Optimal Gaussian Partitions with Application and Open Problems

Elchanan Mossel

UC Berkeley

January 12, 2012

Optimal Gaussian Partitions

Optimal Gaussian Partitions

How to partition

- $\mathbb{R}^n$ (n is unbounded)
- into $r \times q$ parts $f_i^{-1}(a)$ for $1 \leq i \leq r$ and $1 \leq a \leq q$,
- of prescribed Gaussian measures $m_{i,a}$ with $\sum_a m_{i,a} = 1$,
- such that r Gaussian vectors $X_1, \ldots, X_r \in \mathbb{R}^n$ with prescribed covariance structure $Cov(X_i, X_j) = V_{i,j} I_n$
- maximize the expected value of "combinatorial quantity" depending only on $(f_i(X_i))_{i=1}^r$.

Notes

- An asymptotic geometric problem (dimension is unbounded).
- value increases with dimension, maximum is supremum.

Formal definition

Optimal Gaussian Partition

Given:

- $H : [q]^r \to \mathbb{R}$ (combinatorial weights)
- $m \in M_{r \times q}$ a stochastic matrix (parts sizes).
- $0 \le V \in M_{r \times r}$ with $V_{i,i} = 1$ for all i (covariance structure).

Define

$$M(H, m, V) := \sup \mathbb{E}[H(f_1(X_1), \ldots, f_r(X_r))]$$

where the sup is taken over all

- dimensions n,
- $f_i : \mathbb{R}^n \to [q]$ s.t.
- $\mathbb{P}[f_i(X) = a] = m_{i,a}$ for all $1 \le i \le r$ and $1 \le a \le q$.
- $X_1, \ldots, X_r \in \mathbb{R}^n$ are jointly Gaussian with $Cov[X_i, X_j] = V_{i,j} I_n$.

What's known? $q = 2$ parts with $r = 2$

Thm: (C. Borell 1985)

When $r = 2, q = 2$, general m and

$$H(a,b) = 1(a = b), \quad V = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \rho > 0$$

Maximum is obtained in dimension $n = 1$ and

$$f_i(x) = \begin{cases} 1 & x < t. \\ 2 & x \geq t. \end{cases}, \qquad P[X > t] = m_{i,2}.$$

In words

Partition of $\mathbb{R}^n$ into two parts of equal measure which maximizes the probability that two correlated Gaussians will fall in the same part is given by a half-space.

What's known? $q = 2$ parts with general r

Thm: (Isaksson-M 2011)

When $r \geq 2, q = 2, \; m = (m_1, m_2)$,
$H(a, b, c, \ldots) = 1(a = b = c = \ldots)$ and

$$V = \begin{pmatrix} 1 & \rho & \ldots & \rho \\ \rho & 1 & \rho \ldots & \\ \vdots & \ddots & \ddots & \ldots \end{pmatrix}, \rho > 0$$

Maximum is obtained in dimension $n = 1$ and

$$f_i(x) = \begin{cases} 1 & x < t. \\ 2 & x \geq t. \end{cases}, \qquad P[X > t] = m_{i,2}.$$

What else is known?

Nothing.

Proof Techniques

Borell's proof (1985)

Ehrhard symmetrization.

Isaksson-M approach (2011)

- Formulate a spherical statement.
- Prove Spherical Statement using Rearrangement Inequalities.
- Project to a small number of coordinates to obtain Gaussian results

Kindler-O'Donnell (2012)

- Use Gaussian Rotation Invariance + Sub-aditivity.

The Kindler-O'Donnell proof

- Let $P(A) = 1/2$ and $N, M \sim N(0, I)$ with $Cov(N, M) = \rho I$.

$q(\theta, A) := P[N \in A, M \in A^c] = P[N \in A, (\cos\theta)N + (\sin\theta)Z \in A^c]$

$$\textbf{Claim 1: } \forall \text{ integer } k : q(\theta, A) \le kq(\frac{\theta}{k}, A).$$

$$\textbf{Claim 2: } q(\frac{\pi}{2}, A) = \frac{1}{4}$$

$$\implies q(\frac{\pi}{2k}, A) \ge \frac{1}{4k}$$

- For $m = 1(x_1 > 0)$ we have $q(\frac{\pi}{2k}, m) = \frac{1}{4k}$.
- QED (for these special cases).

Isaksson-M approach: Spherical Statement

Spherical Partition Problem

Given n, $0 \le \Sigma \in \mathbb{R}^{k \times k}$, $(m_1, \ldots, m_k) \in (0,1)^k$, Find $\sup P(X_1 \in A_1, \ldots, X_k \in A_k)$ where

- $X'_1, \ldots, X'_k$ are jointly normal with $Cov(X'_i, X'_j) = \Sigma_{i,j} I_n$
- $X_i = \frac{X'_i}{\|X'_i\|_2}$
- sup is over A_i with $\mu(X_i \in A_i) = m_i$ where μ is the Haar measure on the $(n-1)$-sphere.

Thm: Optimal Spherical Partition

If $\Sigma^{-1}_{i,j} \le 0$ for all $i \ne j$ then:

$$P(X_1 \in A_1, \ldots, X_k \in A_k) \le P(X_1 \in H_1, \ldots, X_k \in H_k),$$

where $H_i = \{x : x_1 \le a_1\}$ with $\mu(H_i) = \mu(A_i) = m_i$.

Optimal Spherical Partition - Proof Sketch

Express $P(X_1 \in A_1, \ldots, X_k \in A_k)$ in terms of independent normals $Z_i \sim N(0, c_i I_n)$. Writing $W_i = Z_i / \|Z_i\|_2$ to obtain

$$C_1 \mathbb{E}\left[1_{\{W_1 \in A_1, \ldots, W_k \in A_k\}} \prod_{1 \le i < j \le k} e^{-\left(\Sigma^{-1}\right)_{i,j} \langle Z_i, Z_j \rangle}\right] =$$

$$C_1 \mathbb{E}\left[1_{\{W_1 \in A_1, \ldots, W_k \in A_k\}} \prod_{1 \le i < j \le k} e^{-\left(\Sigma^{-1}\right)_{i,j} \langle W_i, W_j \rangle \|Z_i\|_2 \|Z_j\|_2}\right]$$

Conditioned on $\|Z_i\|_2$, W_i are uniformly distributed on the sphere and $\langle W_i, W_j \rangle$ decreases in $\|W_i - W_j\|$.
Therefore can apply extended Riesz Inequality (Burchard-01, Morpurgo-02) to conclude maximum is obtained for half-spaces H_i.

Optimal Gaussian Partitions

- Take $n \leq m \to \infty$.
- $X_i \in S^{m-1}$, $Y_i \in R^n$ with the same covariance structure Σ.
- $Z_i =$ first n coordinates of X_i.
- $\sqrt{m}(Z_1, \ldots, Z_k) \to_{m \to \infty} (Y_1, \ldots, Y_k)$ in distribution.
- Spherical bound implies Gaussian bound.
- Some approximation agruments needed when sets are not closed.

Open Problem 1 - Finite Dimensionality?

1. Finite dimensionality

Is the supremum $M(H, m, V)$ a maximum? Is it obtained in a finite dimension?

1.a Finite dimensionality variant

Same question assuming $f_s = f_1$ and $m_{s,j} = m_{1,j}$ for $1 \le s \le r$? (Conj. of O. Regev: $n = \infty$ for $r = 2, q = 2, H(a, b) = 1(a \neq b)$).

Comment : Approximate Finite Dimensionality

Find explicit $n(\epsilon, H)$ or $n(\epsilon, H, m, V)$ such that sup in dimension n is ϵ close to $M(H, m, V)$? (Seems doable using dimension reduction ideas (see Raghavendra-Steurer-09)).

Open Problem 2 - Other Optimal partitions?

More Examples

Find other optimal Gaussian partitions!

The Standard Simplex Conjecture (Isaksson-M-11)

Suppose $X, Y \sim N(0, I_n)$ and $Cov(X, Y) = \rho I_n$. Let $A_1, \ldots, A_q \subseteq \mathbb{R}^n$ be a partition of $\mathbb{R}^n$ and $S_1, \ldots, S_q \subseteq \mathbb{R}^n$ a standard simplex partition. Then,

i) If $\rho \geq 0$ and $A_1, \ldots, A_q$ is *balanced*, then

$$\mathbb{P}((X, Y) \in A_1^2 \cup \cdots \cup A_q^2) \leq \mathbb{P}((X, Y) \in S_1^2 \cup \cdots \cup S_q^2) \quad (1)$$

ii) If $\rho < 0$:

$$\mathbb{P}((X, Y) \in A_1^2 \cup \cdots \cup A_q^2) \geq \mathbb{P}((X, Y) \in S_1^2 \cup \cdots \cup S_q^2) \quad (2)$$

The Standard Simplex Partition

definition

For $n+1 \geq q \geq 2$, $A_1, \ldots, A_q$ is a *standard simplex partition* of $\mathbb{R}^n$ if for all i

$$A_i \supseteq \{x \in \mathbb{R}^n | x \cdot a_i > x \cdot a_j, \forall j \neq i\} \tag{3}$$

where $a_1, \ldots a_q \in \mathbb{R}^n$ are q vectors satisfying

$$a_i \cdot a_j = \begin{cases} 1 & \text{if } i = j \\ -\frac{1}{q-1} & \text{if } i \neq j \end{cases} \tag{4}$$