Abstract tube associated with a perturbed polyhedron and multidimensional normal probability calculation

Satoshi Kuriki, T. Miwa and A. J. Hayter
(Inst. Statist. Math., Tokyo)

Joint 2011, Academia Sinica, ROC

December 16, 2011

Contents of talk

1. Introduction
   - Multidimensional normal probability calculation
   - An integration technique for simplicial cones
2. Abstract tubes
   - Inclusion-exclusion
   - Weak abstract tube (Naiman & Wynn)
3. Our proposal and results
   - Lexicographic perturbation in an outer direction
   - Abstract tube
   - Construction of $\mathcal{F}(\varepsilon)$
   - Studentized range statistic — Numerical example

Summary

1. Introduction

Multidimensional normal probability calculation

- A closed convex polyhedron in $\mathbb{R}^n$:

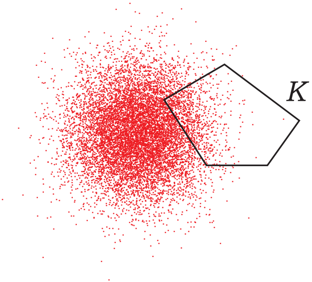
$$K = \{x \in \mathbb{R}^n \mid A^\top x \leq b\}$$

where

$$A = (a_1, \ldots, a_m)_{n\times m}, \quad b = (b_1, \ldots, b_m)^\top_{m\times 1}$$

- Let $x \sim N_n(0, I_n)$, i.e., n-dim standard Gaussian vector. Our primary purpose is to calculate

$$P(K) := \Pr(x \in K)$$



Multidimensional normal probability calculation (cont)

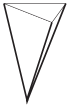
- Application — Multiple comparisons, e.g., distribution of

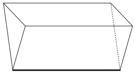
$$\Pr\left(\max_{1\le i<j\le k}\frac{|x_i - x_j|}{\sqrt{\sigma_i^2+\sigma_j^2}} \le c\right)$$

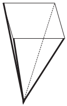
where $x_i \sim N(\mu_i, \sigma_i^2)$ independently.
(Tukey's studentized range statistic)

An integration technique for simplicial cones

- When the column vectors a_i of $A = (a_1, \ldots, a_m)$ are linearly independent, K is a cone.
  Precisely, K is a simplicial cone, or the direct sum of a simplicial cone K_1 and a linear subspace L, i.e., $K = K_1 \oplus L$.







simplical cone

Non-simplical cone

- For such K, Miwa, Hayter and Kuriki (2003, JRSS, B) proposed an "successive numerical integration technique" based on Markov property for computing $\Pr(x \in K)$.
- Implemented as an R library `mvtnorm`
- Then, how to deal with the general case?

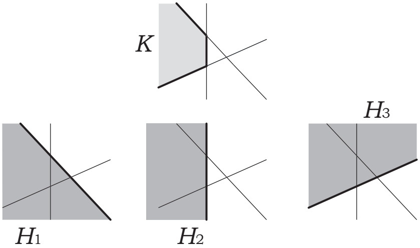
2. Abstract tubes

Inclusion-exclusion

- Abstract tubes = "Polyhedral inclusion-exclusion identity"
- Rewrite the polyhedron K as the intersection of half spaces:

$$K = \bigcap H_i, \quad H_i = \{x \mid a_i^\top x \le b\}$$

- Example: $K = H_1 \cap H_2 \cap H_3$

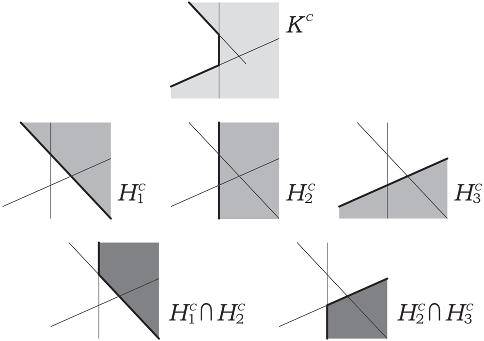


Inclusion-exclusion (cont)

- Since $K^c = H_1^c \cup H_2^c \cup H_3^c$,

$$\begin{aligned} 1 - P(K) = & P(H_1^c) + P(H_2^c) + P(H_3^c) \\ & - P(H_1^c \cap H_2^c) - P(H_2^c \cap H_3^c) \end{aligned}$$

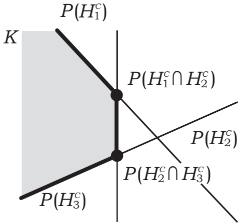
  Note: $-P(H_1^c \cap H_3^c) + P(H_1^c \cap H_2 \cap H_3^c) = 0$



- H_i^c, $H_i^c \cap H_j^c$ are simplicial (or simplicial $\oplus$ linear subspace). "Successive numerical integration technique" can be used.

Inclusion-exclusion (cont)

- $P(H_1^c)$, $P(H_2^c)$, $P(H_3^c)$ correspond to the 3 edges of K. $P(H_1^c \cap H_2^c)$ and $P(H_2^c \cap H_3^c)$ correspond to the 2 vertices of K.



- Note: The identity holds for any probability measure $P(\cdot)$. Not only for Gaussian probability. (e.g., discrete distribution)

Inclusion-exclusion (cont)

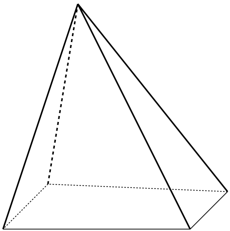
- H_i^c, $H_i^c \cap H_j^c$ are not necessarily simplicial (or simplicial $\oplus$ linear subspace).
- Counter example (Pyramid in $\mathbb{R}^3$)

$$\begin{aligned} H_1 &: \quad -x_1 - x_2 + x_3 \leq 1 \\ H_2 &: \quad -x_1 + x_2 + x_3 \leq 1 \\ H_3 &: \quad +x_1 + x_2 + x_3 \leq 1 \\ H_4 &: \quad +x_1 - x_2 + x_3 \leq 1 \end{aligned}$$

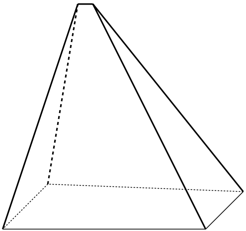
$$\begin{aligned} 1 - P(K) =& P(H_1^c \cup H_2^c \cup H_3^c \cup H_4^c) \\ =& P(H_1^c) + P(H_2^c) + \cdots - P(H_1^c \cap H_2^c) + \cdots \\ & - P(H_1^c \cap H_2^c \cap H_3^c \cap H_4^c) \end{aligned}$$

The last term is not simplicial.

- This difficulty comes from the fact that:
  4 facets in $\mathbb{R}^3$ are not in general position.



Pyramid
(not in general position)



Perturbed pyramid
(in general position)

Weak Absract tube (Naiman & Wynn)

- Perturbed polyhedron

$$K(\epsilon\delta) = \{x \in \mathbb{R}^n \mid A^\top x \leq b + \epsilon\delta\}$$

where $\epsilon \in \mathbb{R}$ and $\delta = (\delta_1, \ldots, \delta_n)^\top$ is a direction vector.

- Let $\mathcal{F}(\epsilon\delta)$ be the set of all faces of $K(\epsilon\delta)$.
- **Proposition** (Naiman & Wynn, 1997)
(i) For a suitable δ and for all ϵ such that $|\epsilon| \ll 1$,

$$1 - P(K) = \sum_{J \in \mathcal{F}(\epsilon\delta)} (-1)^{|J|-1} P(\bigcap_{i \in J} H_i^c)$$

for any continuous probability $P(\cdot)$. "Weak abstract tube"
(ii) $|J| \leq n$

3. Our proposal and results

Lexicographic perturbation in an outer direction

- We propose the use of the following lexicographic perturbation vector

$$\boldsymbol{\varepsilon} = (\varepsilon, \varepsilon^2, \ldots, \varepsilon^n)^\top$$

  where $\varepsilon > 0$ is an infinitesimal positive number.
- Let $\mathcal{F}(\boldsymbol{\varepsilon})$ be the set of all faces of the infinitesimally perturbed polyhedron $K(\boldsymbol{\varepsilon})$.

- **Proposition**

  (i)

$$1 - P(K) = \sum_{J \in \mathcal{F}(\varepsilon)} (-1)^{|J|-1} P\Big(\bigcap_{i \in J} H_i^c\Big)$$

  for all probability measure $P(\cdot)$. “Abstract tube” in the strict sense.

  (ii) $|J| \le \mathrm{rank}(A)$

  (iii) $\bigcap_{i \in J} H_i^c = K_1 \oplus L$, where

$$\begin{aligned} K_1 &: |J|\text{-dim simplicial cone} \\ L &: (n - |J|)\text{-dim linear subspace} \end{aligned}$$

Construction of $\mathcal{F}(\varepsilon)$

- To construction of $\mathcal{F}(\varepsilon)$, for each subset $J \subset \{1, \ldots, m\}$, determine the feasible (exsistence of a solution) of the system:

$$\begin{aligned} a_i^\top x - (b_i + \varepsilon^i) &= 0 \qquad (i \in J) \\ a_i^\top x - (b_i + \varepsilon^i) &\leq 0 \qquad (i \notin J) \end{aligned}$$

  If a solution exists, then $J \in \mathcal{F}(\varepsilon)$. Conducted by the linear programming (LP).
- The term $b_i + \varepsilon^i$ is treated as a polynomial in ε.
  For

$$f(\varepsilon) = \sum c_i \varepsilon^i \quad \text{and} \quad g(\varepsilon) = \sum d_i \varepsilon^i$$

  let

$$f(\varepsilon) \geq g(\varepsilon) \quad \Leftrightarrow \quad (c_0, c_1, \ldots, c_n) \geq (d_0, d_1, \ldots, d_n)$$

  (lexicographically)
- This LP is called lexicographic method.

Studentized range statistic — Numerical example

- The polyhedron defined by the studentized range statistic is

$$K = \left\{ x \in \mathbb{R}^k \;\middle|\; \frac{|\sigma_i x_i - \sigma_j x_j|}{\sqrt{\sigma_i^2 + \sigma_j^2}} \leq c, \ \forall i < j \right\}$$

- In the balanced case $\sigma_1^2 = \ldots = \sigma_k^2$, because

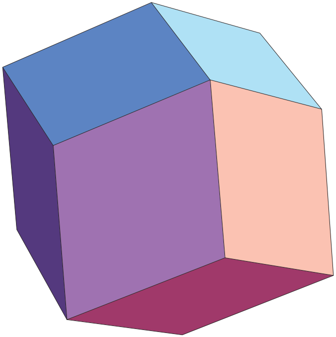
$$\frac{x_1 - x_2}{\sqrt{2}} = c \ \text{ and } \ \frac{x_2 - x_3}{\sqrt{2}} = c \ \text{ and } \ \frac{x_3 - x_4}{\sqrt{2}} = -c$$

$$\text{imply} \ \frac{x_1 - x_4}{\sqrt{2}} = c$$

The facets of K are not in general position.

Studentized range statistic (cont)

- $K_1 = K \cap \{x \mid x_1 = \ldots = x_k\}^\perp$



Studentized range polytope K_1
in the balanced case ($k = 4$)

Studentized range statistic (cont)

- Number of terms in the abstract tube $|\mathcal{F}(\varepsilon)|$

k	2	3	4	5	6
$m = k(k-1)$	2	6	12	20	30
$\lvert\mathcal{F}(\varepsilon)\rvert$	2	12	62	320	1682

m is the number of facets (inequalities)

- Although the lexicographic perturbation depends on the order, $|\mathcal{F}(\varepsilon)|$ looks unchanged even if the order of inequalities is changed.

Studentized range statistic (cont)

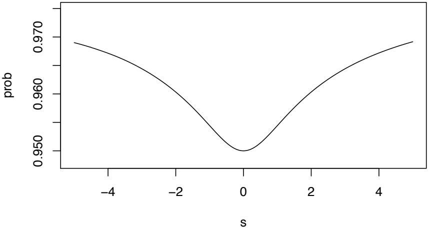
- Hayter (1984) proved the Tukey-Kramer conjecture that

$$\min_{\sigma_1,\ldots,\sigma_k} \Pr(x \in K)$$

is attained iff $\sigma_1 = \ldots = \sigma_k$.

- $k = 5$,

$$(\sigma_i^2) = \left(1, 10^{\frac{1}{4}s}, 10^{\frac{2}{4}s}, 10^{\frac{3}{4}s}, 10^s\right) \quad (s = 0 \Leftrightarrow \sigma_i^2 \equiv 1)$$



Summary

- We discussed a method for computating $\Pr(x \in K)$, where $x \sim N_n(0, I_n)$ and K is any convex polyhedron.
- We proposed the use of the abstract tube with a lexicographic perturbation in an outer direction.
  (i) The lexicographic method of LP is useful in the construction.
  (ii) Each term is simplicial and Miwa, et al. (2003)'s "successive numerical integration technique" can be used for its calculation.
  (iii) The proposed abstract tube is applicable to any probability measures such as discrete distribution.
- Reference:
  S. Kuriki, T. Miwa, and A. J. Hayter (2011), arXiv:1110.2824