Use of Non-convex Optimization to Recover Signals Distorted by Memoryless Non-invertible Sensor Nonlinearities

Sugathevan Suranthiran Dept. of Mechanical Engineering Texas A&M University College Station, TX 77843, USA suren@neo.tamu.edu

Abstract-This paper is concerned with the recovery of bandlimited signals that are distorted by non-invertible sensor nonlinearities. When the nonlinear sensor characteristic is such that the input-output relationship is invertible, the distortion caused by sensor nonlinearity may be compensated and the signals may be reproduced from the distorted sensor measurements. However, such methods may not work if the nonlinear sensor dictates ill-posed input-output behavior such as saturation and dead-band. We propose a new approach that optimizes a non-quadratic criterion to reproduce the signal from the distorted sensor data. The fact that the introduction of a non-quadratic penalty in the cost function does not alter the optimal solution as much as the quadratic penalty does, makes it a better candidate for the problem posed. However, this is a non-convex optimization problem, which implies that multiple local optimal solutions may exist. A systematic solution procedure is presented to interpret the solutions generated by non-quadratic criteria and to identify the global optimal point. Illustrative examples are presented to further explain the proposed framework.

I. INTRODUCTION

A. Problem Description

Sensors are based on mechanisms in which one physical quantity is coupled to another. For example, temperature or pressure at the input to a sensor yields an electrical signal at its output. Most of these physical mechanisms are nonlinear by nature and great effort and ingenuity are used in designing sensors to obtain a linearized regime within which a sensor is to be operated. However, in situations where a sensor is operated outside this linearized regime, the interpretation of the sensor output will be complicated. This places a major limitation on the use of such sensors and the consequences may be intolerable.

The difficult task of achieving a linear regime, which may require costly design procedures and the use of expensive materials, may be simplified or eliminated if sensor nonlinearity is considered a embedded feature of a sensor. This may enhance performance of a sensor since nonlinear input-output characteristics can expand a sensor's operating regime and alleviate the constraints imposed by linearization. For example, High Dynamic Range Complementary Metal-Oxide Semiconductor (CMOS) Sensors utilize logarithmic nonlinear sensor characteristics to increase the input dynamic range and to obtain a good noise response. Fiber optic displacement sensors, which are widely used to obtain approximate displacement measurements at a very low cost, Suhada Jayasuriya Dept. of Mechanical Engineering Texas A&M University College Station, TX 77843, USA sjayasuriya@mengr.tamu.edu

the Hercules Orthoflex capacity sensor, a biosensor that is used to measure the pressure between the foot and shoe [1], and a low cost oxygen sensor that is used for closedloop active combustion control in automobiles [2] are some examples of nonlinear sensors widely used in practice.

Despite many advantages in using nonlinear sensors, distortion caused by nonlinearity may appear at first as a factor that does not encourage its usage. However, a signal recovery setup developed by authors in [3], [4] guarantees that this distortion will not be a problem and a unique signal recovery is possible when the nonlinear sensor is designed so as to satisfy certain requirements. A detailed description of this method and its limitations are given in [5]. This method may not yield accurate results if the sensor characteristics include more severe nonlinearities such as saturation and dead-band. Saturation nonlinearity is a common defect found in most real sensors and its effects are clearly apparent when the sensor is used to pick up high amplitude signals. The main issue with such sensors is that the accurate and unique reproduction of sensor input from the saturated sensor data may not be possible, in general.

B. Motivation

Bearing in mind that retrieving at least some of the original information and eliminating or reducing the distortion caused by saturation nonlinearity are useful contributions towards a "smart sensor technology", we propose a new approach, which optimizes a non-quadratic performance index to recover the original data. The fact that the introduction of a non-quadratic penalty in the cost function does not alter the optimal solution as much as the quadratic penalty does, makes it a better candidate for the problem posed. The nature of solutions generated by non-quadratic optimization suggests its unique ability to solve several practical problems. For example, non-quadratic optimization is applied to decentralize multi-variable model predictive control structures in [6].

We further elaborate the idea of using non-quadratic optimization to solve ill-conditioned equations by considering a simple numerical example.

Consider an optimization problem that requires the following performance index to be minimized:

$$J_o = (x - a)^2 \tag{1}$$

where a is a constant and x is the variable to be optimized.

This is a very simple example that will become clearer later, in which x might correspond to an estimated or measured value and then a would correspond to the true value. Thus the optimum is clearly x = a.

For reasons to appear later, it sometimes proves advantageous in this context to augment the cost function by an amount that is a simple function of the variable, called a "penalty function". To investigate the effects of different penalty functions on the optimal solutions, let us add both quadratic and non-quadratic penalty terms to the original cost function given in Equation (1) and compare the results. We first derive the quadratic cost function:

$$J_q = (x-a)^2 + \lambda_q x^2 \tag{2}$$

whose gradient is given by

$$\nabla J_q = 2(x-a) + 2\lambda_q x \tag{3}$$

where λ_q is the weight on the quadratic penalty.

The non-quadratic cost function and its gradient are derived as follows:

$$J_n = (x - a)^2 + \lambda_n x^{0.5}$$
 (4)

$$\nabla J_n = 2(x-a) + \frac{\lambda_n}{2\sqrt{x}} \tag{5}$$

We assume that the weight, λ_n , on the non-quadratic penalty term, $\lambda_n x^{0.5}$, is reasonably low so that the global minimum does not occur at the origin. Further discussion on the choice of the weights and their effect on the optimal solutions is given in Section 3.

Referring to Equation (3), the gradient of the quadratic cost function is linear and the optimal solution is the point where the terms 2(x-a) and $2\lambda_q x$ are equal but of opposite sign. As schematically shown in Figure 1, the proximity of the optimal solution of the quadratic cost function, J_q , to the actual solution (*a* in this case) is determined by weight, λ_q .



Fig. 1. Comparison of Optimal Solutions obtained with Quadratic and Non-quadratic Cost Functions

Fig. 2. Effect of Weight on Non-quadratic Penalty on Optimal Solutions

The gradient of the non-quadratic function given in Equation (5) has a different shape as shown in Figure 1. Depending on the value of weight, λ_n , the gradient of the non-quadratic term, $\frac{\lambda_n}{2\sqrt{x}}$, has a very high value near zero

and dies out as the value of x increases. We point out that this feature of the non-quadratic function makes it a better candidate than the quadratic function. With a suitable choice of weight, λ_n , the gradient of the non-quadratic penalty can be shaped such that it dies out very soon and consequently has a very small value at the actual solution, a, which will lead to an optimal solution very near a. Figure 1 further illustrates this point. It is worth mentioning that the quadratic penalty does not possess this ability as its gradient is linear.

Another important point is that in order to drive the optimal solution closer to the actual value, the weight on the quadratic penalty must be chosen as low as possible since the accuracy of the solution is simply determined by the weight. This may not be true in the case of non-quadratic penalty. This is because the gradient of the non-quadratic penalty will decrease as x increases and a reasonably low weight may be sufficient to guarantee a small gradient value near the actual solution. It is also pointed out that the shape of the gradient of the non-quadratic penalty is such that it would possibly have optimal solutions only in two neighborhoods: one is in the vicinity of the actual solution when the weight is low and the other is at the origin when the weight on the non-quadratic penalty is high. Figure 2 further supports this point. This fact further encourages the use of non-quadratic cost function as the availability of wide range of suitable weights makes the tuning process very easy and the solution reliable.

Motivated by the above example, we tested this idea to solve the ill-posed signal recovery problem. The results, which are detailed in this paper, have been promising and the optimization of a non-quadratic cost function stands out from other available tools to generate accurate results. The class of non-quadratic functions considered in this paper are those that exhibit infinite gradient at the origin, $\{x^k | 0 < k < 1\}$.

This paper is organized as follows: Section 1 is devoted to the introductory discussion on the topic. The preliminaries and the related previous work are given in Section 2. The main results are derived in Section 3. Illustrative examples are presented in Section 4. Conclusions are drawn in Section 5.

II. PRELIMINARIES



Fig. 3. Nonlinear Sensor

Figure 3 shows the standard sensor measurement setup in which a signal y is measured through a nonlinear sensor g(.) while being corrupted by a noise signal n. Sensor output w is related to the input signal y by

$$w = g(y) + n \tag{6}$$

It is known that even though the input signal y is band-limited with a frequency band $[-\Omega, \Omega]$ the spectrum of the output signal w will not be restricted within this frequency band, in general [4]. This phenomenon is known as nonlinear distortion, which has been the major limiting factor for the successful design and the use of primary nonlinear sensors. Furthermore, nonlinear distortion will lead to inaccurate sensor readings if the sensor being used is assumed to be linear.

However, the signal recovery procedure proposed in [5] suggests that with certain nonlinear sensor characteristics, the nonlinear distortion may be completely compensated and the original signal y may be uniquely recovered. Furthermore, it is also pointed out in the same paper that in order to recover a band-limited signal, complete information about the output signal may not be needed. Only a part of the output signal that lies in the original frequency band $[-\Omega, \Omega]$ is sufficient to retrieve the original band-limited signal. This invention leads to some useful practical consequences such as band-limited sensory operation. The infinite bandwidth nonlinear sensory operation may require a high bandwidth sensor, which may be neither physically realizable nor cost-effective.

The signal recovery scheme proposed in [4] uses the setup that is schematically shown in Figure 4. In this setup, an ideal low pass filter is employed to remove the frequency components, which lie outside the frequency band of signal *y*.



Fig. 4. Signal Recovery Setup

The signal *y* is recovered by solving the following optimization problem:

$$\min_{\mathbf{y}} J = \|z - \mathcal{F}^{-1}\{\mathcal{HF}\{g(\mathbf{y})\}\}\|$$
(7)

where z(t) is the filtered signal, \mathscr{F} is Fourier Transform, \mathscr{F}^{-1} is Inverse Fourier Transform and \mathscr{H} is a low-pass filtering operation defined by

$$\mathscr{H}(w) = \begin{cases} 1 & |\omega| \le \Omega \\ 0 & \text{otherwise} \end{cases}$$

An efficient solution to the unconstrained optimization problem described by the performance index in Equation (7) can be obtained by solving the following recursive equation:

$$y_{k+1}(t) = y_k(t) + \alpha z(t) - \alpha \mathscr{F}^{-1} \{ \mathscr{HF} \{ g(y_k(t)) \} \}$$
(8)

where α is the convergence accelerator.

When the nonlinear operation g is non-singular, the above optimization problem will have a unique solution. The uniqueness condition simplifies to the following requirement:

$$0 < \frac{dg(y)}{dy} < \infty \quad or \quad -\infty < \frac{dg(y)}{dy} < 0 \quad \forall \ y \qquad (9)$$

That is, the signal *y* may be uniquely recovered from the low-pass filtered version of the nonlinear sensor output if the nonlinear sensor function is monotonic.

III. NON-QUADRATIC REGULARIZATION

The method described in the previous section will fail if the nonlinear operation is singular. To solve "ill-conditioned" problems of this kind, the standard regular-ization techniques are applied in general. Tikhonov regular-ization [7], [8] is probably the most commonly used regular-ization method. This technique augments the least-squares performance index given in Equation 7 with an additional term, generally known as a penalty, which incorporates prior information about signal *y* given by

$$\min_{\mathbf{y}} J = \|z - \mathscr{F}^{-1}\{\mathscr{HF}\{g(\mathbf{y})\}\}\|^2 + \lambda_q \|\mathbf{y}\|^2$$
(10)

where λ is a weight on the penalty term. The effect of the penalty function on the optimal solution as well as on the shape of the resultant cost function is graphically shown in Figure 5. For convenience, the variation of the cost function in only one direction is shown in the plot.



Fig. 5. Effect of adding a Quadratic Penalty Function

Fig. 6. Effect of Penalty Function Weight on the optimal solution-Quadratic Penalty Case

It is noted that the constraint enforced by means of a penalty moves the solution away from its actual value. To maintain the accuracy of the solution, the weight on the penalty λ_q must be kept at a low value. When the value of λ_q is increased gradually, the error term on the expression becomes insignificant and the optimal solution moves towards y_o , which is zero in this case. However, it can be shown that the optimal solution will never be exactly zero for a finite value of λ_q . Figure 6 supports this fact.

We have shown that adding a quadratic penalty may not improve the accuracy of the solution, in general. If the accuracy of sensor data is of paramount importance, other approaches should be sought. In this paper, we propose an approach that replaces the quadratic penalty in the cost function (10) by a non-quadratic penalty:

$$\min_{y} J = \|z - \mathscr{F}^{-1}\{\mathscr{HF}\{g(y)\}\}\|^{2} + \lambda_{n} \sum_{i=1}^{N} |y_{i}|^{k}$$
(11)

where $\{k|0 < k < 1\}$ is a non-quadratic index and *N* is the length of the discretized signal *y*. In the above the discrete version of the process is considered and the signals are assumed to be sampled sufficiently fast.

The rationale behind the choice of the non-quadratic penalty has already been explained in Section 1. We provide additional insights by showing how a non-quadratic penalty alters the shape of the cost function, which is otherwise quadratic. Figure 7 shows how this alteration is done. Again, the variation of the cost function only in one direction is shown in the plot.



Fig. 7. Effect of adding a nonquadratic penalty

Fig. 8. Effect of Penalty Function Weight on the optimal solution—Non-Quadratic Penalty Case

Referring to Figures 8 and 9, with a low weight on the penalty, the cost function has two local minimum. One is at zero and the other, which is the global minimum, is very near the actual or expected solution. The local minimum at zero is caused by the very high gradient value of the non-quadratic penalty near zero. The fact is, with low or moderate weight on the non-quadratic penalty, a very low value of the gradient of the non-quadratic penalty when added to the gradient of the error will drive the global minimum of cost function (11) very near the actual value. This feature makes the proposed approach different from other traditional approaches. As shown in Figures 8 and 9, with a high weight on the penalty λ_n , the cost function has only one local minimum, which is therefore the global minimum.

To get further insight about the role of different penalties, we study the effects of quadratic and non-quadratic penalties with the aid of Figure 10. In equations (10) and (11), the cost functions being minimized have the following general form:

$$J = E + P \tag{12}$$

where E is the error function and P is the penalty on y.

Referring to Figure 10, we deduce that when the value of y is high, the penalty on y is higher in the quadratic case than in the non-quadratic case. The higher penalty on



Fig. 9. Optimal solutions for a high and a low weight on the non-quadratic penalty

Fig. 10. Comparison of penalty function values when the variable is of high value

y may result in a less accurate solution because the cost function being minimized is J not E and the optimizer "will choose" the solution that minimizes J, which may not necessarily minimize E. In order to accurately recover signals that exhibit large fluctuations in strength, it is desired that penalties on the variable y be restricted in the working range. This requirement further encourages the use of nonquadratic penalty instead of quadratic penalty. Simulation examples presented in Section 5 support this claim.

IV. ILLUSTRATIVE EXAMPLES

A. Numerical Example

We present a numerical example to explain the proposed methodology. Consider the following sensor model:

$$z = \mathscr{F}^{-1}\{\mathscr{H}\mathscr{F}\{g(y)\}\}$$
(13)

The discrete version of the above equation can be written in the following matrix equation form [4]:

$$z = FGy \tag{14}$$

where G corresponds to nonlinear operation g and F denotes the low pass filtering operation \mathcal{H} .

Taking only the first two time samples, we will use the following ill-conditioned process to explain the algorithm:

$$z = \begin{pmatrix} 2 & 1\\ 2 & 1.1 \end{pmatrix} y + n \tag{15}$$

which gives

$$z_1 = 2y_1 + y_2 + 0.1$$

$$z_2 = 2y_2 + 1.1y_2 + 0.2$$
 (16)

Suppose that the actual samples of the signals are given by

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 10 \\ 8 \end{bmatrix}$$
(17)

The samples of the sensor output for this case are given by,

$$z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 28.1 \\ 29.0 \end{bmatrix}$$
(18)

Let us now present three examples to analyze the effect of quadratic and non-quadratic penalty functions. The following initial values are used to begin the solution search in all three optimization problems.

$$y_o = \begin{bmatrix} -100\\ -1000 \end{bmatrix}$$
(19)

Case 1: Minimization of Quadratic Error Function

$$\min_{y_1, y_2} J = \left\| \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} - \begin{pmatrix} 2 & 1 \\ 2 & 1.1 \end{pmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\|^2$$
$$= (28.1 - 2y_1 - y_2)^2 + (29.0 - 2y_1 - 1.1y_2)^2 20)$$

The optimal solution is,

$$y^* = \left[\begin{array}{c} 9.55\\9.00\end{array}\right] \tag{21}$$

Case 2: Minimization of Quadratic Error Function with Quadratic Penalty

$$\min_{y_1, y_2} J_q = \left\| \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} - \begin{pmatrix} 2 & 1 \\ 2 & 1.1 \end{pmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\|^2 + 0.1 \left(y_1^2 + y_2^2 \right)$$
$$= (28.1 - 2y_1 - y_2)^2 + (29.0 - 2y_1 - 1.1y_2)^2 + 0.1 \left(y_1^2 + y_2^2 \right)$$

The optimal solution is,

$$y^* = \begin{bmatrix} 11.0196\\ 5.9384 \end{bmatrix}$$
(22)

Case 3: Minimization of Quadratic Error Function with Non-quadratic Penalty

$$\min_{y_1,y_2} J_s = \left\| \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} - \begin{pmatrix} 2 & 1 \\ 2 & 1.1 \end{pmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\|^2 + 0.1 \left(|y_1|^{0.5} + |y_2|^{0.5} \right)$$

The optimal solution is,

$$y^* = \begin{bmatrix} 10.0379\\ 8.0687 \end{bmatrix}$$
(23)

Clearly, with non-quadratic cost function, more accurate results are attained.

B. Simulation Example 1

To illustrate performance of the proposed scheme, we present a simulation example in this section. Consider the measurement of a signal shown in Figure 11 through a non-linear sensor whose input-output characteristic is depicted in Figure 13. The distorted sensor output is shown in Figure 12.

The different performance indices formulated in the previous section are optimized to obtain an estimate of the original signal.





Fig. 12. Sensor Output Before Low Pass Filtering, *w*





Fig. 14. Signal Recovered by Optimizing Quadratic Error Function, J_{e_1}

$$J_{q_1} = \|z - \mathscr{F}^{-1} \{ \mathscr{HF} \{ g(y) \} \} \| + \lambda_q \sum_{i=1}^n y_i^2 \qquad (25)$$

$$J_{n_1} = \|z - \mathscr{F}^{-1}\{\mathscr{HF}\{g(y)\}\}\| + \lambda_n \sum_{i=1}^N |y_i|^{0.5}$$
(26)

where N is the number of samples of signal, y.

Figures 14, 15 and 16 show the signals recovered by optimizing the error function J_{e_1} quadratic cost function J_{q_1} and non-quadratic cost function J_{n_1} respectively. A simple comparison with the actual data clearly demonstrates that the signal obtained with non-quadratic criteria is the closest to the actual data.



Fig. 15. Signal Recovered by Optimizing Quadratic Cost Function, J_{q_1}

Fig. 16. Signal Recovered by Optimizing Non-quadratic Cost Function, J_{n_1}

C. Simulation Example 2

We present another example to demonstrate the performance of the proposed scheme. The signal to be measured is shown in Figure 17. We use the same nonlinear sensor as in the previous example. The distorted sensor output is shown in Figure 18.



The following performance indices are derived to estimate the actual signal.

$$J_{q_2} = \|z - \mathscr{F}^{-1}\{\mathscr{HF}\{g(y)\}\}\| + \lambda_q \sum_{i=1}^N y_i^2 \qquad (27)$$

$$J_{n_2} = ||z - \mathscr{F}^{-1} \{ \mathscr{HF} \{ g(y) \} \} || + \lambda_n \sum_{i=1}^N |y_i|^{0.1}$$
(28)

where *N* is the number of samples of signal, *y* and λ_q and λ_n are weights on quadratic and non-quadratic penalties respectively.

The signal recovered by minimizing the quadratic criteria J_{q_2} is shown in Figure 19. Figure 20 shows the signal recovered by optimizing the non-quadratic criteria J_{n_2} . By filtering out the low strength or insignificant frequency components, we obtain signals shown in Figures 21 and 22, which are the solutions to the quadratic and non-quadratic criteria respectively. Clearly, the signal obtained using the non-quadratic criterion is more accurate than that of quadratic performance index.



Fig. 19. Signal Recovered by Optimizing Quadratic Cost Function, J_{q_2}

Fig. 20. Signal Recovered by Optimizing Non-quadratic Cost Function, J_{n_2}

V. CONCLUSIONS

A new signal recovery scheme that eliminates or reduces the distortion caused by ill-posed sensor nonlinearity is proposed. We have discussed some of the advantages in using a non-quadratic performance index in place of the traditional



Quadratic Criteria (J_{q_2}) (after filtering out insignificant frequency components)

Fig. 22. Signal Recovered by Non-quadratic Criteria, (J_{n_2}) (after filtering out insignificant frequency components)

quadratic criterion. When the nonlinear sensory operation characterized by the sensor nonlinearity is ill-posed, the traditional signal recovery schemes fail to produce accurate results. We have derived a new technique that can improve the accuracy of the recovered signal by utilizing the characteristics of non-quadratic criteria. The numerical and simulation experiments suggest that the proposed method is a valuable tool to solve the ill-conditioned problem posed. Further research is needed to guarantee its usage and to identify its limitations.

The proposed scheme appears to be extremely useful in compensating the effects of singular nonlinear sensory operations to a certain extent, which is currently done by expensive ad-hoc signal conditioning techniques. Though the main benefit of the technique is the increased accuracy in sensor readings, the implementation of this recovery scheme would eliminate the need for the expensive signal processors, which would be a tremendous economic benefit.

VI. ACKNOWLEDGMENTS

This work is supported by the National Science Foundation under Grant No. CMS-0097719.

REFERENCES

- [1] M. Kothari, J. Webster, W. Tompkins, J. Wertsch, and P. B. y Rita, "Capacitive sensors for measuring the pressure between the foot and shoe," in Proc. Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 805–806, 1988.
- [2] C. Nwagboso(Editor), Automotive Sensory System. London: Chapman & Hall, 1993.
- [3] S. Jayasuriya and R. Langari, "Nonlinear filtering of signals exhibiting large fluctuations in strength," in Proc. 1996 ASME International Mechanical Engineering Congress and Exposition, pp. 159–163, 1996.
- [4] S. Suranthiran and S. Jayasuriya, "A filtering methodology to recover signals distorted by sensor nonlinearity," in Proc. Japan-USA Symposium on Flexible Automation, pp. 1027–1034, 2002.
- [5] —, "Signal recovery and noise removal with memory-less nonlinear sensors," *in Proc. of American Control Conference*, pp. 4155–4160, 2003.
- [6] S. Suranthiran, "Control with non-quadratic criteria," Master's thesis, University of Cambridge, Cambridge, England, December 1997.
- [7] A. N. Tikhonov and V. Y. Arsenin, Solutions of Ill-posed Problems. 1511 K Street, N.W., Washington D.C. 20005: V. H. Winston & Sons, 1977.
- [8] P. C. Hansen, "Numerical tool for analysis and solution of Fredholm integral equations of the first kind," *Inverse Problems*, vol. 8, pp. 849– 872, 1992.