Elimination of the Measurement System Impact on the Measured Data by Maximum Entropy on the Mean Method

P. Holejšovská, Ing. *), J.-F. Bercher, Prof. Assist. **), Z. Peroutka, Ing. *)**

Západočeská Univerzita v Plzni and Groupe ESIEE Paris, France

Abstract: *Quality of the diagnostic of some system strongly depends on the precision of the data obtained by the measurement. The error of the measurement can lead to the mistakes in the diagnostic process. The aim of this research is to design method that enables to avoid impact of measurement system on experimental evidence, mitigate noises and increases precision of obtained data. One of the possible solutions is the Maximum Entropy on the Mean Method (MEMM) that deals with the solving of the linear and noisy inverse problem of the form* $\mathbf{y=Ax+b}$. *Another method is Expectation-Minimization Algorithm (EM algorithm) consisting of two steps. The expectation step (E step) compute an estimator* $\hat{\mathbf{x}}$ *and is followed by the maximization step (M step) that provides new estimation. Both steps have to be iterated until convergence. In the exponential family, the E step gives similar results as the MEMM that is why we combine both methods in order to explore the advantages of both. The measured data (**y**) processed by proposed algorithm allows to eliminate the noise **b** and the influence of the measurement system properties (presented by degradation matrix **A**). Thus, this algorithm enables to reconstruct real data **x** that object produces. Importance of this method in the signal and image processing and diagnostics is obvious.*

1. MAXIMUM ENTROPY ON THE MEAN METHOD (MEMM)

The Maximum Entropy on the Mean Method (MEMM) solve the linear and noisy inverse problem of the form of $\mathbf{y=Ax+b}$, where $\mathbf{y}$ are the observations, $\mathbf{A}$ is the supposed degradation matrix, vector $\mathbf{x}$ is the measured object (typically a signal or image) and $\mathbf{b}$ is the noise which has to be estimated. The measured data processing by this algorithm allows to eliminate the noise $\mathbf{b}$ and the influence of the measurement system properties (presented by degradation matrix $\mathbf{A}$) to the measured data $\mathbf{y}$ and so to obtain real data that the object produces. Given reference measure $\boldsymbol{\mu}$ defined on the object $\mathbf{x}$ and noise $\mathbf{b}$, the MEMM consists of selecting the distribution $\hat{\boldsymbol{p}}$ which is the closest to $\boldsymbol{\mu}$ according to the Kullback distance and which satisfy a given constraint, in this case the observation equation. The MEMM estimation $\hat{\boldsymbol{x}}$ is the mean of the selected distribution $\hat{\boldsymbol{p}}$. So, it is the minimizer of a convex cost function defined on $\mathbf{x}$ and $\mathbf{b}$.

*) University of West Bohemia in Pilsen, Faculty of Electrical Engineering
Department of Applied Electronics
Univerzitní 8, 306 14 Plzeň, Czech Republic
e-mail: holejsov@kae.zcu.cz

**) École Supérieure d'Ingénieurs en Électrotechnique et Électronique (Groupe ESIEE),
. Département Traitement du signal et Télécommunications et Laboratoire Parole Signal Image (LPSI)
Cité Descartes, BP 99, 93 162 Noisy le Grand Cedex
e-mail: bercherj@esiee.fr

***) University of West Bohemia in Pilsen, Faculty of Electrical Engineering
Department of Electromechanics and Power Electronics
Univerzitní 8, 306 14 Plzeň, Czech Republic
e-mail: peroutka@ieee.org

Let have a linear inverse problem **y=Ax+b**. The observation matrix **A** is supposed to be known and some statistical characteristics of the noise **b** too. When the observation matrix **A** is not regular or ill-conditioned, the problem is ill-posed. It means the convex constraint

$$\boldsymbol{x \in C} \ , \tag{1}$$

where C is a convex set, is necessary. In some specific problems where the lower and upper bound (a_k,b_k) are known, the convex set can be defined as

$$\boldsymbol{C = \left\{ x \in R^N \, / \, x_k \in \langle a_k, b_k \rangle, k = 1..N \right\}}. \tag{2}$$

For the estimation $\hat{x}$ we select the distribution $\hat{p}$ which is the closest to reference measure **μ** according to the Kullback distance. The Kullback distance **D(p||μ)** is defined for a reference measure **μ** and probability measure **P** by

$$\mathbf{D(p} \,\|\, \mu) = \int \log \frac{dP}{d\mu} dP \,. \tag{3}$$

Thus, the MMEM method begins by the specification of the convex set **C** and the reference measure **dμ(x)** over it. The actual observations **y** are considered as the mean $\mathbf{E_p(X)}$ of the probability distribution **P** defined on **C**. The distribution **P** is selected as the minimizer of the μ-entropy submitted on the constraints of the mean $\mathbf{AE_p(x) = y}$ in the noiseless case. It means the **P** is the nearest distribution respect to the Kullback distance **D(p||μ)** to the reference measure **μ** satisfying equation $\mathbf{AE_p(x) = y}$. MEMM problem in the noiseless case is given by:

$$\begin{cases} \hat{p} = \arg\min\limits_{P} \int \log \frac{dP}{d\mu}(x) dP(x), \\ such \ that \ \ y = A\int x dP(x) \\ and \ \ x = E_p(x) \end{cases} \tag{4}$$

The solution exists if it is belongs the exponential family. For each $\boldsymbol{x \in C}$, the $\mathbf{AE_p(x) = y}$. We define the function **F(x)** as the optimum value of the Kullback distance.

$$F(x) = \operatorname*{Inf}_{P \in P_X} \mathbf{D(p} \,\|\, \mu), \ \ where \ \ P_X = \left\{ P : E_p(x) = x \right\} . \tag{5}$$

Than the problem can be posed as:

$$\begin{cases} \hat{x} = \arg\min\limits_{x} F(x), \\ such \ that \ \ y = Ax. \end{cases} \tag{6}$$

This method amounts to minimizing the convex criterion and so admits the dual formulation. The dual function is defined as

$$D(\lambda) = \lambda^t y - F^*(\lambda^t A), \tag{7}$$

and allows to calculate numerically the expectation $\mathbf{E_p(x)}$. The function $F^*(\lambda^t A)$ is the convex conjugate of the function **F(x)**. It means that dual formulation of the problem that have to be solved is to find estimator $\hat{\lambda}$ by maximizing the dual function **D(λ)**:

$$\left\{ \hat{\lambda} = \sup_{\lambda} \left(\lambda^t y - F^*(\lambda^t A) \right), \right. \tag{8}$$

The problem number one of the MEMM is to define and the function $F^*(\lambda^t A)$ and to maximize it. The problem is that in the noisy case the function $F^*(\lambda^t A)$ has to be separated:

$$F^*(\lambda^t A) = F_x^*(\lambda^t A) + F_b^*(\lambda^t). \tag{9}$$

The solution of this problem is really complicated. In the exponential family, the MEMM gives similar results as the E step of the EM algorithme. That is why we tried to use the EM algorithm to eliminate the need of the function $\boldsymbol{F}^*\left(\lambda^t \boldsymbol{A}\right)$ separation in the noisy case.

2. E-M ALGORITHM

Expectation-Minimization Algorithm (EM algorithm) consists of two steps: the expectation step followed by the maximization step. The E step computes an estimation $\hat{\boldsymbol{x}}$ using the current estimation of the parameter satisfying the conditions upon the observations. Let have **x** the data, the **y** observations, **f(x|θ)** probability density function and **θ** set of parameters of the density, then for E step we compute:

$$\boldsymbol{Q}\left(\theta \middle| \theta^{[k]}\right) = \boldsymbol{E}\left[\log \boldsymbol{f}(\boldsymbol{x} \mid \theta) \mid \boldsymbol{y}, \theta^k\right], \text{ it means } \boldsymbol{Q}\left(\lambda, \lambda^k, \boldsymbol{y}\right) = \lambda^t \boldsymbol{A}\boldsymbol{E}\left[\boldsymbol{x} \mid \boldsymbol{y}, \lambda^k\right]. \qquad (10)$$

In the exponential family, the MEMM gives similar results as the E step of the EM algorithm that is why the MEMM was used as the E step.

The M step provides new estimation.

$$\boldsymbol{Q}\left(\lambda, \lambda^k, \boldsymbol{y}\right) = \lambda^t \boldsymbol{A}\boldsymbol{E}\left[\boldsymbol{x} \mid \boldsymbol{y}, \theta^k\right], \text{ it means } \lambda^{k+1} = \underset{\lambda}{\mathbf{arg\,max}}\, \lambda^t \boldsymbol{A}\boldsymbol{E}\left[\boldsymbol{x} \mid \boldsymbol{y}, \lambda^k\right] - \boldsymbol{F}^*\left(\lambda^t \boldsymbol{A}\right). \qquad (11)$$

These two steps must be iterated until convergence occurs - it may be determined as:

$$\left\|\theta^{[k]} - \theta^{[k-1]}\right\| < \varepsilon, \text{ it means } \left\|\lambda^{[k]} - \lambda^{[k-1]}\right\| < \varepsilon. \qquad (12)$$

3. RESULTS

In the exponential family, the E step gives similar results as the MEMM that is why for the first estimation of the estimator $\hat{\boldsymbol{x}}$ the MEMM has bee used. Than the new estimation was proposed by the M step of EM algorithm and the process was iterated until convergence condition occurred. Results of the data reconstruction are well seen in the simulations below. The simulations were made to reconstruct real signal **x** (x_{true}) from the measured observations **y** ($y_{measured}$). On the right side of each figure you can see the iteration process of estimating $\hat{\boldsymbol{x}}$. The real signal **x** is presented by dotted line, last iteration of the computed signal $\mathbf{x_{ME}}$ by bold. On the left side, there is the comparison of the real signal **x** (normal) and signal $\mathbf{x_{ME}}$ (bold) computed by the MEMM with EM algorithm. Experimental evidence **y** is presented by points. The noise is supposed to be Gaussian. Various values of the σ^2 have been considered.

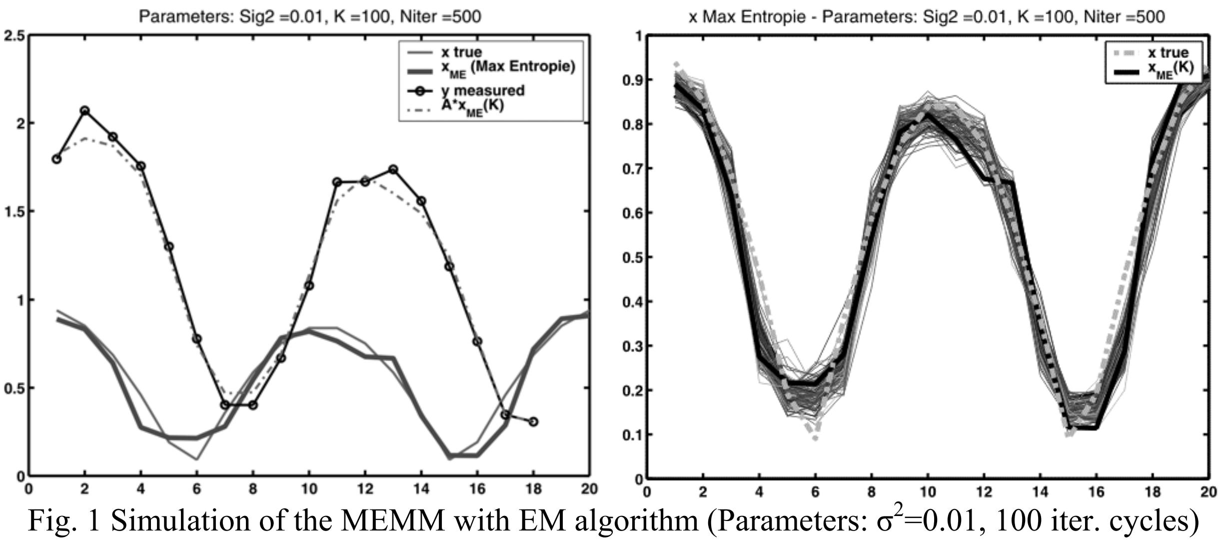

Fig. 1 Simulation of the MEMM with EM algorithm (Parameters: σ^2=0.01, 100 iter. cycles)

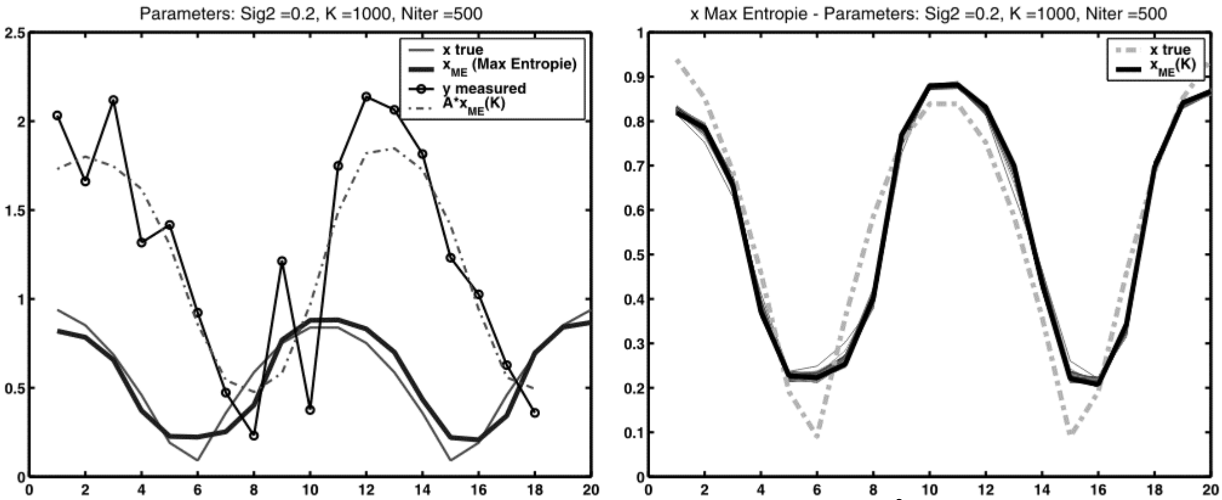

Fig. 2 Simulation of the MEMM with EM algorithm (Parameters: σ^2=0.2, 1000 iter. cycles)

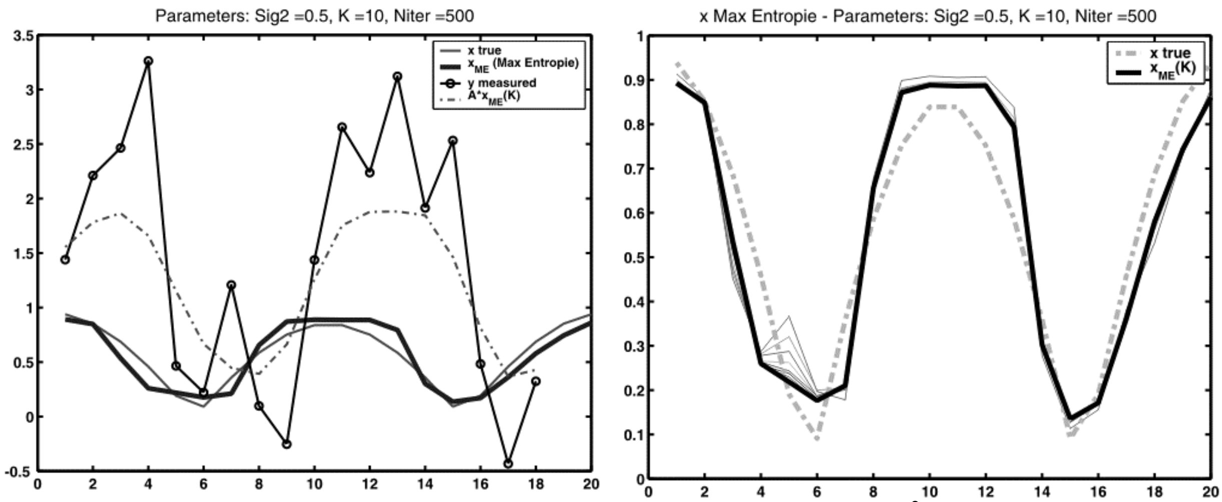

Fig. 3 Simulation of the MEMM with EM algorithm (Parameters: σ^2=0.5, 10 iter. cycles)

4. CONCLUSIONS

The proposed algorithm reconstructs the real data **x** from measured, noisy observations **y**. This method enables to mitigate the noise **b** and the measurement system impact. The problem is to find the best estimation $\hat{\boldsymbol{x}}$ of the real data and satisfy the given constraints. Because of difficulties in the computation process in the noisy case, the MEMM is used as the first step of the EM algorithm that allows to converge to the real signal **x** by the iteration process. The method is still in progress.

Acknowledgement
Research is performed at Groupe ESIEE Paris, France and is partly supported by project FRVŠ 2328/2003/G1.

References
[1] J.-F.Bercher, G.Le Besnerais, G.Demoment: *The Maximum Entropy on the Mean Method, Noise and Sensitivity.* Maximum Entropy and Bayesian Methods, Kluwer Academic Publishers, Cambridge,U.K.,1995.
[2] C.Heinrich, J.-F.Bercher, G.Demoment: *The Maximum Entropy on the Mean Method, Correlations and Implementation Issues.* In: IEEE Transactions on Information Theory,1995.
[3] Todd K. Moon: *The Expectation-Maximization Algorithme..* In: IEEE Signal Processing Magazine. November 1996, IEEE1996. p.47-60. 1053-5888/96.