DESY HERA 93-07
July 1993

Trajectory Correction with Sequential Filtering

R. H. Manukian*

Deutsches Elektronen-Synchrotron DESY
Hamburg, Germany

*Permanent address: Yerevan Physics Institute, Armenia

Trajectory Correction with Sequential Filtering

R.H.Manukian
Yerevan Physics Institute, Armenia

1 Introduction

With the increase of the sizes of the charged particles transport lines the manual correction of the particle trajectories in the lines becomes inefficient. This is the reason that the automatic trajectory correction programes are now developed which make the beam losses in the transfer lines minimal. The described program is developed for the control of proton and electron beam transfer lines from *PETRA* to *HERA* at *DESY.*

We use the state space formalism for the description of the beam dynamics in the transfer line and the methods of *Sequential Filtering* are used for estimating the control kicks for the trajectory correction. We show, that the application of these methods gives a flexible tool for accelerator control and identification. The structure of algorithms is general and can be used for different requirements upon the control problems.

These algorithms and their applications are described in [2] and [3]. Here we give the survey of algorithms and the description of the program.

2 The Mathematical Background

The state space formalism and sequential algorithms are well developed in control and estimation theory [1]. The sequential filtering algorithms handle dynamical systems, which are described by differential (continuous systems) or difference (discrete systems) equations. The particle motion in the beamline (ring) and particle position measurement can be described by difference equations and this description gives an interesting opportunity to apply the sequential filtering algorithms for beam-kick errors, beam-focuse errors estimation.

We show the possibility of application of estimation algorithms for the trajectory correction problems.

2.1 State Space Formalism

The *state space* formalism describes the object (system) by

Z vector of output. In the accelerator it can be the reading of beam position monitors (BPM);

X vector of states. The states can be the position, angle, energy of the particle at any point of the beamline, kicks given to the beam by magnets.

These vectors are related by matrix equation

$$\mathrm{Z_i} = \mathrm{M_i X} + v_\mathrm{i} \qquad \mathrm{i} = 1, 2, \ldots, \mathrm{N} \tag{1}$$

where $\mathrm{Z_i}$ is the measured vertical and horizontal displacement of the beam on i BPM

$$\mathrm{Z_i} = \begin{bmatrix} x_i \\ y_i \end{bmatrix}.$$

X is the state vector. $\mathrm{M_i}$ is the influence matrix of X state vector on Z measurement (we assume the coupled case also). v_i is additive white noise with 0 mean value. The main problem of estimation theory is finding of estimation of system state X using measurement Z. In the case of beamline the X state vector can be the beam parameters

$$\mathrm{X} = \begin{bmatrix} x \\ x' \\ y \\ y' \\ \Delta p/p \end{bmatrix},$$

kicks in the magnets, errors in BPM, corrector's callibration coefficients.

2.2 The Cost Function

For estimating X we define the cost function. By *cost function* for estimation we mean the cost or penalty for not achieving correct estimation. The choice of cost function depends on prior knowledge about the stochastic features of X and v noise. In the following $\hat{\mathrm{X}}(\mathrm{Z})$ is the estimate of X based upon some observation Z.

1. *Least Square Cost Function* (LS).

Commonly used costs functions are the Least Square cost functions

$$\mathcal{J} = \sum_{i=1}^{N} [\mathrm{Z_i} - \mathrm{M_i X}]^\mathrm{T} \mathrm{S_i} [\mathrm{Z_i} - \mathrm{M_i X}], \tag{2}$$

where $\mathrm{S_i}$ is nonnegative definite symmetric weight matrix. The superscript T means the transpose of the vector or the matrix. The LS estimate $\hat{\mathrm{X}}_\mathrm{LS}(\mathrm{Z})$ is found from minimization of $\mathcal{J}$ and is obtained from solution of

$$\left.\frac{\partial \mathcal{J}}{\partial \mathrm{X}}\right|_{X=\hat{X}_{LS}(Z)} = 0. \tag{3}$$

For the LS estimation no prior information about X and v is required.

2. *Maximum Liklihood Cost Function* (ML).

The function P(Z|X) is probability density of the observation Z conditioned upon parameter X. In this case we desire to maximize P(Z|X) with respect to X. $\hat{X}(Z)$ is obtained from

$$\left.\frac{\partial P(Z|X)}{\partial X}\right|_{X=\hat{X}_{ML}(Z)} = 0 \tag{4}$$

and is estimate of X which most likely coused Z to occure. If P(Z|X) is Gaussian then maximization of P(Z|X) is equivalent to minimization of cost function

$$\mathcal{J} = \sum_{i=1}^{N} [Z_i - M_i X]^T R_i^{-1} [Z_i - M_i X]. \tag{5}$$

Here $R_i = \mathcal{E}[v_i, v_i^T]$ is the covariance matrix of v noise. $\mathcal{E}$ means the mean value of random number. For Maximum Liklihood estimation we need information about statistical features of v white noise.

3. *Maximum A Posteriori Cost Function* (MAP).

The other possible cost function can be conditional density function P(X|Z). Again $\hat{X}(Z)$ is obtained by maximizing P(X|Z). It is the solution of equation

$$\left.\frac{\partial P(X|Z)}{\partial X}\right|_{X=\hat{X}_{MAP}(Z)} = 0 \tag{6}$$

and is called the maximum a posteriori estimate (MAP). We need more statistical information to accomplish the intended maximization. Specifically, we need the density of X. We assume it to be Gaussian with $\mathcal{E}[X]$ mean and P_0 covariance matrix. The maximizing of P(X|Z) with respect to the choice of X is equivalent to minimizing

$$\mathcal{J} = \Delta X_0^T P_0^{-1} \Delta X_0 + \sum_{i=1}^{N} [Z_i - M_i X]^T R_i^{-1} [Z_i - M_i X]. \tag{7}$$

Here $\Delta X_0 = \mathcal{E}[X] - X$ is the error of estimation. For the MAP estimation we need knowledge about the covariance matrices P_0 and R. In fact the matrix P_0 shows the uncertainty of our knowledge about the value of X.

The mathematical form derived for MAP cost function includes those derived for LS and ML cost functions. One of these cost functions can be used depending on the knowledge about the state and noise. The MAP cost function can be used if matrices P_0 and R are known. The MAP cost function mathematicaly reduces to ML cost function if no prior information about X state is available i.e. $P_0 \to \infty$ or $P_0^{-1} \to 0$. The LS cost function is obtained from MAP cost function if $P_0^{-1} \to 0$ and no information about matrix R is available. In that case insted of matrix R_i^{-1} any nonegative definite symmetric weight matrix S_i can be used.

2.3 Sequential Filtering

So far we described the cost functions. Now we shall describe the algorithms for obtaining the estimation of X by minimizing functions (2) (5) (7). We use the algorithms based on sequential filtering. For obtaining the optimum solution these filters need information about mean value and deviation of X state and v noise. Since we use less information the estimation will be suboptimal, but the general structure of the filter is preserved and it is easy to get the optimum estimate if more information is available.

2.3.1 The Filter

We start from the solution of LS or ML case. We do trajectory measurement on K BPM. This is described by set of equations

$$Z_i = M_i X + v_i \qquad i = 1, 2, \ldots, K \tag{8}$$

From the data Z_i from all K BPM it is possible to calculate X unknown vector by solving equ.(3)

$$\hat{X}^K = [\mathcal{M}_K^T \mathcal{R}_K^{-1} \mathcal{M}_K]^{-1} \mathcal{M}_K^T \mathcal{R}_K^{-1} \mathcal{Z}_K$$

where

$$\mathcal{Z}_K = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_K \end{bmatrix}, \quad \mathcal{M}_K = \begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_K \end{bmatrix}, \quad \mathcal{R}_K = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_K \end{bmatrix}.$$

Here $\hat{X}^K$ means that the estimate of X is obtained from K measurements. We then consider that an additional measurement Z_{K+1} is added and find the correction $\Delta\hat{X}^{K+1}$ which updates the estimate $\hat{X}^K$, given information up to the K^{th} measurement. The correction is given by [1]

$$\Delta\hat{X}^{K+1} = P_{K+1} M_{K+1}^T R_{K+1}^{-1} (Z_{K+1} - M_{K+1}\hat{X}^K),$$

and this yields the sequential estimate

$$\hat{X}^{K+1} = \hat{X}^K + P_{K+1} M_{K+1}^T R_{K+1}^{-1} (Z_{K+1} - M_{K+1}\hat{X}^K),$$

$$\tag{9}$$

$$P_{K+1} = P_K - P_K M_{K+1}^T [M_{K+1} P_K M_{K+1}^T + R_{K+1}]^{-1} M_{K+1} P_K,$$

and

$$P_K = [\mathcal{M}_K^T \mathcal{R}_K^{-1} \mathcal{M}_K]^{-1}.$$

If X vector consists of n unknowns and Z_i is just measurement of horizontal or vertical displacement of the trajectory, M_i is $1 \times n$ vector and R_i is scalar,

$$M_{K+1} P_K M_{K+1}^T + R_{K+1}$$

is scalar too and no matrix inversion is necessary. Anyway this derivation is general and not restricted to scalar $\mathbf{R_i}$ however.

The equ.(9) is the sequential filter. For obtaining the filter's algorithm we assume that K measurements have already been done. We can use this algorithm if $K = 0$. In that case the matrix $\mathbf{P_0}$ is the covariance matrix described for MAP cost function and $\hat{\mathbf{X}}^0 = \mathcal{E}[\mathbf{X}]$. If these values are known equ.(9) gives the MAP estimate of $\hat{\mathbf{X}}$. Again equ.(9) includes the algorithms for ML and LS estimates as it was described for the cost functions.ML estimate will be obtained if no information about $\mathbf{X}$ unknown vector is available. In that case we put $\hat{\mathbf{X}}^0 = 0$ into equ.(9) and $\mathbf{P_0} \to \infty$. In numerical realization matrix $\mathbf{P_0}$ is a diagonal matrix, which diagonal elements are chosen to be $10^2 - 10^4$ order of magnitude of expected values of X. If $\mathbf{P_0} \to \infty$ and R is not known, any nonegative definite symmetric weight matrix can be used instead of $\mathbf{R_i}$ and LS estimate will be obtained.

2.4 Trajectory Correction with the Filter

Now we shall show the application of the filter algorithm (9) for the trajectory correction.

We describe the influence of the correctors kicks on the trjectory measurement on i BPM by equation

$$\mathrm{Z_i} = \mathrm{M_i X} + v_\mathrm{i} \tag{10}$$

$\mathbf{Z_i}$ is the measured horizontal or vertical beam displacement on i BPM. $\mathbf{X}$ is the vector of correctors kicks. $\mathrm{M_i}$ is the influence of X vector on trajectory on i BPM. v_i is the noise on i BPM.

The trajectory or closed orbit control problem is the minimaization of measurement $\mathbf{Z_i}$ with the help of correctors. One of the possible ways is the minimization of measurement r.m.s. value i.e. the minimization of the function

$$\mathcal{J} = \sum_{i=1}^{N} [\mathrm{Z_i} - \mathrm{M_i X}]^\mathrm{T} [\mathrm{Z_i} - \mathrm{M_i X}] \tag{11}$$

with respect to the choice of X. This is the LS cost function, where $\mathbf{S_i}$ is unit matrix. For generality we can include the weight matrix $\mathrm{S_i}$ or $\mathrm{R_i^{-1}}$. For obtaining stable solution, i.e. solution for the kicks which have not big deviations, we desire to have set of kicks with minimum r.m.s. value. With these assumptions the function (11) will have a form

$$\mathcal{J} = \mathrm{X^T P_0^{-1} X} + \sum_{i=1}^{N} [\mathrm{Z_i} - \mathrm{M_i X}]^\mathrm{T} \mathrm{R_i^{-1}} [\mathrm{Z_i} - \mathrm{M_i X}]. \tag{12}$$

It is obviouse that this function has the same form as cost function (7) obtained for MAP estimation, if $\mathcal{E}[\mathrm{X}] = 0$. Here $\mathrm{P_0}$ and $\mathrm{R_i}$ are weight marices. We can use the filter (9) for the minimization of this function with respect to X. We must start from $K = 0$. The vector of the kicks $\hat{\mathrm{X}}^0 = 0$. The choice of weight matrices $\mathrm{P_0}$ and $\mathrm{R_i}$ will be described later. After the measurement of the trajectory on K^{th} BPM the optimum kicks $\hat{\mathrm{X}}^\mathrm{K}$ are calculated with (9).

2.4.1 Weight Matrices

We describe the choice of the weight matrices $\mathbf{P_0}$ and $\mathbf{R_i}$ in details, since they give the possibility of the flexible application of the filter (9). In fact these matrices are weights upon the values of the kicks and measurements respectivly. But from the other side they preserve their meaning of covariance matrices defined for MAP or ML estimators.

The cost function (12) can be derived from (7) with assumption that $\mathcal{E}[\mathbf{X}] = \mathbf{0}$. This assumption means that we vary the kicks arround 0 value and the optimum solution will have the possible minimum r.m.s. of the kicks. The matrix $\mathbf{P_0}$ shows the freedom of variation.

In the correction procedure we chose $\mathbf{P_0}$ as diagonal matrix, if the weights on the correlation between the correctors are not desired. The diagonal elements define the weights on the kicks of the certain correctors. The influence of $\mathbf{P_0}$ matrix on the result can be interpreted in different wayes, but they produce the same results, however.

If we interpret $\mathbf{P_0}$ as simple weight matrix, then the big diagonal element means that the constraint of minimum r.m.s. of kicks has less weight for the certain corrector. By making the diagonal element 0 it is possible to exclude the corresponding corrector. From the other side if we interpret $\mathbf{P_0}$ as covariance matrix of kicks which have 0 mean value, then the big diagonal element means that the corrector has big variance and it can have big changes.

So with the choice of diagonal elements of $\mathbf{P_0}$ matrix it is possible to increase or decrease the role of the correctors in the correction procedure.

It was decided to use this feature of $\mathbf{P_0}$ matrix to solve the trajectory correction problem, when the correctors' currents are limited. In the first approach all diagonal elements of $\mathbf{P_0}$ are put the same value. Then the optimum kicks are calculated. If the corrector's calculated current exceed the limit, the corresponding diagonal element in $\mathbf{P_0}$ is decreased. Then the correction procedure is repeated, until all correctors are in their limits. This procedure is heuristic, but in practice it gives reasonable result in some iterations, however.

The $\mathbf{R}$ matrix can be treated with the same iterpretations as it was done for $\mathbf{P_0}$. The diagonal elements of this matrix put weights on corresponding BPM measurements. Big diagonal element in $\mathbf{R}$ decreases the influence of corresponding BPM measurement on the result of the correction procedure. The big diagonal element means that the variance of the error on the BPM is big and that measurement must have small influence on the decision.

With these considerations elements of $\mathbf{R}$ matrix can be chosen in a way to exclude the measurement of doubtful BPM or increase the influence of those, where the accurate trajectory correction is desirable.

With the help of $\mathbf{R}$ weight matrix it is possible to solve constrained minimization problems. Sometimes local trajectory correction in beamline using arbitrary number of correctors is preferable i.e. trajectory correction in region without desturbing it out of the region. This means that X and X' must not be changed at the last BPM of the region. We shall have two constraints

$$0 = \mathbf{MX}$$

and

$$0 = \mathrm{M'X}.$$

Here X is vector of the correctors kicks. M is the vector of influence of X vector on beam position (X) on the last BPM. M$'$ is vector of influence of X vector on beam angle (X') on the last BPM. These two constraints are included into function (12) with high weights. The new cost function is

$$\begin{aligned}\mathcal{J} &= \mathrm{X^TP_0^{-1}X} + \sum_{i=1}^{N}[\mathrm{Z_i - M_iX}]^\mathrm{T}\mathrm{R_i^{-1}}[\mathrm{Z_i - M_iX}] + \\ &\quad [\mathrm{MX}]^\mathrm{T}\mathrm{R_{N+1}^{-1}}[\mathrm{MX}] + [\mathrm{M'X}]^\mathrm{T}\mathrm{R_{N+2}^{-1}}[\mathrm{M'X}]. \qquad (13)\end{aligned}$$

Here $\mathrm{R_{N+1}}$ and $\mathrm{R_{N+2}}$ have small values, which means that the satisfuction of the two constraints has big priority. This function can be written in the form of function (12)

$$\mathcal{J} = \mathrm{X^TP_0^{-1}X} + \sum_{i=1}^{N+2}[\mathrm{Z_i - M_iX}]^\mathrm{T}\mathrm{R_i^{-1}}[\mathrm{Z_i - M_iX}]. \qquad (14)$$

Here $\mathrm{Z_{N+1}} = 0$, $\mathrm{Z_{N+2}} = 0$ and $\mathrm{M_{N+1}} = \mathrm{M}$, $\mathrm{M_{N+2}} = \mathrm{M'}$. And (14) function can be easily minimized with filter (9).

3 The Description of the Program

The program for the correction of the trajectory is developed for the control of proton line from *PETRA* to *HERA* at *DESY*.

The length of this beam transfer line is $\sim$ 400m. The beam trajectory is measured with the help of 20 screen monitors. The measurements are giving the beam center coordinates with respect to the vacuum chamber. For the correction procedure the program is using 19 vertical correctors, 20 horizontal correctors, 2 septum magnets at *PETRA* and 3 power supplies, which control 30 main bending magnets in the transfer line.

The program has two functional parts: Simulation and Correction.

3.1 Simulation

This module of the program simulates the trajectory with the kicks on the correctors, septums and main bending magnets. So the operator can have an idea about the influence of the certain change on the actual trajectory before doing changes in the power supplies.

3.2 Correction

The correction procedures are based on the algorithms described above. Here the advatages of the sequential algorithms become obvious. This algorithms do not require the matrix inversion and they can treat both underdetermined and overdetermined problems.

In the correction procedure the trajectory is minimized with respect to reference (desired) trajectory. The reference trajectory is the one which was obtained empirically to be the best. Sometimes it is better to steer the beam not to the centre of the screen monitor, but to other position of the screen monitor, which will result in better beam transparency threw the beam line. The reference trajectory is stored in a file and can be updated.

Since the transfer line is long, the trajectory correction region by region is preferable. The operator can choose the region where the trajectory correction is desired, do the correction and then move to the correction of the other region. It is possible to exclude certain monitors, where the information is not reliable. The certain correctors can also be disabled and they will not take part in the correction. There are four correction options available, which give the possibility of flexible operation in different conditions and requirements:

1. *Best Corrector*

The option finds one best corrector in the chosen region, which gives the minimum r.m.s value of corrected trajectory. This option can be used for diagnosing the changes in *PETRA*.

2. *Local Correction*

This option gives the kicks which minimize the trajectory in the chosen region without disturbing it out of the region. Here all correctors in the region are used. This is useful for the situations when the injection to *HERA* is optimized and it is desired to correct the trajectory in the transfer line without changing the injection conditions.

3. *Global Correction*

Here the trajectory in the chosen region is minimized with the correctors of the region. In this option the last screen monitor in the region has comparatively small weight. This will avoide of having big angle on the exit of the region.

4. *Correction with Main Bending Magnets*

Really this option is fitting. It calculates the kicks on main bending magnets for best fit with trajectory. This can be used for finding the energy mismatch.

4 Conclusions

In this paper the application of sequential algorithms for the trajectory correction is described. These algorithms are stable to measurement errors and give flexible tool for the correction procedures. Here the matrix inversion is not required. They do not need big computer memory due to sequential processing of measurements and self-learning structure.

5 Acknowledgement

The auther wish to thank Mr.Glöer and Mr.Holm for help and guidance in development of software in Pocal. The auther is indebted to Dr.J.Rossbach for continuous interest and encouragement.

References

[1] A.P.Sage "Optimum Systems Control", Prentice-Hall,Inc. Englewood Cliffs, N.J., 1968.

[2] S.H.Ananian and R.H.Manoukian "The Program for Automatic Control of Beam Transfer Lines," *DESY-M-90-15,* November 1990.

[3] S.H.Ananian and R.H.Manoukian "The Beam Steering in Beamlines and Rings Using the Methods of Sequential Filtering" Proc. 3rd European Particle Accelerator Conference, Berlin 1992, pp750-752.