then we can solve (8), the solution is:

$$\boldsymbol{x} = -(\boldsymbol{R}^T \boldsymbol{W} \boldsymbol{R} + \boldsymbol{R}^T \boldsymbol{W}^T \boldsymbol{R} + \boldsymbol{K} + \boldsymbol{K}^T)^{-1} \boldsymbol{R}^T (\boldsymbol{W} + \boldsymbol{W}^T) \boldsymbol{\epsilon} \qquad (9)$$

The various optimal solutions are found by varying the $\boldsymbol{W}$ and $\boldsymbol{K}$ matrices.

The existence of a solution for (9) requires that *det(**H**)* is not null. The requirement that $\boldsymbol{H}$ is positive definite is stronger.

Notice that for $\boldsymbol{W}=\boldsymbol{I}$ and $\boldsymbol{K}=\boldsymbol{0}$ we find the classical linear least squares problem. But in this case we have more insight in choosing the possible trade-offs.

# APPLICATION TO ORBIT CORRECTION IN A SYNCHROTRON RADIATION SOURCE

In this type of machines one is interested in the stability of the photon beam at the entrance of the beam line. The electron beam at the centre of the corresponding insertion device straight section must be corrected in both position and angle, with more importance to the angle due to the amplification determined by length of the optical path. If we ideally assume that we have 2 BPMs at the ends of the $k_{th}$ straight section of length $l$, and that there are no other magnetic elements that modify the beam orbit between them, we can calculate the beam position $p$ and angle $\alpha$ from the BPM readings:

$$p = \frac{\epsilon_k}{2} + \frac{\epsilon_{k+1}}{2} \qquad \alpha = \frac{-\epsilon_k}{l} + \frac{\epsilon_{k+1}}{l}$$

or, in matrix form:

$$\begin{vmatrix} p \\ \alpha \end{vmatrix} = \begin{vmatrix} 1/2 & 1/2 \\ -1/l & 1/l \end{vmatrix} \begin{vmatrix} \epsilon_k \\ \epsilon_{k+1} \end{vmatrix} \qquad (10)$$

We now have the *form* of a suitable sub matrix to be inserted into the $\boldsymbol{W}$ matrix in order to take into account the cost of photon beam position. The sub matrix is inserted at position (*k,k*).

## *Simulation results*

We have carried out some simulations using the measured horizontal response matrix of the Elettra storage ring. For our tests we wanted to have a stricter control on the position and angle of the orbit at the center of the straight sections between BPM number 43 and 44; for this purpose, following formula (10), we empirically set the corresponding $\boldsymbol{W}$ matrix elements to: $w_{43,43}$=10, $w_{43,44}$=10, $w_{44,43}$=-100, $w_{44,44}$=100. Simulations have been carried out using 3 strategies to build the rest of the weight matrix $\boldsymbol{W}$:

Case 1): $w_{i,j}$=0. See Fig.1. The goal is to control the orbit only between the 2 said BPMs.

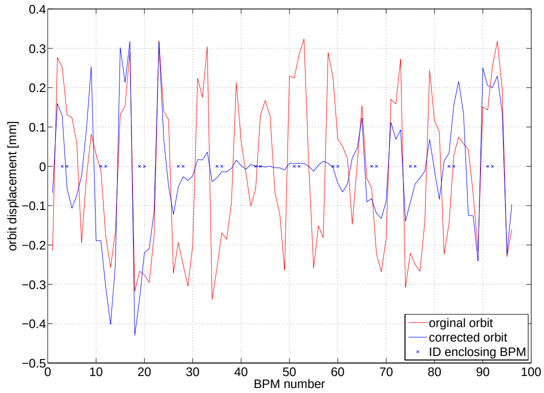


Figure 1: local orbit control

Case 2): as 1) but with $w_{h,h}$=10, where h are the indexes of BPM at the sides of an insertion device. See Fig.2. The goal is to control the orbit only in correspondence of the insertion device locations.

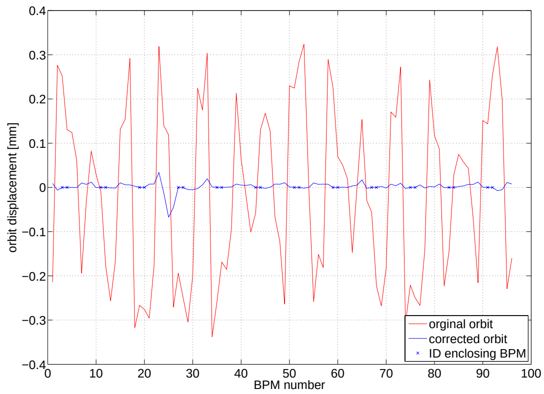


Figure 2: orbit control at ID locations

Case 3): as 1) but with $w_{k,k}$=1. See Fig.3. The Goal is to control the orbit globally.

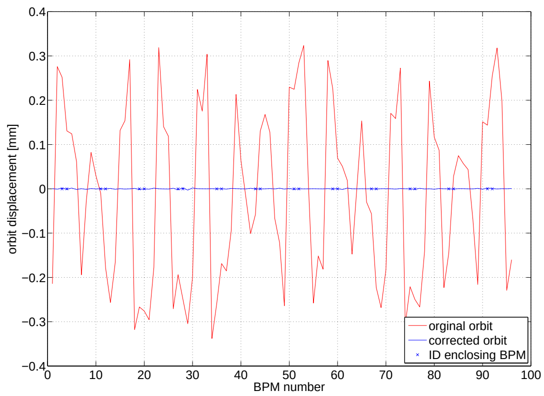


Figure 3: gloal orbit control

In all 3 cases the correction cost matrix $\boldsymbol{K}$ was set to $\boldsymbol{K}=\boldsymbol{I} * 100$.

The results show that the proposed correction scheme allows us to shape the global correction and concentrate it in some zones of the ring. The correction effort, measured by the rms of the required kicks is for the 3 cases:

1) 7 mA
2) 19 mA
3) 127 mA

As expected the correction effort grows as we increase the constraint on the orbit, but remains always very reasonable and well within the limits of our correctors.

## CONCLUSION

The proposed algorithm behaves as expected. It is a global correction scheme, but has the added benefit that we can that we can very clearly and intuitively distribute the correction effort. It will be tested as an operational correction scheme with the new Global Orbit Feedback.

## APPENDIX

### *Demonstration of formulas (5)*

Let us start from the expression of a generic bilinear form:

$$J=\boldsymbol{x}^T \boldsymbol{A}\, \boldsymbol{y} \qquad \boldsymbol{A}=[\, m\times n]$$

We can also write the bilinear form as:

$$\boldsymbol{J}=\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} x_i y_j$$

The partial derivatives respect to vector $\boldsymbol{y}$ are:

$$\frac{\partial J}{\partial y_{\bar{j}}}=\sum_{i=1}^{m} a_{i\bar{j}} x_i$$

The gradient vector is thus:

$$\left|\begin{matrix}\frac{\partial J}{\partial y_1}\\ \dots\\ \frac{\partial J}{\partial y_n}\end{matrix}\right|=\boldsymbol{x}^T \boldsymbol{A}$$

Similarly, the partial derivatives respect to vector $\boldsymbol{x}$ are:

$$\frac{\partial J}{\partial x_{\bar{i}}}=\sum_{j=1}^{n} a_{\bar{i}j} y_j$$

The gradient vector is thus:

$$\left|\begin{matrix}\frac{\partial J}{\partial x_1}\\ \dots\\ \frac{\partial J}{\partial x_m}\end{matrix}\right|=\boldsymbol{A}\, \boldsymbol{y}$$

If $m=n$ we have: $\boldsymbol{x}^T \boldsymbol{A}=\boldsymbol{A}\, \boldsymbol{x}$ So we can write:

$$\frac{\partial}{\partial \boldsymbol{y}}(\boldsymbol{x}^T \boldsymbol{A}\, \boldsymbol{y})=\boldsymbol{A}^T \boldsymbol{x} \qquad \frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T \boldsymbol{A}\, \boldsymbol{y})=\boldsymbol{A}\, \boldsymbol{y}$$

For quadratic forms:

$$J=\boldsymbol{x}^T \boldsymbol{A}\, \boldsymbol{x} \qquad \boldsymbol{A}=[\, n\times n]$$

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T \boldsymbol{A}\, \boldsymbol{x})=(\boldsymbol{A}+\boldsymbol{A}^T)\boldsymbol{x}$$

## REFERENCES


[1] M. Lonza et al. “Status of the Elettra Global Orbit Feedback Project”, this proceedings.