Solution to least square problems with stochastic uncertainties by polynomial chaos

Pham Luu Trung Duong and Moonyong Lee*

School of chemical engineering, Yeungnam Univeristy, Gyeongsan, South Korea

(Tel: 053-810-3241; Email: mynlee@yu.ac.kr)

Keywords—Stochastic least square, polynomial chaos, cubature

I. Introduction

In the standard least square problem, the objective is to minimize the squared Euclidean norm $\|Ax-y\|^2$ of the residual of a system of a linear equation. However, normally the data matrices $A = A(\boldsymbol{\xi}) \in R^{m,n}; y = y(\boldsymbol{\xi}) \in R^m$ are not known exactly, and are nonlinear function of vector uncertain real parameters $\boldsymbol{\xi} \in \Theta$. Hence, robust worst-case or probabilistic approach is used to take into account of these uncertainties when solving least square problem with uncertainties [1, 2]. The worst- case approach is discussed in the works [1, 3, 4]. The probabilistic framework is discussed in [1, 2]. In the worst-case approach solution of min-max problem is sought: let

$$f(x,\boldsymbol{\xi}) = \|A(\boldsymbol{\xi}) - y(\boldsymbol{\xi})\|^2 \tag{1}$$

then the worst-case solution is

$$x^*_{wc} = \arg\min_x \max_{\boldsymbol{\xi}\in\Theta} f(x,\boldsymbol{\xi}) \tag{2}$$

From other view point, it is natural to look for a solution minimizing expected value of the LS residual:

$$x_E^* = \arg\min_x E_\xi[f(x,\boldsymbol{\xi})] \tag{3}$$

The problem (3) is referred as stochastic robust approximation problem in [1]. In [1], probabilistic framework consider examples with the *A, y* are linear function of uncertainties since it is very difficult to evaluate the objective. In [2], the Monte Carlo (MC) is used for estimation and optimizing the expected value of the residual. Unfortunately, the computational cost of the MC can be very high.

The generalized polynomial chaos (gPC) [5] approach is an effective method for estimating the moment of complex quantity. In this paper, we adopt the gPC approach for evaluation the objective and solving stochastic robust least square problems. Moreover, the method reserves the structure of LS problem, which lead to an exact minimizer.

II. Problem Formulation

Assume that the $\boldsymbol{\xi} = (\xi_1,...,\xi_n)$ consists of independent random parameters with probability density functions (pdf) $\rho_i(\xi_i):\Gamma_i \to R^+$ .The joint pdf is $\boldsymbol{\rho} = \prod_{i=1}^n \rho_i$ with the support $\boldsymbol{\Gamma} = \prod_{i=1}^n \Gamma_i \in R^{+n}$. The stochastic LS optimization problem can be solved using algorithm below:

- Construct a suitable *n* dimension cubature with single graded lexicographic index j [5]: $\{\boldsymbol{\xi}^{(j)}, \mathbf{w}^{(j)}\}_{j=1}^Q$ (4)
- The expected value of LS residual is approximated as: $\sum_{\mathbf{j}=1}^{Q} f(x,\boldsymbol{\xi}^{(\mathbf{j})})\mathbf{w}^{(\mathbf{j})}$ (5)

  Note that the accuracy of the approximation in (5) is given by the degree of exactness of the cubature.

- The objective has sum of square structure and can be cast into new form [6]: $\left\| \Lambda x - \Upsilon \right\|^2$ (6)
  with

$$\Lambda = \begin{vmatrix} \sqrt{\mathbf{w}^{(1)}} A(\xi^{(1)}) \\ \vdots \\ \sqrt{\mathbf{w}^{(Q)}} A(\xi^{(Q)}) \end{vmatrix}; \quad \Upsilon = \begin{vmatrix} \sqrt{\mathbf{w}^{(1)}} y(\xi^{(1)}) \\ \vdots \\ \sqrt{\mathbf{w}^{(Q)}} y(\xi^{(Q)}) \end{vmatrix} \tag{7}$$

Thus, an minimizer to stochastic least square can be approximated as $x^{*}_{E} = \Lambda^{*}\Upsilon$, where Λ^{*} is pseudo inverse of matrix Λ .

III. SPARSE GRID

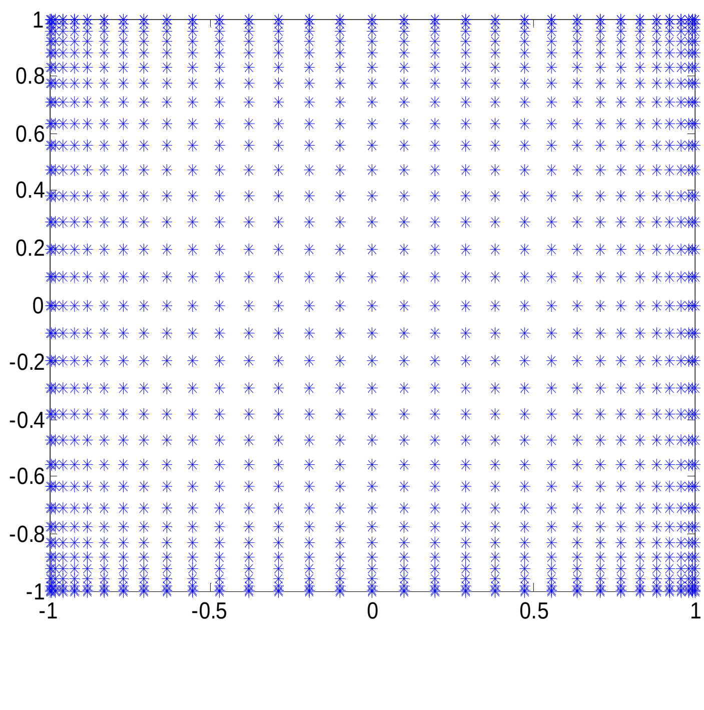
The tensor quadrature rule can become intractable due to the exponential growth of the number of functional evaluation as dimension of random space increases. In [5],[10] quadrature rule based on sparse grid is recommended for problem with high dimension of random space.
Starting with the one dimensional rule, the Smolyak quadrature is given by:

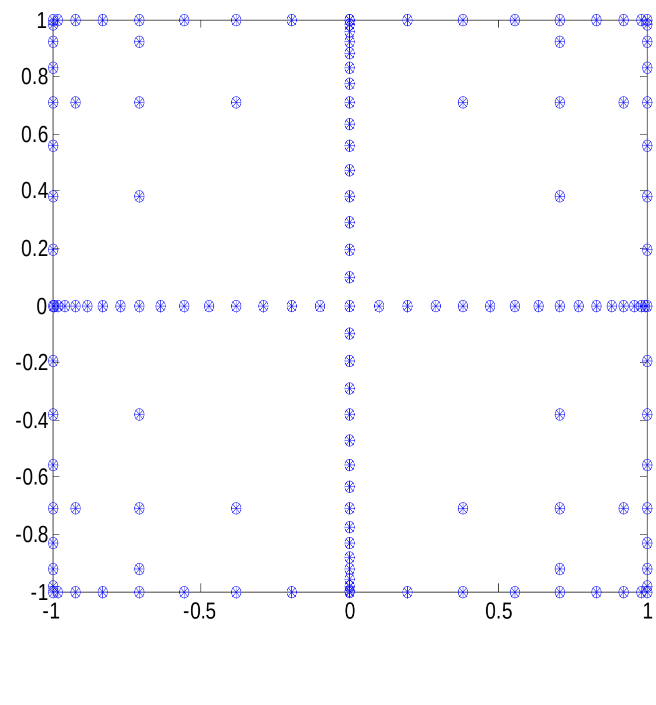
$$U^{Q}[f] = (Q_{i_1}{}^{(1)} \otimes \ldots \otimes Q_{i_d}{}^{(1)}) f = \Big| \; \Big| \sum_{J-N+1 \le \mathbf{i} \le l} (-1)^{J-|\mathbf{i}|} \binom{N-1}{J-|\mathbf{i}|} (Q_{i_1}{}^{(1)} \otimes \ldots \otimes Q_{i_d}{}^{(1)})$$

where (8)

$$|\mathbf{i}| = i_1 + i_2 + \ldots + i_n$$

The number of node in Smolyak is depend on dimension d and level index l.Fig.1 provides an example for comparison tensor quadrature and Smolyak quadrature for $d=3, l=5$





(a) *(b)*

Figure 2:Tensor quadrature v.s sparse grid quadrature: (a):Tensor quadrature, (b):Smolyak sprase grid with Clenshaw-Curtis(CC) as 1D integration rule

Table 1 gives number of nodes using Tensor quadrature, Smolyak quadrature with Gaussian and Clenshaw-Curtis as 1D integration rule.

Table 1 Number of node for different quadrature rules.

	$d=2$	$d=3$	$d=4$	$d=5$	$d=6$	$d=7$
Tensor Gauss	25	125	625	3125	15625	78125
Smolyak Gauss	17	31	49	71	97	127
Smolyak CC	13	25	41	61	85	113

IV. EXAMPLES

In this section, several examples are considered for demonstration the usefulness of the proposed method.

a. Example 1

Consider the LS problem taken from [2]:

$$A = A_0 + \sum_{i=1}^{3} \xi_i A_i, \; y^T = [0 \;\, 2 \;\, 1 \;\, 3]$$
$$A_0 = \begin{vmatrix} 3 & 1 & 4 \\ 0 & 1 & 1 \\ -2 & 5 & 3 \\ 1 & 4 & 5.2 \end{vmatrix}, A_1 = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix}, A_2 = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}, A_3 = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} \qquad (9)$$

with ξ_i are Gaussian random uncertainties with zeros mean and standard deviations $\sigma_1 = 0.067; \sigma_1 = 0.1; \sigma_1 = 0.2$. In this case, the exact solution from theorem 1 in [2] and results to be: $x^* = [-2.352; -2.076; 2.481]$. The result using gPC method is: $x_{gPC}{}^* = [-2.350; -2.0747; 2.480]$. From the results, it can be seen that the gPC method can give quite accurate results. Figure 2 shows the histogram of the residual error with normal LS solution and stochastic LS solution by gPC. From the Figure, it can be seen that, the normal LS is more sensitive to uncertainties.

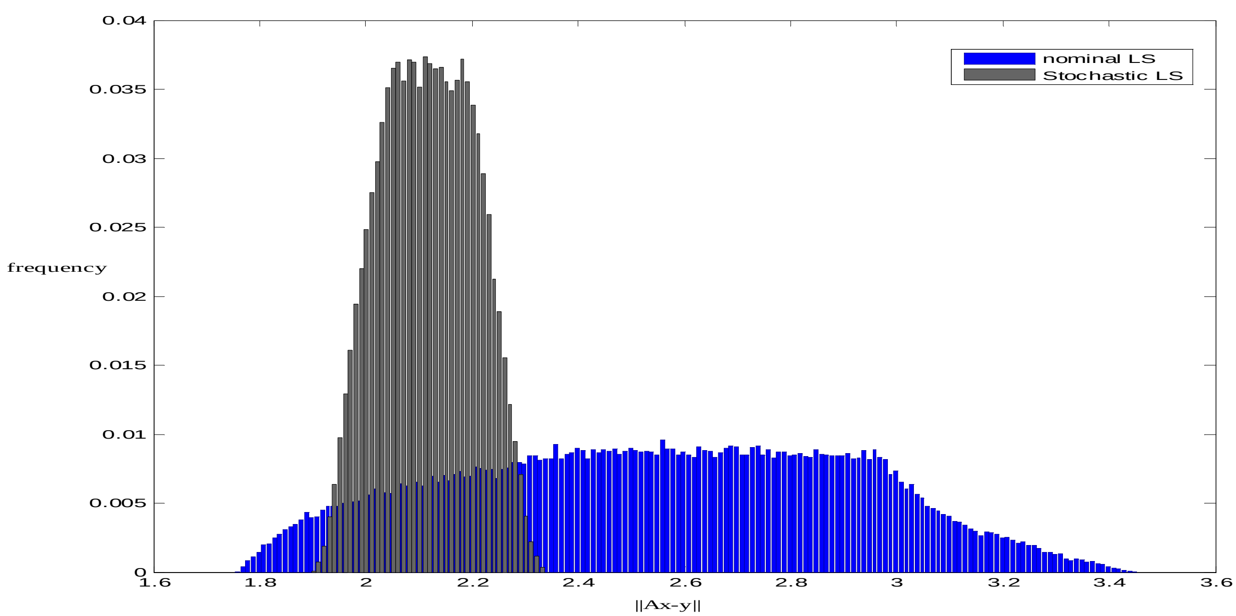


Figure 2 Distribution of the residual error for the two solutions of least square problem in example 1

b. Example 2

Consider LS problem taken from [1]with dimension *m=50, n=20*. The data matrices have given norm as: $\|\overline{A}\|_2 = 10; \|A_1\|_2 = \|A_2\|_2 = 1$. The uncertainties u_1, u_2 lie in the unit disk in R^2 .The generalize spherical coordinate is used to transform the uniform box uncertainties to a uniform disk uncertainties [7] .The solution of stochastic least square was obtained by the algorithm described in the previous part. The sensitivity of the solution to the stochastic least square is illustrated in Figure 3. The sensitivity of other methods such as: Tikhonov regularized solution, robust worst-case approach [1] are also shown in this figure. It can be seen that the distribution of the residual for nominal LS is wide spread, hence is very sensitive to parameter variation. The robust worst-case solution gives the least sensitive solution with some trade off on the residual value. The Tikhonnov and stochastic LS provide more flexible tradeoff between the robustness and residual errors. Both Tikhonov and stochastic LS reserve the LS structure of the problem, while the robust worst- case approach transform the problem into a semidefinite programming problem. More detail on the robust worst-case approach can be found in [1] and references therein. The cubature can be obtained using existing software in [8, 9].

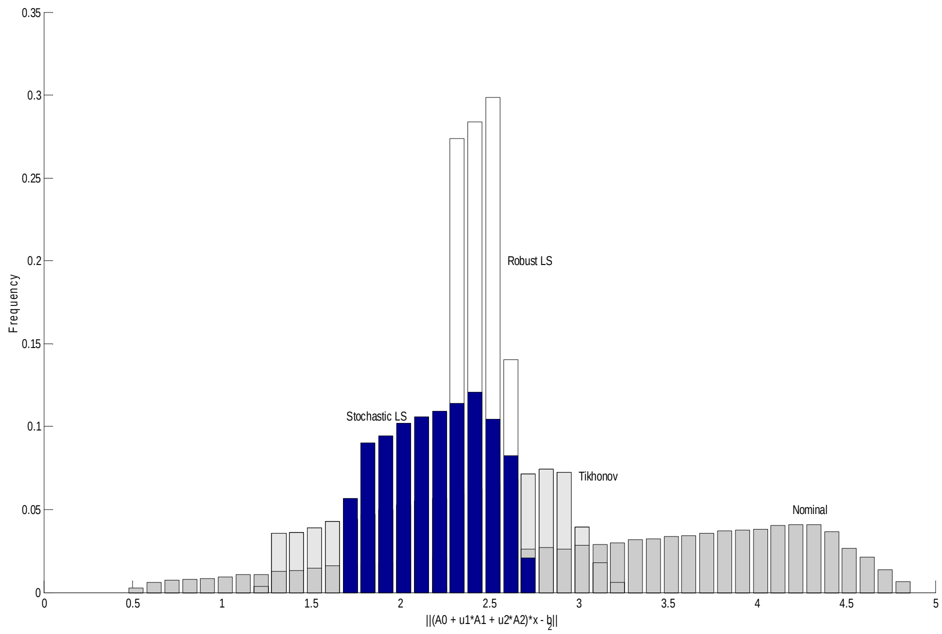


Figure 3 Distribution of the residual error for the four examples of least square problem in example 2

V. CONCLUSIONS

This paper presented an alternative approach for stochastic robust LS problem based on minimization the mean of the residual with respect to the uncertainties .By using polynomial chaos method, a cubature is obtained, then the stochastic robust LS is transform into an equivalent deterministic LS with larger dimension. The polynomial chaos method reserves the LS structure of the problem, hence leads to an exact minimizer. The proposed method is compared with existing techniques to demonstrate its usefulness in several numerical examples. Due to the property of cubature with full tensorization used in this paper; the number of uncertainties can be handle in the proposed method is small. For uncertainties with larger dimension, the sparse grid cubature [10] can be used and this will be incorporated in future works.

REFERENCES

[1] S. Boyd , L. Vandeberghe, *Convex optimization*, Camridge University Press, 2004.

[2] G. Calafiore, F. Dabbene, Near optimal solution to least square problems with stochastic uncertainties, *System & Control Letters*, 54,2005,1219-1232.

[3] L. El Ghaoui, H. Lebret, Robust solutions to least-squares problems with uncertain data, *SIAM J. Matrix Anal. Appl.* , 18 (4), 1997, 1035–1064.

[4] S. Chandrasekaran, G.H. Golub, M. Gu, A.H. Sayed, Parameter estimation in the presence of bounded data uncertainties, *SIAM J. Matrix Anal*. Appl., 19, 1998, 235–252.

[5] D. Xiu, *Numerical method for stochastic computations: A spectral method approach*, Princeton university press, 2010.

[6] F .Dabbene, Exact solution of uncertain convex optimization problems, *in: Proceeding of American Control Conference*, 2007, 2654-2659.

[7] R.Tempo, G. Calfiore, F. Dabbene, *Randomize algorithms for analysis and control of uncertain systems*, Springer, 2003.

[8] K.A. Puvkov, N.D. Egupov, A.M. Makarenkov, A.I. Trofimov . *Theory and Numerical Methods for Studying Stochastic Systems*, Fizmatlits: Moscow, **2003**.

[9] W. Gaustchi, *Orthogonal Polynomials: Computation and approximations*, Oxford University Press,2003. OPQ suite http://www.cs.purdue.edu/archives/2002/wxg/codes/OPQ.html

[10] M. Holtz, *Sparse grid quadrature in high dimensions with application in finance and insurance*, Springer, 2010.