1 26th of September 2018 — F. Poloni

In the previous lecture we introduced some sufficent conditions for matrix orthogonality.

Theorem 1.1. *Let $U \in M(n, \mathbb{R})$ be an orthogonal matrix and let $x \in \mathbb{R}^n$. Then $\|Ux\| = \|x\|$.*

Proof. $\|Ux\| = (Ux)^T \cdot (Ux) \stackrel{(1)}{=} x^T U^T U x = x^T I_n x = x^T x = \|x\|$
where $\stackrel{(1)}{=}$ follows from the definition of transpose of a product. □

Geometrically, an orthogonal matrix represents a symmetry or a rotations and these operations do not alter the sixe of vectors.

Definition 1.1 (Orthonormality)**.** *Let $x, y \in R^n$ we say that x and y are **orthonormal** if $< x, y >= 0$ and $\|x\| = \|y\| = 1$.*

Fact 1.2. *Let us take $U \in M(n, \mathbb{R})$ such that U is orthogonal. Then its columns $U^1, U^2, \ldots, U^n$ are **orthonormal** and the same holds for its rows.*

$$U^{i^T} U^j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

and

$$U_i {U_j}^T = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Fact 1.3. *Let $U, V \in M(n, \mathbb{R})$, such that U and V are orthogonal, then $U \cdot V$ is orthogonal.*

Proof. $(UV)^T \cdot (UV) = V^T U^T U V = V^T I_n V = V^T V = I_n$ □

Definition 1.2 (Eigenvectors and eigenvalues)**.** *Let $A \in M(n, \mathbb{R})$ and let $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$. If $Ax = \lambda x$ we say that x is an **eigenvector** of **eigenvalue** λ.*

Fact 1.4. *Let $A \in T(n, \mathbb{R})$ (real triangular matrix). The eigenvalues of A are the scalars on the diagonal.*

Under some conditions we may obtaine a **diagonal form** of A, namely:

$$A = V \Lambda V^{-1} = \begin{pmatrix} V^1 & V^2 & \cdots & V^n \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$$

where $\forall i = 1, \ldots, n$ v_i are eigenvectors of A of eigenvalue λ_i.

This is due to the fact that $\forall B \in R^n$ s.t. $Vx = b$, $x = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, where the 1 is found at position i and so $AV^i = V\Lambda V^{-1}V^i = V\Lambda e_i = \lambda_i V^i$.

Another way to see the diagonalized form of A is the following:

$$\begin{aligned} A &= V\Lambda V^{-1} = \sum_{i=1}^{n} V^i \lambda_i w_i \\ &= \boxed{V^1} \cdot \boxed{\lambda_1} \cdot \boxed{\quad w_1 \quad} + \boxed{V^2} \cdot \boxed{\lambda_2} \cdot \boxed{\quad w_2 \quad} + \cdots + \boxed{V^n} \cdot \boxed{\lambda_n} \cdot \boxed{\quad w_n \quad} \end{aligned} \tag{1.1}$$

Something on Matlab . . .

Notice that in Matlab the eigenvalues and eigenvectors of a matrix are computed using the command `[V, Lambda] = eig(U)` and this operation has a computational complexity of $O(n^3)$.

Notice that not all matrices $A \in M(n, \mathbb{R})$ allow a diagonal decomposition. It may happen that such a matrix is diagonalizable in $\mathbb{C}$ and its eigenvalues are complex.

Fact 1.5. *If $A \in M(n, \mathbb{R})$ is diagonalizable (aka may be written as $A = V\Lambda V^{-1}$) $A^k x = \sum_{i=1}^{n} \lambda_i^k \alpha_i V^i$, for some $\alpha_i \in \mathbb{R}$.*

Proof. Algebraic view point: Let us write x in the base of R^n made of the linearly independent columns of V:

$$x = V^1\alpha_1 + V^2\alpha_2 + \cdots + V^n\alpha_n$$

for some $\alpha_i \in \mathbb{R}$.

$$\begin{aligned} Ax &= A \cdot (V^1\alpha_1 + V^2\alpha_2 + \cdots + V^n\alpha_n) \\ &= AV_1\alpha_1 + AV_2\alpha_2 + \cdots + AV_n\alpha_n \\ &= \lambda_1 V^1\alpha_1 + \lambda_2 V^2\alpha_2 + \cdots + \lambda_n V^n\alpha_n \\ &= V^1(\lambda_1\alpha_1) + V^2(\lambda_2\alpha_2) + \cdots + V^n(\lambda_n\alpha_n) \end{aligned} \tag{1.2}$$

Then

$$
\begin{aligned}
A^2x &= A \cdot \Big(V^1(\lambda_1\alpha_1) + V^2(\lambda_2\alpha_2) + \cdots + V^n(\lambda_n\alpha_n)\Big) \\
&= AV^1\lambda_1\alpha_1 + AV^2\lambda_2\alpha_2 + \cdots + AV^n\lambda_n\alpha_n \\
&= {\lambda_1}^2V^1\alpha_1 + {\lambda_2}^2V^2\alpha_2 + \cdots + A{\lambda_n}^2V^n\alpha_n
\end{aligned} \tag{1.3}
$$

Inductively, we have proved the theorem.

Linear algebra view point:

$$
\begin{aligned}
A^kx &= A \cdot A \cdot \ldots \cdot A \cdot x \\
&= V\Lambda \cancel{V^{-1}}\cancel{V}\Lambda\cancel{V^{-1}} \ldots \cancel{V}\Lambda V^{-1}x \\
&= V\Lambda^k V^{-1}x \\
&= V \begin{pmatrix} {\lambda_1}^k & & & \\ & {\lambda_2}^k & & \\ & & \ddots & \\ & & & {\lambda_n}^k \end{pmatrix} V^{-1}x \\
&= V \begin{pmatrix} {\lambda_1}^k & & & \\ & {\lambda_2}^k & & \\ & & \ddots & \\ & & & {\lambda_n}^k \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ddots \\ \alpha_n \end{pmatrix}
\end{aligned} \tag{1.4}
$$

□

Thanks to this proposition we can prove the following

Theorem 1.6. *Let $A \in M(n, \mathbb{R})$. If $\forall \lambda_i$ eigenvalues of A $|\lambda_i| < 1$ then $\lim_{k\to\infty} A^kx = 0$.*

Theorem 1.7. *Let $A \in M(n, \mathbb{R})$. If $\forall \lambda_i$ eigenvalues of A $|\lambda_i| < |\lambda_1|$ then $A^kx \approx V^1{\lambda_1}^k\alpha_1$.*

Fact 1.8. *Let $A \in M(n, \mathbb{R})$ be a diagonalizable matrix and let*

$$
A = V\Lambda V^{-1} = \begin{pmatrix} V^1 & V^2 & \cdots & V^n \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}
$$

Let us now consider a reordering of V's columns and apply the same permutations to the "diagonal vector" of Λ such that

$$
\hat{V} = \begin{pmatrix} V^2 & V^1 & V^3 \cdots & V^n \end{pmatrix} \text{ and } \hat{\Lambda} = \begin{pmatrix} \lambda_2 & & & & \\ & \lambda_1 & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ & & & & \lambda_n \end{pmatrix}
$$

A can be diagonalized through such $\hat{V}$ and $\hat{\Lambda}$: $A = V\Lambda V^{-1} = \hat{V}\hat{\Lambda}\hat{V}^{-1}$.

Moreover, in the case of repeated eigenvalues

Fact 1.9. *Let $A \in M(n, \mathbb{R})$ a diagonalizable matrix such that $A = V\Lambda V^{-1}$, where $\lambda_1 = \lambda_2$ (without loss of generality). Then V can be replaced by* $\widetilde{V} = \begin{pmatrix} \\ V^1 + V^2 & V^1 - V^2 & V^3 & \cdots & V^n \\ \\ \end{pmatrix}$.

Theorem 1.10 (Spectral theorem)**.** *Let $A \in S(n, \mathbb{R})$ (A is a real symmetric matrix). Then A is diagonalizable, its eigenvalues are real numbers and V is an orthogonal matrix.*

Fact 1.11. *Let $A \in S(n, \mathbb{R})$ and let $x \in \mathbb{R}^n$. $\lambda_{min}\|x\|^2 \leq x^T Ax \leq \lambda_{max}\|x\|^2$, where λ_{max} and λ_{min} are respectively the eigenvalue of maximum value and the eigenvalue fo minimum value.*

Proof. Diagonal A: $x^T Ax = x^T \cdot \begin{pmatrix} \lambda_2 & & & & \\ & \lambda_1 & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ & & & & \lambda_n \end{pmatrix} \cdot x = \lambda_1 {x_1}^2 + \lambda_2 {x_2}^2 + \cdots + \lambda_n {x_n}^2$

It is obvious that this sum is bounded by:

$$\lambda_{\min} \cdot ({x_1}^2 + {x_2}^2 + \cdots + {x_n}^2) \leq \lambda_1 {x_1}^2 + \lambda_2 {x_2}^2 + \cdots + \lambda_n {x_n}^2 \leq \lambda_{\max} \cdot ({x_1}^2 + {x_2}^2 + \cdots + {x_n}^2)$$

The following holds: $\lambda_{\min} \cdot ({x_1}^2 + {x_2}^2 + \cdots + {x_n}^2) = \lambda_{\min} \cdot x^T x = \lambda_{\min} \cdot \|x\|^2$ and, on the other hand, $\lambda_{\max} \cdot ({x_1}^2 + {x_2}^2 + \cdots + {x_n}^2) = \lambda_{\max} \cdot x^T x = \lambda_{\max} \cdot \|x\|^2$ and this proves the fact in the special case of diagonal matrix A.

General case: Let us represent A through its eigendecomposition: $A = U\Lambda U^{-1} = U\Lambda U^T$, where U is an orthogonal matrix.

$x^T Ax = x^T U\Lambda U^T x \overset{(1)}{=} y^T \Lambda y$

where $\overset{(1)}{=}$ is due to the change of variable $y = U^T x$ (that implies $y^T = x^T U$).

By the same argument used in the diagonal case,

$\lambda_{\min} \cdot \|y\|^2 \leq y^T \Lambda y \leq \lambda_{\max} \cdot \|y\|^2$.

Now the point is that if we can replace $\|y\|^2$ with $\|x\|^2$ we have proved the theorem. In fact this is true, due to the orthogonality of matrix U.

□

Corollary 1.12. *Let $A \in S(n, \mathbb{R})$ and let $x \in \mathbb{R}^n$. If $x \neq 0$, $\lambda_{min} \leq \frac{x^T Ax}{\|x\|^2} \leq \lambda_{max}$, where λ_{max} and λ_{min} are respectively the eigenvalue of maximum value and the eigenvalue fo minimum value.*

Definition 1.3 (Positive (semi)definite)**.** *Let $A \in S(n, \mathbb{R})$. We say that A is* ***positive semidefinite*** *if $\forall x \in R^n$ $x^T Ax \geq 0$.*

On the other hand,A is termed ***positive definite*** *if $x^T Ax > 0 \forall x \in R^n \backslash \{0\}$.*

Fact 1.13. *Let $A \in S(n, \mathbb{R})$. $\forall \lambda$ eigenvalue of A $\lambda \geq 0$ iff A is* ***positive semidefinite****.*

On the other hand, all eigenvalues are ***strictly*** *positive iff A is positive definite.*

Proof.

($\Rightarrow$) $x^T Ax \geq \lambda_{\min}$.

($\Leftarrow$)Let v_i be an eigenvector of A.

$0 \leq {v_i}^T A v_i = {v_i}^T \cdot (\lambda_i v_i) = \lambda_i {v_i}^T v_i = \lambda_i \|v_i\|^2 \Rightarrow \lambda_i \geq 0$. □

Fact 1.14. *Let $B \in M(m, n, \mathbb{R})$, $B^T B \in S(n, \mathbb{R})$ is positive semidefinite.*

Proof. Symmetry: $(B^T B)^T = B^T \cdot (B^T)^T = B^T B$.

Positive definite: $x^T B^T Bx = (Bx)^T (Bx) = \|Bx\|^2 \geq 0$

□

Corollary 1.15. *The same holds for BB^T, since we can define $C = B^T$.*

Fact 1.16. *Let $A \in S(n, \mathbb{R})$. $A \succeq 0$ and A invertible iff A is* ***strictly positive definite****.*

Something on Matlab ...

In ordert to check if a matrix A is positive definite in Matlab we can look at its eigenvalues (cfr. `eig(A)`).