18 Mar 2022

Estimating singular vectors from samples, Part I

Announcements:

1. Quiz 2 on Wed 3/23 in class.
2. Write to rdk2, as2626 if you have a time conflict.
3. Practice problems on Ed.
4. Solutions, formula sheet forthcoming on Ed.
5. If you're on Zoom and can't hear me, type in chat that you can't hear.

Problem to solve. We are given i.i.d. samples $a_1, a_2, \ldots, a_n \in \mathbb{R}^d$ each has $N(0, BB^T)$ distribution. Try to find a unit vector in $\mathbb{R}^d$ that is close to the top right sing. vector of B^T, v_1.

Goal. Output $\hat{v} \in \mathbb{R}^d$ s.t. $\|\hat{v}\|_2 = 1$ and

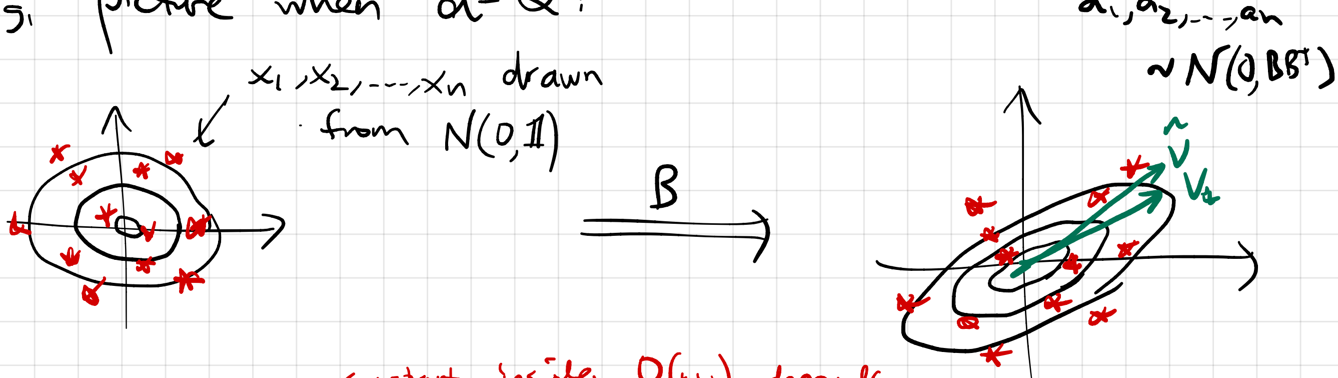
$$|\langle \hat{v}, v_1 \rangle| > 1 - \varepsilon.$$

Succeed with probability $1-\delta$.

How large does n need to be??
If n is large enough, what algorithm should we use?

Algorithm will be: form the data matrix A^T whose rows $a_1^T, a_2^T, \dots, a_n^T$ are the row vectors formed from the data points.
Output $\hat{v}$ = top right sing vector of A^T.

E.g. picture when $d=2$:



constant inside $O(\dots)$ depends on σ_1, σ_2, tends to ∞ as $\sigma_1 - \sigma_2 \to 0$.

Goal: show $n \approx O_{\sigma_1,\sigma_2}\left(d \log\left(\frac{d}{\varepsilon \delta}\right)\right)$ samples suffice.

Game plan:

1. Analyze distribution of $\|A^T x\|^2$ for an arbitrary unit vector x.

$$A^T x = \begin{bmatrix} \langle a_1, x\rangle \\ \langle a_2, x\rangle \\ \vdots \\ \langle a_n, x\rangle \end{bmatrix} \qquad \|A^T x\|^2 = \sum_{i=1}^{n} \langle a_i, x\rangle^2$$

Prove $\sum_{i=1}^{n} \langle a_i, x\rangle^2$ is unlikely to be far from its expected value.

Step 2. $\mathbb{E}\left[\langle a_i, x\rangle^2\right] = \mathbb{E}\left[\langle Bx_i, x\rangle^2\right]$ $\qquad x_i \sim N(0,1)$

$$= \mathbb{E}\left[\langle x_i, B^T x\rangle^2\right]$$

has distribution $N(0, \|B^T x\|^2)$

$$= \|B^T x\|^2$$

Step 3. Singular vector v_1 maximizes $\|B^T x\|^2$ among all unit vectors x.

And $\hat{v}$ maximizes $\|A^T x\|^2$ among all unit vectors x.

In Step 2 we calculated $\mathbb{E}\|A^T x\|^2 = n\|B^T x\|^2$.

In step 1 we asserted that the actual (random) value of $\|A^T x\|^2$, for any individual x, is likely to be close to its expectation.

Step 4. Approximate $\{$all unit vectors in $\mathbb{R}^d\}$ with a finite subset $\{x_1, \dots, x_m\}$.

Argue that if $\|A^T x\|^2 \overset{1\pm\frac{\varepsilon}{4}}{\approx} \mathbb{E}\|A^T x\|^2$ for all $x \in \{x_1, \dots, x_m\}$ then

$\|A^T x\|^2 \overset{1\pm\frac{\varepsilon}{2}}{\approx} \mathbb{E}\|A^T x\|^2$ for all unit vectors x.

$$m = \left(\frac{4d^2}{\varepsilon^2}\right)^d \quad \text{suffices.}$$