Properties and Calculation of Singular Normal Distributions

René Henrion[1], Tamas Szántai[2]

[1] Weierstrass Institute Berlin
10117, Berlin, Mohrenstr. 39, Germany
henrion@wias-berlin.de
[2] Technical University Budapest, Operations Research Group
1111, Budapest, Muegyetem rakpart 3. H. ép. 47., Hungary
szantai@math.bme.hu

Abstract. The need for calculating and characterizing singular normal distributions arises in a natural way when considering probabilistic constraints of the type , where A is a rectangular matrix having more rows than columns, b is some function and ξ has a nondegenerate multivariate normal distribution. The talk provides structural results (criterion for Lipschitz continuity and differentiability) as well as a formula for the calculation of singular normal distribution functions.

Keywords. singular normal distributions, probabilistic constraints, probability of polyhedra

1 Introduction

An m-dimensional random vector η is said to have a singular normal distribution if there exists some s-dimensional random vector ξ having a nondegenerate normal distribution such that

$$\eta = A\xi + b, \tag{1}$$

where A is an (m, s)-matrix with rank smaller than m and b is an m-vector. Singular normal distributions are those normal distributions whose covariance matrix has a rank strictly smaller than the dimension of the random vector. Such seemingly artificial distributions arise in a natural way in problems of stochastic optimization, where a relatively small (nondegenerate-) normally distributed random vector induces a large number of linear inequality constraints. As an example, we refer to the problem of optimal capacity expansion in a network with stochastic demands (see [4], p. 453), which leads to a probabilistic constraint of the form

$$P(A\xi \leq b(x)) \geq p \Longleftrightarrow P(\eta \leq \tilde{b}(x)) \geq p \Longleftrightarrow \Phi(\tilde{b}(x)) \geq p,$$

where η is defined in (1), Φ refers to its distribution function and $\tilde{b}(x) = b(x) - b$. In order to cope with such constraints, it is important to be able to calculate

Dagstuhl Seminar Proceedings 05031
Algorithms for Optimization with Incomplete Information
http://drops.dagstuhl.de/opus/volltexte/2005/109

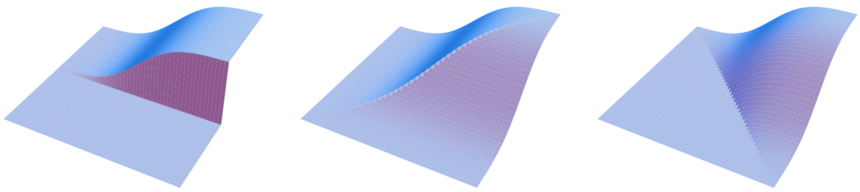

Fig. 1. Distribution functions of 2-dimensional singular normal distributions with covariance matrix having rank one (see text).

values and gradients of singular normal distribution functions. As the latter need not exist in general, it is of interest to characterize differentiability of such functions. If differentiability fails to hold, one could rely on more general tools from nonsmooth optimization (both for algorithmic purposes and optimality conditions). In such constellation, local or global Lipschitz continuity is a favorable property. Whether a singular normal distribution function is discontinuous or not does not depend on the rank of the covariance matrix. Figure 1 shows (from the left to the right) the distribution functions of 2-dimensional normal distributions with zero mean and covariance matrices all of which have rank one:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

2 Results

The following result on (Lipschitz-) continuity is a special case of a characterization provided in [2] for the broader class of so-called quasi-concave measures which singular normal distributions belong to (see [4]).

Theorem 1. *A singular normal distribution function is (globally) Lipschitz continuous if and only if it is continuous and if and only if its covariance matrix does not contain zero elements in the diagonal.*

The Theorem explains why in the three examples of Figure 1, the first distribution function is discontinuous whereas the other two are Lipschitz continuous.

In order to establish a condition for differentiability, consider the system (A, b) of linear inequalities induced by the rows a_i of an (m, s)- matrix A and the components b_i of b:

$$\langle a_i, z \rangle \leq b_i \quad (i = 1, \ldots, m).$$

With (A, b) we associate a family of index sets defined by

$$\begin{aligned} I(A,b) := \{I \subseteq \{1, \ldots, m\} | \exists z \in \mathbb{R}^s : \; & \langle a_i, z \rangle = b_i \quad (i \in I), \\ & \langle a_i, z \rangle < b_i \quad (i \in \{1, \ldots, m\} \backslash I)\}. \end{aligned}$$

The system (A, b) is said to be *nondegenerate*, if

$$\operatorname{rank} \{a_i | i \in I\} = \#I \quad \forall I \in I(A,b).$$

Theorem 2. *Let ξ be an s-dimensional random vector having a nondegenerate normal distribution. Denote by Φ the distribution function of η in (1). Then, Φ is smooth (infinitely many times differentiable) at any point $\bar{x}$ for which the system $(A, \bar{x} - b)$ is nondegenerate.*

The last theorem essentially relies on the following formula for the probability of polyhedra proved in [3] by means of the so-called abstract-tube theory (a recent proof based on more elementary arguments like duality of linear programming can be found in [1]):

Theorem 3. *Let ξ be an s-dimensional random vector. If the system (A, b) is nondegenerate, then the probability of the polyhedron induced by (A, b) equals*

$$P(\langle a_i, \xi \rangle \leq b_i \quad (i = 1, \ldots, m)) = \sum_{I \in I(A,b)} (-1)^{\#I} P(\langle a_i, \xi \rangle > b_i \quad (i \in I)).$$

Specializing this formula to the case of ξ having a (nondegenerate) normal distribution and translating it to the transformed random vector η in (1), one obtains the following formula for the explicit calculation of the singular normal distribution function (of η):

$$\Phi(x) = \sum_{I \in I(A, \bar{x} - b)} (-1)^{\#I} F^I \left(b^I - x^I\right) \quad \forall x \in U. \tag{2}$$

Here, F denotes the (regular) normal distribution function of ξ and the upper index I refers to a selection of corresponding components.

3 Conlusions

If one is able to calculate the index sets $I(A, \bar{x} - b)$ (which amounts to the determination of corners of polyhedra), and if, moreover, a code for the calculation of nondegenerate normal distribution functions is available, then (2) may be used for the calculation of values and gradients (by derivation of (2)) of singular normal distribution functions. First numerical experiments show that this approach is very efficient in moderate dimensions.

References

1. Bukszár, J., Henrion, R., Hujter, M., Szántai,T.: Polyhedral Inclusion-Exclusion, Weierstrass Institute Berlin, Preprint No. 913, 2004.
2. Henrion, R., Römisch, W.: Lipschitz and differentiability properties of quasi-concave and singular normal distribution functions, Preprint 200, DFG Research Center Matheon "Mathematics for key technologies", 2005, submitted.
3. Naiman, D.Q., Wynn, H.P.: Abstract tubes, improved inclusion-exclusion identities and inequalities and importance sampling, *Annals of Statistics* 25 (1997), 1954–1983.
4. Prékopa, A.: *Stochastic Programming*, Kluwer, Dordrecht, 1995.