Topics in the Local Theory of Normed Spaces
Lecture 9

1 Gaussian process and Slepian's lemma

Let X be an n-dimensional subspace of $L_p(\Omega, \mu)$, where $\Omega = \{1, \ldots, m\}$. Our goal will be to find a sequence of signs $\{\epsilon_i\}_{i=1}^m$ such that for every $x = (x_1, \ldots, x_m)$ the expression

$$\sup_{x \in X,\ \|x\| \le 1} \left| \sum_{i=1}^m \epsilon_i |x_i|^p \mu(i) \right|$$

will be "very small" as a function of $\frac{n}{m}$. It's enough for us to show that in average this quantity behaves nicely, meaning

$$\mathbb{E}_\epsilon \sup_{x \in X,\ \|x\| \le 1} \left| \sum_{i=1}^m \epsilon_i |x_i|^p \mu(i) \right| \le \delta(n, m),$$

where $\delta(n, m)$ is "very small". Our first step will be to pass from signs $\epsilon = \pm 1$ to standard gaussian variables.

Claim 1.1. *Let $\{g_i\}_{i=1}^m$ be independent identically distributed normal random variables. Then*

$$\mathbb{E}_\epsilon \sup_{x \in X,\ \|x\| \le 1} \left| \sum_{i=1}^m \epsilon_i |x_i|^p \mu(i) \right| \le \sqrt{\frac{2}{\pi}} \cdot \mathbb{E} \sup_{x \in X,\ \|x\| \le 1} \left| \sum_{i=1}^m g_i |x_i|^p \mu(i) \right|$$

Proof. First of all we use the fact that the distribution of g_i is the same as the distribution of $|g_i|\epsilon_i$, where $\mathbf{P}(\epsilon_i = 1) = \mathbf{P}(\epsilon_i = -1) = \frac{1}{2}$ and $g_1, \ldots, g_m, \epsilon_1, \ldots, \epsilon_m$ are independent. Hence,

$$\mathbb{E} \sup_{x \in X,\ \|x\| \le 1} \left| \sum_{i=1}^m g_i |x_i|^p \mu(i) \right| = \mathbb{E}_\epsilon \mathbb{E}_g \sup_{x \in X,\ \|x\| \le 1} \left| \sum_{i=1}^m |g_i| \epsilon_i |x_i|^p \mu(i) \right| \ge$$

$$\ge \mathbb{E}_\epsilon \sup_{x \in X,\ \|x\| \le 1} \left| \mathbb{E}_g \sum_{i=1}^m |g_i| \epsilon_i |x_i|^p \mu(i) \right| = \sqrt{\frac{2}{\pi}} \cdot \mathbb{E}_\epsilon \sup_{x \in X,\ \|x\| \le 1} \left| \sum_{i=1}^m \epsilon_i |x_i|^p \mu(i) \right|,$$

where the inequality is the result of the convexity of the supremum function and Jensen's inequality.

■

Definition 1.2. *A Gaussian process $\{G_t\}_{t\in T}$ is a collection of random variables indexed by some finite or infinite set T such that every finite linear combination of it $\sum a_j G_{t_j}$ is normally distributed.*
A Gaussian vector is a finite sequence $(X_1, \ldots, X_k)$ of random variables such that any linear combination of them $\sum a_j X_j$ is normally distributed.
A Gaussian vector centered at zero is just a Gaussian vector satisfying $\mathbb{E}X_i = 0$ for all i.

Claim 1.3. *$X = (X_1, \ldots, X_k)$ is a Gaussian vector centered at zero if and only if there exists a k independent distributed $N(0,1)$ random variables $G = (G_1, \ldots, G_k)$ and a matrix $T = (t_{ij})_{i,j=1}^k$ satisfying*

$$X = TG,$$

or in other words $X_i = \sum_{j=1}^k t_{ij} G_j$ for all $i = 1, \ldots, k$.

Remark: Clearly, linear combination of random variable of the form $X_i = \sum_{j=1}^k t_{ij} g_j$ is Gaussian with mean 0.
In the opposite direction, consider $X_1, \ldots, X_k \in L_2(\Omega, \mathbf{P})$, where $\|X\|_{L_2(\Omega,\mathbf{P})} = (\mathbb{E}|X|^2)^{1/2}$. The dimension of the subspace $span\{X_1, \ldots, X_k\}$ is $\leq k$ and g_i-s are members of the orthonormal basis of that space.

Consider the covariance matrix of $X_1, \ldots, X_k$, $\Gamma = (\gamma_{ij})_{i,j=1}^k$, where $\gamma_{ij} = \mathbb{E}X_i X_j$. Then using the last claim we have

$$\gamma_{uv} = \mathbb{E}(\sum_{j=1}^k t_{uj} g_j)(\sum_{l=1}^k t_{vl} g_l) = \sum_{j,l} t_{uj} t_{vl} \mathbb{E}(g_j g_l) = \sum_{j=1}^k t_{uj} t_{vj} = (TT^*)_{uv}.$$

What is the density function of $X = (X_1, \ldots, X_k)$? Suppose, that T (and Γ) is regular (equivalently $\dim span\{X_1, \ldots, X_k\} = k$). Then, for every set $A \in \mathbb{R}^k$

$$\begin{aligned}\mathbf{P}(X \in A) = \mathbf{P}(G \in T^{-1}A) &= \frac{1}{(2\pi)^{k/2}} \int_{T^{-1}A} e^{-<u,u>/2} du = \\ &= \frac{1}{(2\pi)^{k/2}} \frac{1}{\det(T)} \int_A e^{-<T^{-1}v, T^{-1}v>/2} dv = \\ &= \frac{1}{(2\pi)^{k/2}} \frac{1}{\det(\Gamma)^{1/2}} \int_A e^{-<\Gamma^{-1}v, v>/2} dv,\end{aligned}$$

where we used the fact that $< T^{-1}v, T^{-1}v >=< (T^*)^{-1}T^{-1}v, v >=< \Gamma^{-1}v, v >$. Hence we showed, that the density of X is

$$f_X(v) = f_\Gamma(v) = \frac{1}{(2\pi)^{k/2}} \frac{1}{\det(\Gamma)^{1/2}} e^{-<\Gamma^{-1}v,v>/2}. \tag{1.1}$$

Next, we will compute the characteristic function of X. Let $\xi \in \mathbb{R}^k$. Then

$$\widehat{f}(\xi) = \mathbb{E}e^{i<\xi,X>} = \frac{1}{(2\pi)^{k/2}} \frac{1}{(\det\Gamma)^{1/2}} \int_{\mathbb{R}^k} e^{i<\xi,v>} e^{-<\Gamma^{-1}v,v>/2} dv$$

In the case of X being a vector of independent standard normally distributed variables ($\Gamma = I$)

$$\widehat{f_I}(\xi) = \frac{1}{(2\pi)^{k/2}} \int_{\mathbb{R}^k} e^{i<\xi,v>} e^{-|v|^2} dv = \frac{1}{(2\pi)^{k/2}} \prod_{j=1}^{k} \int_{\mathbb{R}} e^{i\xi_j v_j - v_j^2/2} dv_j =$$
$$= \frac{1}{(2\pi)^{k/2}} \prod_{j=1}^{k} \int_{\mathbb{R}} e^{-(v_j - i\xi_j)^2/2} dv_j \prod_{j=1}^{k} e^{-\xi_j^2/2} = e^{-|\xi|^2/2}.$$

Making the change of variables $v = Tu$ and using the result for $\Gamma = I$ we get

$$\widehat{f}(\xi) = \frac{1}{(2\pi)^{k/2}} \int_{\mathbb{R}^k} e^{i<\xi,Tu>} e^{-|u|^2} du = e^{-|T^*\xi|^2/2} = e^{-<TT^*\xi,\xi>/2} = e^{-<\Gamma\xi,\xi>/2}.$$

Lemma 1.4. *(Slepian's lemma) Let $X = (X_1, \ldots, X_k)$ and $Y = (Y_1, \ldots, Y_k)$ be two Gaussian vectors centered at zero. Assume*

1. $\mathbb{E}X_i^2 = \mathbb{E}Y_i^2$.
2. $\mathbb{E}X_iX_j \le \mathbb{E}Y_iY_j$ *for all* $i \neq j$.

Then for all $t \in \mathbb{R}$

$$\mathbf{P}(\max X_i > t) \ge \mathbf{P}(\max Y_i > t)$$

and in particlar

$$\mathbb{E}(\max X_i) \ge \mathbb{E}(\max Y_i).$$

Proof. Without loss of generality assume that X and Y are independent. For every $0 < \Theta < 1$ define

$$X_\Theta = \Theta X + (1 - \Theta^2)^{1/2} Y,$$

and for every $1 \le i \le k$ define

$$X_{\Theta_i} = \Theta X_i + (1 - \Theta^2)^{1/2} Y_i.$$

Clearly, X_Θ is also a gaussian vector. Let Γ_X and Γ_Y be the covariance matrices of X and Y. Then the covariance matrix of X_Θ is

$$\Gamma_\Theta(i,j) = \mathbb{E}X_{\Theta_i}X_{\Theta_j} = \mathbb{E}(\Theta X_i + (1-\Theta^2)^{1/2}Y_i)(\Theta X_j + (1-\Theta^2)^{1/2}Y_j) = \\ = \Theta^2\Gamma_X(i,j) + (1-\Theta^2)\Gamma_Y(i,j).$$

Hence,

$$\Gamma_\Theta = \Theta^2\Gamma_X + (1-\Theta^2)\Gamma_Y.$$

Denote by $f_\Theta = f_{\Gamma_\Theta}$ the density function of X_Θ. Fix $t_1, \dots, t_k \in \mathbb{R}$ and consider $\mathbf{P}(\Theta) = \mathbf{P}(X_{\Theta_1} \le t_1, \dots, X_{\Theta_k} \le t_k)$:

$$\mathbf{P}(\Theta) = \int_{-\infty}^{t_1} \dots \int_{-\infty}^{t_k} f_\Theta(u)du_1 \dots du_k.$$

In order to prove the lemma, it is enough to show that $\mathbf{P}'(\Theta) \le 0$ for all $0 < \Theta < 1$. So we have to evaluate

$$\mathbf{P}'(\Theta) = \int_{-\infty}^{t_1} \dots \int_{-\infty}^{t_k} \sum_{i,j} \frac{\partial f_\Theta(u)}{\partial \gamma_\Theta(i,j)} \frac{d\gamma_\Theta(i,j)}{d\Theta} du_1 \dots du_k.$$

First of all by definition of $\gamma_\Theta(i,j)$ and by the second condition of the Lemma

$$\frac{d\gamma_\Theta(i,j)}{d\Theta} = 2\Theta(\gamma_X(i,j) - \gamma_Y(i,j)) \le 0,$$

for $i \ne j$ and $\frac{d\gamma_\Theta(i,j)}{d\Theta} = 0$ if $i = j$.

Suppose, $i \ne j$. Then

$$\frac{\partial f_\Theta(u)}{\partial \gamma_\Theta(i,j)} = \frac{\partial}{\partial \gamma_\Theta(i,j)} c_k \int_{\mathbb{R}^k} e^{-i<\xi,v>} e^{-<\Gamma\xi,\xi>/2} d\xi = -c_k \int_{\mathbb{R}^k} e^{-i<\xi,v>} e^{-<\Gamma\xi,\xi>/2} \xi_i \xi_j d\xi.$$

On the other hand

$$\frac{\partial^2 f_\Theta(v)}{\partial v_i \partial v_j} = \frac{\partial}{\partial v_j} c_k \int_{\mathbb{R}^k} -i\xi_i e^{-i<\xi,v>} e^{-<\Gamma\xi,\xi>/2} d\xi = \\ = c_k \int_{\mathbb{R}^k} \xi_i \xi_j e^{-i<\xi,v>} e^{-<\Gamma\xi,\xi>/2} d\xi = \frac{\partial f_\Theta(u)}{\partial \gamma_\Theta(i,j)}.$$

Substituting we get

$$\mathbf{P}'(\Theta) = \sum_{i \ne j} \int_{-\infty}^{t_1} \dots \int_{-\infty}^{t_k} \frac{\partial^2 f_\Theta(v)}{\partial v_i \partial v_j} 2\Theta(\gamma_X(i,j) - \gamma_Y(i,j)) dv_1 \dots dv_k = \\ = \sum_{i \ne j} \int_{-\infty}^{t_1} \dots \int_{-\infty}^{t_k} f_\Theta(v_1, \dots, t_i, \dots, t_j, \dots, v_k) 2\Theta(\gamma_X(i,j) - \gamma_Y(i,j)) dv_1 \dots dv_k,$$

where the integration is over $(k-2)$ variables $(v_l)_{l=1,l\neq i,j}^{k}$. This is clearly ≤ 0. Which means that

$$\mathbf{P}(0) = \mathbf{P}(Y_1 \leq t_1, \ldots, Y_k \leq t_k) \geq \mathbf{P}(X_1 \leq t_1, \ldots, X_k \leq t_k) = \mathbf{P}(1).$$

This is even stronger statement than we need to prove

$$\mathbf{P}(\max X_i > t) \geq \mathbf{P}(\max Y_i > t).$$

To prove the second part of the lemma, we approximate X_is and Y_is by the random variables bounded from below by some M. Hence, $X_i - M$ and $Y_i - M$ are positive random variable and the second statement of the lemma follows

$$\mathbb{E}(maxX_i - M) = \int_0^\infty \mathbf{P}(maxX_i > t + M) \geq \int_0^\infty \mathbf{P}(maxY_i > t + M) = \mathbb{E}(maxY_i - M).$$

■

Lemma 1.4 has the following geometrical meaning. Consider in $\mathbb{R}^k$ two families of balls $\{B(x_i, r_i)\}_{i=1}^l$ and $\{B(y_i, r_i)\}_{i=1}^l$. If for every $i \neq j$, $\|x_i - x_j\|_2 \geq \|y_i - y_j\|_2$ then according to Lemma 1.4

$$Vol\left(\bigcap_{i=1}^{l} B(x_i, r_i)\right) \leq Vol\left(\bigcap_{i=1}^{l} B(y_i, r_i)\right).$$

References

[MS] V. Milman and G. Schechtman, *Asyptotic theory of finite-dimensional normed spaces*, Lecture Notes in Mathematics, 1200, Springer-Verlag, Berlin, 1986

[P] G. Pisier, *The volumes of convex bodies and Banach space geometry*, Cambridge University Press, Cambridge 1989

[S] G. Schechtman, *Concentration, results and applications*, available on-line at: http://www.wisdom.weizmann.ac.il/ gideon/papers/concentrationNov19.ps