Random Spherical Uncertainty in Estimation and Robustness

Boris T. Polyak and Pavel S. Shcherbakov
Institute for Control Science, Moscow, Russia, 117806, Profsojuznaya, 65
phone: 7-(095)-334-7641; fax: 7-(095)-420-2016
e-mail: boris@ipu.rssi.ru e-mail: sherba@ipu.rssi.ru

Abstract

A theorem is formulated that gives an exact probability distribution for a linear function of a random vector uniformly distributed over a ball in n-dimensional space. This mathematical result is illustrated via applications to a number of important problems of estimation and robustness under spherical uncertainty. These include parameter estimation, characterization of attainability sets of dynamical systems, and robust stability of affine polynomial families.

1. Introduction

Different fields of control theory exploit different models for the uncertainty. In parameter and state estimation, random perturbations is a tradition, and least squares and Kalman filtering are the standard tools for estimation. In parametric robustness, deterministic uncertainty models are of most concern ([1]). Often, hard bounds on the uncertain parameters are not known; instead, their probabilistic characteristics may be available, and the solution often involves Monte Carlo simulations ([2]). Main benefits are the low computational complexity and a considerable enhancement of admissible uncertainty domains in exchange of a small risk of violation of deterministic specifications. The results obtained so far relate to independent random variables.

We follow the probabilistic approach and work with dependent random parametric uncertainty, — the uniform distribution on a ball in l_2-norm.

2. A linear function in a uniform distribution on a ball

Notation: $\mathbf{B} \doteq \{x \in \mathbb{R}^n : \|x\| \le 1\}$; $\partial\mathbf{B} \doteq \{x \in \mathbb{R}^n : \|x\| = 1\}$; $\mathcal{U}(\mathbf{Q})$ is the uniform distribution on the set $\mathbf{Q} \subset \mathbb{R}^n$; $\mathcal{N}(0, I_m)$ is the standard m-dimensional Gaussian distribution; $\mathcal{C}(m)$ is the χ^2 distribution with m degrees of freedom; "$\sim$" means "distributed as."

Lemma 1. *Let $\xi \sim \mathcal{N}(0, I_n)$. Then $\xi/\|\xi\| \sim \mathcal{U}(\partial\mathbf{B})$. If, in addition, a random variable $\rho \sim \mathcal{U}([0,1])$ is independent of ξ, then $\rho^{1/n}\xi/\|\xi\| \sim \mathcal{U}(\mathbf{B})$.*

The first proposition of the lemma traces back to [3].

Theorem 1. *Let $q \sim \mathcal{U}(\mathbf{B})$, $\mathbf{B} \subset \mathbb{R}^n$; let a matrix $A \in \mathbb{R}^{m\times n}$ has rank $m \le n$. Then the random variable*

$$\tau \doteq \Big((AA^{\mathrm{T}})^{-1}Aq,\, Aq\Big)$$

has beta *distribution $\tau \sim \mathcal{B}(\frac{m}{2}, \frac{n-m}{2}+1)$ with density*

$$f_\tau(x) = \begin{cases} \dfrac{\Gamma(\frac{n}{2}+1)}{\Gamma(\frac{m}{2})\Gamma(\frac{n-m}{2}+1)} x^{\frac{m}{2}-1}(1-x)^{\frac{n-m}{2}}, & x \in (0,1) \\ 0, & x \notin (0,1) \end{cases}$$

Theorem 1 provides an explicit description of a confidence ellipsoid for a random vector which is a linear transformation of a uniform distribution on a ball.

Theorem 2 (Asymptotic form). *Assume that for every $n \ge m$, a matrix $A_n \in \mathbb{R}^{m\times n}$ has rank m, and $q^{(n)} \sim \mathcal{U}(\mathbf{B})$, $\mathbf{B} \subset \mathbb{R}^n$. Then the random vector $\eta^{(n)} \doteq n^{1/2}(A_nA_n^{\mathrm{T}})^{-1/2}A_nq^{(n)}$ tends in distribution to $\mathcal{N}(0, I_m)$ as $n \to \infty$.*

Since $\|\xi\|^2 \sim \mathcal{C}(m)$ for $\xi \sim \mathcal{N}(0, I_m)$, we observe that $n\Big((A_nA_n^{\mathrm{T}})^{-1}A_nq_n, A_nq_n\Big) \to \mathcal{C}(m)$ in distribution. Hence, with increase of $\dim q$, the transformed vector Aq tends to concentrate closer the center of the image of the support set. Below, we exploit this effect when constructing *probabilistic predictors* of sets in $\mathbb{R}^n$.

3. Parameter estimation

We consider the linear regression model

$$y_i = a_i^{\mathrm{T}}c^* + \xi_i, \qquad i = 1, \dots, n, \tag{1}$$

where $c^* \in \mathbb{R}^m$, $m \le n$, is the vector of unknown parameters, $a_i \in \mathbb{R}^m$, $i = \overline{1,n}$, are fixed known regressors, y_i, $i = \overline{1,n}$, are observations, and $\xi_i \in \mathbb{R}$, $i = \overline{1,n}$, is noise such that $\xi \doteq (\xi_1, \dots, \xi_n) \sim \mathcal{U}(r\mathbf{B}) \subset \mathbb{R}^n$. Assuming that there are m linearly independent vectors among the a_i, the least squares estimate for c^* is $\widehat{c} = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}y$, where $y \doteq (y_1, \dots, y_n)^{\mathrm{T}}$ and $A = [a_1^{\mathrm{T}}\, a_2^{\mathrm{T}}\, \dots\, a_n^{\mathrm{T}}] \in \mathbb{R}^{n\times m}$. The theorem below provides a closed-form description of confidence ellipsoids for c^*.

Theorem 3. *Let $0 \le p \le 1$ and let τ_p denote the $100p\%$ quantile of the beta distribution $\mathcal{B}(\frac{m}{2}, \frac{n-m}{2}+1)$. Then, under the conditions above, the ellipsoid*

$$\mathbf{E}_p \doteq \Big\{x \in \mathbb{R}^k : \Big(A^{\mathrm{T}}A(x-\widehat{c}),\, x-\widehat{c}\Big) \le r^2\tau_p\Big\} \tag{2}$$

is a $100p\%$ confidence domain for the vector c^* in (1).

4. Attainability sets of dynamic systems

We consider a discrete-time dynamic system

$$x_{k+1} = Ax_k + Bw_{k+1},\ k = 0, 1, \dots, \quad w_k \in \mathbb{R}^m,$$

with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, where the uncertainty $\{w_1, \ldots, w_N\}$ accumulated by the N-th step satisfies $\sum_{k=1}^N \|w_k\|^2 \le c^2$ for some $c > 0$

For every $N > 0$, the *attainability set* of this system is the set of all its possible states which can be attained by the Nth step [4]. For a controllable pair (A, B), this set is the ellipsoid $\mathbf{Q}_N = \mathbf{E}_N \doteq \{x \in \mathbb{R}^n : (S_N^{-1}(x - \overline{x}_N), x - \overline{x}_N) \le c^2\}$ with center $\overline{x}_N \doteq A^N x_0$ and matrix $S_N = M_N M_N^{\rm T}$, where $M_N \doteq [A^{N-1}B \;\; A^{N-2}B \; \ldots \; AB \;\; B] \in \mathbb{R}^{n\times mN}$ is the controllability matrix of the system.

Let the uncertainty be random; namely, the mN-dimensional vector $w \doteq (w_1; w_2; \ldots; w_N)$ is $w \sim \mathcal{U}(\mathbf{B})$, $\mathbf{B} \subset \mathbb{R}^{mN}$. We represent $x_N = \overline{x}_N + cM_N w$ and consider the random variable $\tau_N \doteq \Big((c^2 M_N M_N^{\rm T})^{-1}(x_N - \overline{x}_N), x_N - \overline{x}_N\Big)$. Theorem 1 yields $\tau_N \sim \mathcal{B}(\frac{n}{2}, \frac{mN-n}{2} + 1)$. Next, we specify the confidence probability p and denote the $100p\%$ quantile for beta distribution by $\tau_{N,p}$.

Theorem 4. *The ellipsoid*

$$\mathbf{E}_{N,p} = \Big\{x \in \mathbb{R}^n : \; \Big(S_N^{-1}(x - \overline{x}_N), x - \overline{x}_N\Big) \le c^2 \tau_{N,p}\Big\}$$

is a $100p\%$ predictor of the attainability set $\mathbf{E}_N$ in the sense that $\mathrm{Prob}\{x_N \in \mathbf{E}_{N,p}\} = p$.

With increase of N, the quantile $\tau_{N,p}$ decreases, and Theorem 4 gives a considerable *volumetric reduction*, $\mathbf{Vol}(\mathbf{E}_N)/\mathbf{Vol}(\mathbf{E}_{N,p}) = 1/\tau_{N,p}^n$; i.e., the probabilistic approach allows for the high-confidence replacement of the state uncertainty $\mathbf{E}_N$ with the much smaller predictor $\mathbf{E}_{N,p}$.

5. Robust stability of polynomials

We consider the following polynomial family:

$$p(s, q) = p_0(s) + \sum_{i=1}^n q_i p_i(s), \qquad (3)$$

where $p_0(s)$ is the nominal and $p_i(s)$, $i = \overline{1, n}$, are known perturbation polynomials with real coefficients, $q = (q_1, \ldots, q_n)^{\rm T}$ is the vector of uncertain parameters subjected to spherical constraints

$$q \in \gamma\mathbf{B} \doteq \Big\{q \in \mathbb{R}^n : \sum_{i=1}^n q_i^2 \le \gamma^2\Big\}, \qquad (4)$$

and $\gamma > 0$ is the range of uncertainties. We consider the continuous-time case and Hurwitz stability. The nominal polynomial $p_0(s)$ is assumed to be stable. The goal is to check if (3)–(4) is robustly stable and to determine the robust stability radius $\gamma_{\max} \doteq \max\{\gamma : p(s, q)$ is stable for all $q \in \gamma\mathbf{B}\}$.

The traditional approach uses the value set concept and the zero exclusion principle, [1]. The value set of (3)–(4) is the ellipse $\mathbf{E}(\omega) = \{x \in \mathbb{R}^2 : \Big(C(x - p_0(j\omega)), x - p_0(j\omega)\Big) \le \gamma^2\}$, where $C \doteq (AA^{\rm T})^{-1}$ and

$$A = A(\omega) = \begin{bmatrix} \mathrm{Re}\, p_1(j\omega) & \ldots & \mathrm{Re}\, p_n(j\omega) \\ \mathrm{Im}\, p_1(j\omega) & \ldots & \mathrm{Im}\, p_n(j\omega) \end{bmatrix}.$$

The deterministic robust stability radius is found to be

$$\gamma_{\max} = \inf_{\omega \ge 0} \Big(Cp_0(j\omega), \, p_0(j\omega)\Big)^{1/2}, \qquad (5)$$

which is obtained by performing the frequency sweep with $\mathbf{E}(\omega)$, see [5] for the results of such kind.

Assume that the specified uncertainty radius γ is greater than the robustness margin $\gamma_{\max}$ obtained via a deterministic test of the sort (5) and let $q \sim \mathcal{U}(\gamma\mathbf{B})$. Given *probability risk* $\varepsilon \in [0, 1]$, for any fixed $\omega \ge 0$ we introduce a $100\%(1 - \varepsilon)$ *probabilistic predictor* of the value set $\mathbf{E}(\omega)$ as a domain $\mathbf{E}_{1-\varepsilon}(\omega) \subset \mathbf{E}(\omega)$ such that $\mathrm{Prob}\{p(j\omega, q) \in \mathbf{E}_{1-\varepsilon}(\omega)\} = 1 - \varepsilon$. Theorem 1 yeilds $\mathbf{E}_{1-\varepsilon}(\omega) \doteq \Big\{x \in \mathbb{R}^2 : \Big(C(x - p_0(j\omega)), x - p_0(j\omega)\Big) \le \gamma_\varepsilon^2\Big\}$, where $\gamma_\varepsilon = \gamma\sqrt{1 - \varepsilon^{2/n}}$, i.e., $\mathbf{E}(\omega)$ and its predictor are similar ellipses and $\gamma_\varepsilon/\gamma = (1 - \varepsilon^{2/n})^{1/2} < 1$ is the similarity coefficient. Hence, if $q \sim \mathcal{U}(\alpha(\varepsilon)\gamma\mathbf{B})$, where $\alpha(\varepsilon) \doteq \gamma/\gamma_\varepsilon > 1$, then $\mathrm{Prob}\{p(j\omega, q) \in \mathbf{E}(\omega)\} = 1 - \varepsilon$. We then perform the frequency sweep with the predictor $\mathbf{E}_{1-\varepsilon}(\omega)$. To overcome the so-called *cross-frequency effect*, we formulate

Lemma 2. *Let a symmetric matrix $B(\omega) \in \mathbb{R}^{n\times n}$ have rank 2 for all $\omega \in \Omega \doteq [\omega_1, \omega_2]$ and $\omega_0 \in \Omega$. Let $\|B(\omega) - B(\omega_0)\| \le \delta$ for all $\omega \in \Omega$ and some small $\delta > 0$. Then for $q \sim U(\mathbf{B})$ and any $\varepsilon \in [0, 1]$, we have*
$\mathrm{Prob}\Big\{\Big(B(\omega)q, q\Big) \le 1 - \varepsilon^{2/n} \; \forall \omega \in \Omega\Big\} \ge 1 - (\varepsilon^{2/n} + \delta)^{n/2}$.

Using Lemma 2, the relative volume of the violating portion of the ball $\gamma\mathbf{B}$ in the q-space can be estimated quite accurately. Numerical experiments confirm the validity of the approach.

Acknowledgements

Funding for this research was provided by RFBR, Grant 99-01-00340; NSF, Grant ECS-9418709 and INTAS, Grant 97-10782.

References

[1] B. R. Barmish, *New Tools for Robustness of Linear Systems,* New York: Macmillan, 1994.

[2] X. Zhu, Y. Huang, and J. Doyle, “Soft vs. hard bounds in probabilistic robustness analysis,” in *Proc. 35-th CDC,* Kobe, Japan, 1996.

[3] M. E. Muller, “A note on a method for generating random points uniformly on n-dimensional spheres,” *Comm. ACM,* vol. 2, pp. 19–20, 1959.

[4] F. C. Schweppe, *Uncertain Dynamic Systems,* Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

[5] C.Soh, C.Berger, K.Dabke, “On the stability properties of polynomials with perturbed coefficients,” *IEEE TAC,* No. 10, 1033–1036, 1985.

[6] S. Wilks, *Mathematical Statistics,* New York: John Wiley, 1963.