Beginning of Lecture 38

Let $T \in \mathcal{L}(V, W)$ and recall the singular value decomposition of T:

- $\sigma_1, \ldots, \sigma_n$ the singular values of T with $\sigma_1, \ldots, \sigma_r$ the nonzero singular values.
- $e_1, \ldots, e_n$ an ONB of V consisting of eigenvectors of T^*T with $T^*Te_k = \sigma_k^2 e_k$
- $f_1, \ldots, f_r \in W$ orthonormal and defined as $f_k := (1/\sigma_k)Te_k$.

Then:

$$Tv = \sum_{k=1}^{r} \sigma_k \langle v, e_k \rangle f_k, \quad \forall\, v \in V$$

If $T \in \mathcal{L}(V)$ then $f_1, \ldots, f_r \in V$. In this case:

$$T^m v = \sum_{k=1}^{r} \sigma_k \langle v, e_k \rangle T^{m-1} f_k \neq \sum_{k=1}^{r} \sigma_k^m \langle v, e_k \rangle f_k$$

unlike an ONB $g_1, \ldots, g_n$ of eigenvectors of T in which

$$v = \sum_{k=1}^{n} \langle v, g_k \rangle g_k \Longrightarrow T^m v = \sum_{k=1}^{n} \langle v, g_k \rangle T^m g_k = \sum_{k=1}^{n} \lambda_k^m \langle v, g_k \rangle g_k$$

Nevertheless the SVD is still very useful! Indeed, the SVD tells us a lot about the “metric properties” of a linear transformation.

Computational Remark: The SVD requires finding the eigenvalues and eigenvectors of T^*T. In general computing eigenvalues of an operator (matrix) is hard, but for self-adjoint operators (like T^*T) there are algorithms that can do it very effectively. This will be good to keep in mind, even though we will not say any more on the subject.

Application 1: Image of the unit ball

Let $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and let

$$B = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$$

be the closed unit ball. We want to describe the shape of B after it is transformed by T, i.e., we want to know what $T(B)$ looks like.

Suppose first that $T \in \mathcal{L}(\mathbb{R}^n)$ and T takes a diagonal form, i.e., $\tilde{e}_1, \ldots, \tilde{e}_n$ with

$$\tilde{e}_k = (0, \ldots, 0, \underbrace{1}_{k}, 0, \ldots, 0)$$

are eigenvectors of T with eigenvalues $\sigma_1, \ldots, \sigma_n$, each $\sigma_k > 0$, so that in particular for any $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$,

$$Tx = T(x_1, \ldots, x_n) = (\sigma_1 x_1, \ldots, \sigma_n x_n)$$

Therefore, if $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$, then

$$y = (y_1, \ldots, y_n) = T(x_1, \ldots, x_n) = Tx, \quad \text{for } x \in B$$

if and only if

$$\sum_{k=1}^{n} \frac{y_k^2}{\sigma_k^2} \leq 1 \tag{21}$$

Indeed, $T(x_1, \ldots, x_n) = (\sigma_1 x_1, \ldots, \sigma_n x_n) = (y_1, \ldots, y_n)$ if and only if $y_k = \sigma_k x_k$, or equivalently $x_k = y_k / \sigma_k$. But then:

$$x \in B \Leftrightarrow \|x\| \leq 1 \Leftrightarrow \|x\|^2 \leq 1 \Leftrightarrow \sum_{k=1}^{n} x_k^2 \leq 1 \Leftrightarrow \sum_{k=1}^{n} \frac{y_k^2}{\sigma_k^2} \leq 1$$

The set of points satisfying (21) is called an ellipsoid. In $\mathbb{R}^2$ it is an ellipse (with its interior) with half-axes σ_1 and σ_2 [draw a picture]. The vectors $\tilde{e}_1, \ldots, \tilde{e}_n$ defined above as the standard ONB of $\mathbb{R}^n$ are the principal axes.

Now consider $T \in \mathcal{L}(V)$ with singular values $\sigma_1, \ldots, \sigma_n$ and $\sigma_k > 0$ for each $k = 1, \ldots, n$. Let $e_1, \ldots, e_n \in \mathbb{R}^n$ and $f_1, \ldots, f_n \in \mathbb{R}^n$ be the two ONBs associated to the SVD of T so that

$$Tx = \sum_{k=1}^{n} \sigma_k \langle x, e_k \rangle f_k, \quad \forall\, x \in \mathbb{R}^n$$

Note that:

$$x \in B \Leftrightarrow \|x\| \leq 1 \Leftrightarrow \|x\|^2 \leq 1 \Leftrightarrow \sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \leq 1$$

We also have:

$$\begin{aligned} y = Tx, \text{ for } x \in B &\Leftrightarrow \sum_{k=1}^{n} \langle y, f_k \rangle f_k = \sum_{k=1}^{n} \sigma_k \langle x, e_k \rangle f_k, \text{ for } x \in B \\ &\Leftrightarrow \langle y, f_k \rangle = \sigma_k \langle x, e_k \rangle, \text{ for } x \in B, \ \forall\, k = 1, \ldots, n \\ &\Leftrightarrow \sum_{k=1}^{n} \frac{|\langle y, f_k \rangle|^2}{\sigma_k^2} \leq 1 \end{aligned}$$

But that is also an ellipsoid! It is just rotated so that its principal axes are $f_1, \ldots, f_n$.

Now consider the fully general case of $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ with nonzero singular values $\sigma_1, \ldots, \sigma_r$. We have:

$$Tx = \sum_{k=1}^{r} \sigma_k \langle x, e_k \rangle f_k$$

Notice this implies that $\operatorname{range} T = \operatorname{span}(f_1, \ldots, f_r)$. In particular,

$$y = Tx \text{ for } x \in \mathbb{R}^n \Leftrightarrow y \in \operatorname{range} T \Leftrightarrow y \in \operatorname{span}(f_1, \ldots, f_r)$$

Now we can use essentially the same calculation as before to get:

$$\begin{aligned} y = Tx, \text{ for } x \in B &\Leftrightarrow \sum_{k=1}^{r} \langle y, f_k \rangle f_k = \sum_{k=1}^{r} \sigma_k \langle x, e_k \rangle f_k, \text{ for } x \in B \\ &\Leftrightarrow \langle y, f_k \rangle = \sigma_k \langle x, e_k \rangle, \text{ for } x \in B, \ \forall\, k = 1, \ldots, r \\ &\Leftrightarrow \sum_{k=1}^{r} \frac{|\langle y, f_k \rangle|^2}{\sigma_k^2} \leq 1 \end{aligned}$$

Thus we have shown:

Theorem 37. *The image $T(B)$ of the closed unit ball B is an ellipsoid in* $\operatorname{range} T$ *with half axes $\sigma_1, \ldots, \sigma_r$ along the principal axes $f_1, \ldots, f_r$, where $\sigma_1, \ldots, \sigma_r$ are the nonzero singular values of T and $f_k := (1/\sigma_k) T e_k$ with $e_1, \ldots, e_n$ an ONB of V consisting of eigenvectors of $T^* T$.*

End of Lecture 38