High-Dimensional Probability and Statistics

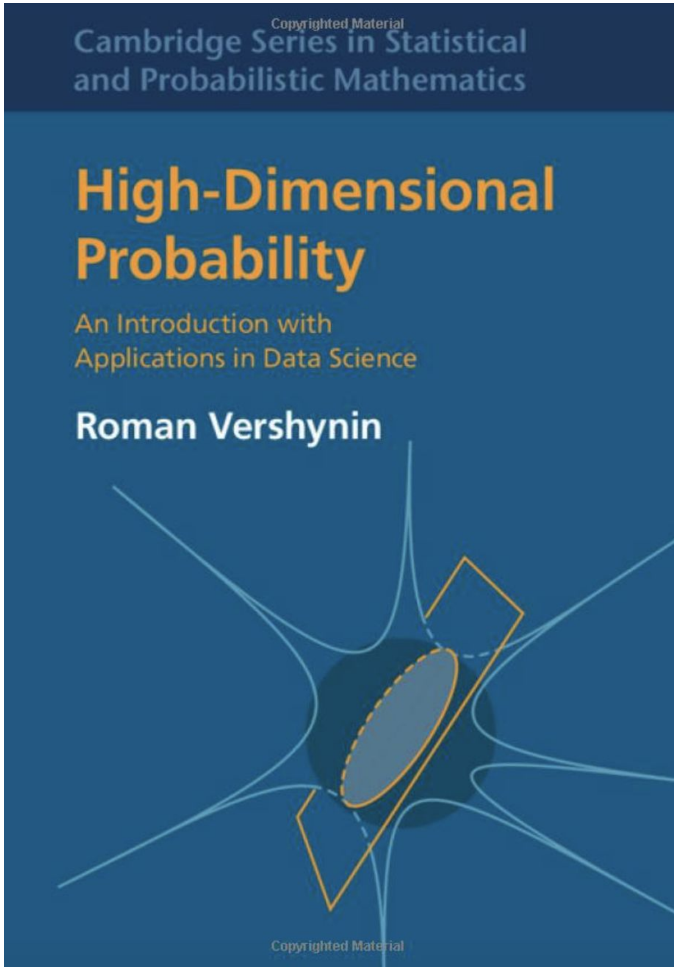
MATH/STAT/ECE 888: Topics in Mathematical Data Science

Sebastien Roch (Math+Stat)
UW-Madison

Fall 2021

Lecture 19 (10/20/21)

Today’s slides based on Vershynin



Isometries

Distortion

The extreme singular values $s_1(A)$ and $s_n(A)$ have an important geometric meaning. They are respectively the smallest number M and the largest number m that make the following inequality true:

$$m\|x\|_2 \leq \|Ax\|_2 \leq M\|x\|_2 \quad \text{for all } x \in \mathbb{R}^n. \tag{4.5}$$

(Check!) Applying this inequality for $x-y$ instead of x and with the best bounds, we can rewrite it as

$$s_n(A)\|x-y\|_2 \leq \|Ax - Ay\|_2 \leq s_1(A)\|x-y\|_2 \quad \text{for all } x \in \mathbb{R}^n.$$

This means that the matrix A, acting as an operator from $\mathbb{R}^n$ to $\mathbb{R}^m$, can only change the distance between any points by a factor that lies between $s_n(A)$ and $s_1(A)$. Thus the extreme singular values control the *distortion* of the geometry of $\mathbb{R}^n$ under the action of A.

Isometries

Exercise 4.1.4 (Isometries). ☕ Let A be an $m \times n$ matrix with $m \geq n$. Prove that the following statements are equivalent.

(a) $A^\mathsf{T} A = I_n$.
(b) $P := AA^\mathsf{T}$ is an *orthogonal projection*[1] in $\mathbb{R}^m$ onto a subspace of dimension n.
(c) A is an *isometry*, or isometric embedding of $\mathbb{R}^n$ into $\mathbb{R}^m$, which means that

$$\|Ax\|_2 = \|x\|_2 \quad \text{for all } x \in \mathbb{R}^n.$$

(d) All singular values of A equal 1; equivalently

$$s_n(A) = s_1(A) = 1.$$

Approximate isometries

Quite often the conditions of Exercise 4.1.4 hold only approximately, in which case we regard A as an *approximate isometry.*

Lemma 4.1.5 (Approximate isometries)**.** *Let A be an $m \times n$ matrix and $\delta > 0$. Suppose that*

$$\|A^\mathsf{T} A - I_n\| \le \max(\delta, \delta^2).$$

Then

$$(1-\delta)\|x\|_2 \le \|Ax\|_2 \le (1+\delta)\|x\|_2 \quad \textit{for all } x \in \mathbb{R}^n. \tag{4.6}$$

Consequently, all singular values of A are between $1-\delta$ and $1+\delta$:

$$1-\delta \le s_n(A) \le s_1(A) \le 1+\delta. \tag{4.7}$$

Approximate isometries cont'd

Proof To prove (4.6), we may assume without loss of generality that $\|x\|_2 = 1$. (Why?) Then, using the assumption, we get

$$\max(\delta, \delta^2) \geq \left|\langle (A^\top A - I_n)x, x\rangle\right| = \left|\|Ax\|_2^2 - 1\right|.$$

Applying the elementary inequality

$$\max(|z-1|, |z-1|^2) \leq |z^2 - 1|, \quad z \geq 0 \tag{4.8}$$

for $z = \|Ax\|_2$, we conclude that

$$|\|Ax\|_2 - 1| \leq \delta.$$

This proves (4.6), which in turn implies (4.7) as we saw in the beginning of this section. □

Converse

Exercise 4.1.6 (Approximate isometrics). ☕☕ Prove the following converse to Lemma 4.1.5: if (4.7) holds, then

$$\|A^\mathsf{T} A - I_n\| \leq 3 \max(\delta, \delta^2).$$

[1] Recall that P is a projection if $P^2 = P$, and P is called orthogonal if the image and kernel of P are orthogonal subspaces.