Restricted Maximal Singular Value of Randomized Low-rank Approximations

Jonathan Lacotte
Stanford University

Notations

We say that $S \in \mathbb{R}^{d \times m}$ is a test matrix if S has i.i.d. standard normal entries. We denote the set of orthogonal matrices operating over $\mathbb{R}^d$ by $\mathcal{O}_d$.

For two real-valued random variables X and Y, we say that X is stochastically dominated by Y if $\mathbb{P}(X \geqslant \tau) \leqslant \mathbb{P}(Y \geqslant \tau)$ for any $\tau \in \mathbb{R}$, and we write $X \overset{d}{\leqslant} Y$. In particular, this ordering relationship is transitive, i.e., if $X \overset{d}{\leqslant} Y$ and $Y \overset{d}{\leqslant} Z$, then $X \overset{d}{\leqslant} Z$.

Given a bounded set $T \subset \mathbb{R}^d$, we define the Gaussian width of T as $w(T) = \mathbb{E}\left[\sup_{x \in T} \langle g, x \rangle\right]$, where $g \sim \mathcal{N}(0, I_d)$. The radius of T is defined as $\mathrm{rad}(T) = \sup_{x \in T} \|x\|_2$.

Problem statement

Let $A \in \mathbb{R}^{n \times d}$ be a given matrix, with rank r. Let $\mathcal{C}$ be a compact subset of $\mathbb{R}^d$. We introduce a functional of interest, and which depends on the matrix A and a test matrix S, that is,

$$f(A, S) = (I - P_{AS})\, A. \tag{1}$$

where $P_{AS} = AS(S^\top A^\top AS)^\dagger S^\top A^\top$ is the linear orthogonal projector onto the range of AS. We are interested in characterizing the maximal singular value of $f(A, S)$ restricted to the set $\mathcal{C}$,

$$\sigma_{\mathcal{C}}(A, S) = \sup_{x \in \mathcal{C}} \|f(A, S)x\|_2. \tag{2}$$

Example 1. Consider a vector of observations $y = Ax^* + w$ where $x^* \in \mathbb{R}^d$ is a planted vector known to lie in a constraint set $\mathcal{C}$, and $w \sim \mathcal{N}_n(0, \sigma^2 I)$. We aim to recover the in-sample responses Ax^*. Given a low-rank matrix approximation $A \approx QQ^\top A$ where $Q \in \mathbb{R}^{n \times m}$ has orthogonal columns, one might prefer, from a computational standpoint, to compute the low-dimensional solution $\widehat{\alpha} := \mathrm{argmin}_{\alpha \in \mathbb{R}^m} \|Q\alpha - y\|_2^2$, and then use $Q\widehat{\alpha}$ as an estimator of Ax^*. The worst-case risk of $Q\widehat{\alpha}$ is given by

$$\sup_{x^* \in \mathcal{C}} \mathbb{E}_w \left[\|Q\widehat{\alpha} - Ax^*\|_2^2\right] = m\sigma^2 + \sup_{x \in \mathcal{C}} \|(I - QQ^\top)A\|_2^2\,. \tag{3}$$

$\square$

Analysis

Let $A = UDV^\top$ be a singular value decomposition of A, where $U \in \mathbb{R}^{n \times r}$ and $V \in \mathbb{R}^{d \times r}$ have orthonormal columns, and $D \in \mathbb{R}^{r \times r}$ is a diagonal matrix, with entries $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_r > 0$.

Let us fix an integer $1 \leqslant k \leqslant \frac{r}{2}$ – which we refer to as the target rank –, and an oversampling parameter $p \geqslant 1$. Then, we consider the sketch size $m = k + p$, and we introduce the block decomposition

$$D = \begin{pmatrix} D_k & 0 \\ 0 & D_{r-k} \end{pmatrix}, \qquad V = \begin{bmatrix} V_r & V_{r-k} \end{bmatrix},$$

where $D_k = \mathrm{diag}(\sigma_1, \ldots, \sigma_k)$, $D_{r-k} = \mathrm{diag}(\sigma_{k+1}, \ldots, \sigma_r)$, $V_k \in \mathbb{R}^{d \times k}$ consists of the first k columns of V and $V_{r-k} \in \mathbb{R}^{d \times (r-k)}$ of its last $(r-k)$ columns.

Lemma 1. *For any test matrix $S \in \mathbb{R}^{d \times m}$, it holds that $f(A, S) = U f(\Sigma, V^\top S) V^\top$. Setting $G = V^\top S$, it follows, by rotational invariance of the Gaussian distribution, that $G \in \mathbb{R}^{r \times m}$ is a Gaussian test matrix, and hence,*

$$f(A, S) \overset{d}{=} U f(D, G) V^\top. \tag{4}$$

Consequently,

$$\sigma_{\mathcal{C}}(A, S) \overset{d}{=} \sup_{x \in V^\top \mathcal{C}} \|f(D, G)x\|_2. \tag{5}$$

We introduce a family of auxiliary matrices, which will be useful in our analysis, and which were first proposed in [6]. For $t > 0$, we set

$$M(t) = \begin{pmatrix} tI_k & 0 \\ 0 & D_{r-k} \end{pmatrix}. \tag{6}$$

We observe that $D \preceq M(t)$ for $t \geqslant \sigma_1$.

Lemma 2. *Let $\Sigma_1, \Sigma_2 \in \mathbb{R}^{r \times r}$ be two diagonal matrices. If $\Sigma_1 \preceq \Sigma_2$, then, for any $w \in \mathbb{R}^r$,*

$$\|f(\Sigma_1, G)w\|_2 \leqslant \|f(\Sigma_2, G)w\|_2. \tag{7}$$

Consequently, for any $t \geqslant \sigma_1$,

$$\sigma_{\mathcal{C}}(A, S) \overset{d}{\leqslant} \sup_{w \in V^\top \mathcal{C}} \|f(M(t), G)w\|_2. \tag{8}$$

The following fact has been proved in [6], and is key to our analysis. Let $G \in \mathbb{R}^{r \times m}$ be a test matrix. Then, it holds that

$$\lim_{t \to +\infty} f(M(t), G) = \begin{bmatrix} 0 & 0 \\ -f(D_{r-k}, X_2) X_1 \Sigma^{-1} \Omega & f(D_{r-k}, X_2) \end{bmatrix}, \tag{9}$$

where $X_1 \in \mathbb{R}^{(r-k) \times k}$, $X_2 \in \mathbb{R}^{(r-k) \times p}$, $\Sigma \in \mathbb{R}^{k \times k}$ and $\Omega \in \mathbb{R}^{k \times k}$ are independent random matrices. Further, X_1 and X_2 have independent i.i.d. Gaussian entries, Σ is diagonal with diagonal entries distributed as the first k singular values of a $k \times m$ Gaussian test matrix, and $\Omega \overset{d}{\sim} \mathrm{Unif}(\mathcal{O}_k)$. Combining (8) and Lemma 2, we immediately obtain the next result.

Lemma 3. *For any test matrix $G \in \mathbb{R}^{r \times m}$, and $t > 0$, we have*

$$f(M(t), G) \leqslant \mathit{rad}(V_k^\top \mathcal{C}) \, \|f(D_{r-k}, X_2) X_1 \Sigma^{-1} \Omega\|_2 + \sup_{x \in \mathcal{C}} \|D_{r-k} V_{r-k}^\top x\|_2 . \tag{10}$$

Thus, in order to provide a high-probability upper bound on the restricted singular value $\sigma_{\mathcal{C}}(A, G)$, it suffices to control the right-hand side of (10). First, we focus on the term $\|f(D_{r-k}, X_2) X_1 \Sigma^{-1} \Omega\|_2$. A classical result – known as Chevet inequality [2], whose constants were improved by Gordon [3] – yields that, with high-probability,

$$\|f(D_{r-k}, X_2) X_1 \Sigma^{-1} \Omega\|_2 \leqslant \|D_{n-k}\|_2 \|\Sigma^{-1}\|_F + \|D_{n-k}\|_F \|\Sigma^{-1}\|_2 . \tag{11}$$

We leverage standard facts on the singular values of Gaussian matrices (see Propositions A.4 and A.6 in [4]). Choosing an oversampling parameter $p = k+1$, we get that, with high probability,

$$\|f(D_{r-k}, X_2) X_1 \Sigma^{-1} \Omega\|_2 \leqslant c_0 \left(\sigma_{k+1} + \sqrt{\frac{1}{k} \sum_{j=k+1}^{r} \sigma_j^2} \right) . \tag{12}$$

where c_0 is some universal constant. Combining everything together, we obtain the following result.

Theorem 1. *Choosing an oversampling parameter $p = k+1$, it holds that*

$$\sigma_{\mathcal{C}}(A, S) \lesssim \mathit{rad}(V_k^\top \mathcal{C}) \left(\sigma_{k+1} + \sqrt{\frac{1}{k} \sum_{j=k+1}^{r} \sigma_j^2} \right) + \sup_{x \in \mathcal{C}} \|D_{r-k} V_{r-k}^\top x\|_2 , \tag{13}$$

with probability at least $1 - c_0 e^{-c_1 k}$ for some universal constants $c_0, c_1 > 0$.

Random right singular matrix V

We focus on controlling the remaining error terms $\mathrm{rad}(V_k^\top \mathcal{C})$ and $\sup_{x \in \mathcal{C}} \|D_{r-k} V_{r-k}^\top x\|_2$, which appear in the right-hand side of (13). These two terms depend on the coupling between the matrix A – through its right singular matrix V – and the constrained set $\mathcal{C}$. In a worst-case sense, the former term is of order $\mathrm{rad}(\mathcal{C})$, and the latter term is of order $\sigma_{k+1}\mathrm{rad}(\mathcal{C})$, which yields an upper bound scaling as $\sigma_{k+1}\mathrm{rad}(\mathcal{C})$. On the other hand, under reasonable assumptions, these error terms can be significantly reduced, up to a quantity which involves two essential dimensions of the least-squares problem, that is, the effective dimension of the matrix A and the Gaussian width of the set $\mathcal{C}$.

We make the assumption that the matrix $V \in \mathbb{R}^{d \times r}$ is a projection onto a uniformly random m-dimensional subspace in $\mathbb{R}^n$. As an immediate consequence (see, for instance, Theorem 7.7.1 in [5]), it holds that

$$\mathrm{rad}\big(V_k^\top \mathcal{C}\big) \lesssim \frac{w(\mathcal{C}) + \mathrm{rad}(\mathcal{C})\sqrt{k}}{\sqrt{n}} , \tag{14}$$

with probability at least $1 - 2e^{-m}$.

It remains to control the term $\sup_{x \in \mathcal{C}} \|D_{r-k} V_{r-k}^\top x\|_2$. Again, we will use a Gaussian Chevet inequality. In order to do so, we transform the latter quantity into a standard Gaussian form. Let us introduce an orthogonal matrix $\Omega \sim \mathrm{Unif}(\mathcal{O}_{r-k})$. Then, we have $\Omega V_{r-k}^\top \stackrel{\mathrm{d}}{=} V_{r-k}$, and further,

$$\mathbb{E}\left[\sup_{x \in \mathcal{C}} \|D_{r-k} V_{r-k}^\top x\|_2\right] = \mathbb{E}\left[\sup_{x \in \mathcal{C}} \|D_{r-k} \Omega V_{r-k}^\top x\|_2\right] . \tag{15}$$

Let $G \in \mathbb{R}^{(r-k) \times d}$ be a matrix with i.i.d. entries $\mathcal{N}(0, \frac{1}{d})$, independent of Ω and V_{r-k}, and let $\Lambda_n = \mathrm{diag}(\eta_1, \dots, \eta_{r-k})$ be its matrix of singular values. For $i = 1, \dots, r-k$, we set $\Lambda_n^i = \mathrm{diag}(\eta_{1+i}, \dots, \eta_{r-k+i})$ the diagonal matrix which entries are equal to that of Λ_n, shifted by i units. It holds that $\Lambda_n^i \stackrel{\mathrm{d}}{=} \Lambda_n$. By independence of Ω, Λ_n and V_{r-k}, it follows that $G \stackrel{\mathrm{d}}{=} \Omega \Lambda_n^i V_{r-k}^\top$. Writing $\overline{\Lambda}_n = \frac{1}{r-k}\sum_{i=1}^{r-k} \Lambda_n^i$, we have that

$\mathbb{E}\,\overline{\Lambda}_n = \gamma\, I_{r-k}$, where $\gamma = \frac{1}{r-k}\mathbb{E}\left[\sum_{i=1}^{r-k} \eta_i\right]$. Hence,

$$\begin{aligned}
\sup_{x\in\mathcal{C}} \|D_{r-k}\Omega V_{r-k}^\top x\|_2 &= \frac{1}{\gamma}\sup_{x\in\mathcal{C}} \|D_{r-k}\Omega\,\gamma\, I_{r-k} V_{r-k}^\top x\|_2 &(16)\\
&= \frac{1}{\gamma}\sup_{x\in\mathcal{C}} \left\|D_{r-k}\Omega\,\mathbb{E}\left[\frac{1}{r-k}\sum_{i=1}^{r-k}\Lambda_n^i\right] V_{r-k}^\top x\right\|_2 &(17)\\
&\underset{(i)}{=} \frac{1}{\gamma}\sup_{x\in\mathcal{C}} \left\|D_{r-k}\Omega\,\mathbb{E}\left[\frac{1}{r-k}\sum_{i=1}^{r-k}\Lambda_n^i \mid \Omega, V_{r-k}\right] V_{r-k}^\top x\right\|_2 &(18)\\
&\underset{(ii)}{\leqslant} \frac{1}{\gamma}\sup_{x\in\mathcal{C}} \mathbb{E}\left[\left\|D_{r-k}\Omega\,\frac{1}{r-k}\sum_{i=1}^{r-k}\Lambda_n^i V_{r-k}^\top x\right\|_2 \mid \Omega, V_{r-k}\right] &(19)\\
&\underset{(iii)}{\leqslant} \frac{1}{\gamma}\sup_{x\in\mathcal{C}} \frac{1}{r-k}\sum_{i=1}^{r-k}\mathbb{E}\left[\left\|D_{r-k}\Omega\,\Lambda_n^i V_{r-k}^\top x\right\|_2 \mid \Omega, V_{r-k}\right] &(20)\\
&\underset{(iv)}{\leqslant} \frac{1}{\gamma}\frac{1}{r-k}\sum_{i=1}^{r-k}\mathbb{E}\left[\sup_{x\in\mathcal{C}}\left\|D_{r-k}\Omega\,\Lambda_n^i V_{r-k}^\top x\right\|_2 \mid \Omega, V_{r-k}\right]. &(21)
\end{aligned}$$

Equality (i) holds by independence of the Λ_n^i with Ω and V_{r-k}, (ii) and (iii) both follow from Jensen's inequality, and (iv) follows from two successive applications of Fatou's lemma. Then, taking expectations with respect to Ω and V_{r-k}, using the fact that $\Omega\Lambda_n^i V_{r-k}^\top \overset{\mathrm{d}}{=} G$ and recalling equality (15), we get

$$\sup_{x\in\mathcal{C}} \|D_{r-k}V_{r-k}^\top x\|_2 \leqslant \frac{1}{\gamma}\mathbb{E}\,\left[\left\|D_{r-k}Gx\right\|_2\right]. \tag{22}$$

It is a standard result that the minimal singular value[1] of G is greater than $1-\sqrt{\frac{r-k}{d}}$, which further implies that $\gamma \geqslant 1-\sqrt{\frac{r-k}{d}}$, and

$$\mathbb{E}\,\left[\sup_{x\in\mathcal{C}} \|D_{r-k}V_{r-k}^\top x\|_2\right] \leqslant \frac{1}{1-\sqrt{\frac{r-k}{d}}}\mathbb{E}\,\left[\left\|D_{r-k}Gx\right\|_2\right]. \tag{23}$$

Finally, we can apply Gordon's inequality to obtain the upper bound

$$\mathbb{E}\left[\sup_{x\in\mathcal{C}} \|D_{r-k}V_{r-k}^\top x\|_2\right] \leqslant \frac{1}{1-\sqrt{\frac{r-k}{d}}}\frac{1}{\sqrt{d}}\left(w(\mathcal{C})\,\mathrm{rad}(\mathcal{E}) + w(\mathcal{E})\,\mathrm{rad}(\mathcal{C})\right). \tag{24}$$

where we introduced the ellipsoid $\mathcal{E} = \{D_{r-k}z \mid \|z\|_2 \leqslant 1\}$. Using the fact that $w(\mathcal{E}) \leqslant \left(\sum_{j=k+1}\sigma_j^2\right)^{\frac{1}{2}}$ and $\mathrm{rad}(\mathcal{E}) = \sigma_{k+1}$, the latter inequality becomes

$$\mathbb{E}\left[\sup_{x\in\mathcal{C}} \|D_{r-k}V_{r-k}^\top x\|_2\right] \leqslant \frac{1}{1-\sqrt{\frac{r-k}{d}}}\left(\sigma_{k+1}\frac{w(\mathcal{C})}{\sqrt{d}} + \mathrm{rad}(\mathcal{C})\sqrt{\frac{k}{d}}\sqrt{\frac{1}{k}\sum_{j=k+1}^{r}\sigma_j^2}\right).$$

Combining the previous inequalities and taking a union bound over the several events we considered, we get the following result.

[1] This lower bound could be significantly improved. In particular, for square Gaussian matrices $d\times d$, it is known that, asymptotically, $\gamma = \frac{8}{3\pi}$ [1].

Theorem 2. *Choosing an oversampling parameter $p = k + 1$, it holds that*

$$\sigma_{\mathcal{C}}(A, S) \lesssim \frac{1}{1 - 2\sqrt{\frac{r-k}{d}}} \left(\frac{w(\mathcal{C}) + rad(\mathcal{C})\sqrt{k}}{\sqrt{d}} \right) \left(\sigma_{k+1} + \sqrt{\frac{1}{k} \sum_{j=k+1}^{r} \sigma_j^2} \right), \tag{25}$$

with probability at least $1 - c_1 e^{-c_2 m}$, where c_1 and c_2 are universal constants.

Let us comment on the above result. First, we mention that, in an overparameterized setting where $d \gg r$, the first term $\left(1 - 2\sqrt{\frac{r-k}{d}}\right)^{-1}$ is negligible. Thus, let us focus on the factor term involving $\sqrt{k}$ and the Gaussian width $w(\mathcal{C})$. For instance, if $\mathcal{C}$ is the Euclidean unit ball, then $w(\mathcal{C}) \approx \sqrt{d}$ so that $\frac{w(\mathcal{C})+\sqrt{k}}{\sqrt{d}} \approx 1$, and we recover the standard result

$$\|P_{AS}^{\perp} A\|_2 \lesssim \left(\sigma_{k+1} + \sqrt{\frac{1}{k} \sum_{j=k+1}^{r} \sigma_j^2} \right). \tag{26}$$

On the other hand, if the Gaussian width of the set $\mathcal{C}$ is much smaller than $\sqrt{d}$, then the term $\sigma_{\mathcal{C}}(A, S)$ can be much smaller than $\|P_{AS}^{\perp} A\|_2$. For instance, if $\mathcal{C}$ is the L_1-ball, then $w(\mathcal{C}) \approx \sqrt{\log d}$, and we get

$$\sigma_{\mathcal{C}}(A, S) \lesssim \left(\frac{\sqrt{\log d} + \sqrt{k}}{\sqrt{d}} \right) \left(\sigma_{k+1} + \sqrt{\frac{1}{k} \sum_{j=k+1}^{r} \sigma_j^2} \right). \tag{27}$$

If the effective dimension d_e – which we informally define as as the smallest index such that $\sigma_{d_e} \gtrsim \sqrt{\frac{1}{d_e} \sum_{j=d_e+1}^{r} \sigma_j^2}$ – is smaller than $\log d$, then one should pick a sketch size $k \lesssim \log d$, which results in an error approximation

$$\sigma_{\mathcal{C}}(A, S) \lesssim \sqrt{\frac{\log d}{d}}\, \sigma_{k+1}. \tag{28}$$

More generally, a 'favorable' regime is $d_e \leqslant w(\mathcal{C})$.

References

[1] A. S. Bandeira, C. Kennedy, and A. Singer. Approximating the little grothendieck problem over the orthogonal and unitary groups. *Mathematical programming*, 160(1-2):433–475, 2016.

[2] S. Chevet. Séries de variables aléatoires gaussiennes à valeurs dans E $\hat{\otimes}_\varepsilon$ F. Application aux produits d'espaces de Wiener abstraits. In *Proc. Séminaire sur la Géométrie des Espaces de Banach (1977-1978)*. Ecole Polytechnique, Palaiseau, 1978.

[3] Y. Gordon. On Milman's inequality and random subspaces which escape through a mesh in $\mathbb{R}^n$. In *Geometric Aspects of Functional Analysis*, pages 84–106. Springer, 1988.

[4] N. Halko, P.-G. Martinsson, and J. A. Tropp. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM review*, 53(2):217–288, 2011.

[5] R. Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2018.

[6] R. Witten and E. Candes. Randomized algorithms for low-rank matrix factorizations: sharp performance bounds. *Algorithmica*, 72(1):264–281, 2015.