30. Lecture 30: 8.3 Polar Decomposition and Singular Value Decomposition

In the book the singular value decomposition is proven in detail and the polar decomposition is only outlined in the exercises. One can prove one from the other in any order. Here we first prove the polar decomposition since it is more natural though the proof is more abstract.

A complex number can be written in polar form $z = r\,e^{i\theta}$. Similarly we can write:
Th (Polar Decomposition) If A is an $n\times n$ matrix then there is an orthogonal $n\times n$ matrix Q and a symmetric positive definite $n\times n$ matrix P such that $A = QP$.
Pf The are several steps to the proof:
For a complex number we first define $r^2 = |z|^2 = \overline{z}z$ and then we define $e^{i\theta} = z/|z|$.

Lemma There is a unique positive definite symmetric matrix P such that $P^2 = A^TA$.
We have $\|P\mathbf{v}\| = \|A\mathbf{v}\|$, for all $\mathbf{v}$.
Pf Since the matrix A^TA is symmetric it can be diagonalized $A^TA = VDV^T$, where V is orthogonal and $D = \operatorname{diag}\{\lambda_1, \dots, \lambda_n\}$. Here $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ are the eigenvalues of the positive definite matrix A^TA. Hence if we define $\sigma_i = \sqrt{\lambda_i}$, for $i = 1, \dots, n$ and $\Sigma = \operatorname{diag}\{\sigma_1, \dots, \sigma_n\}$ then $\Sigma^2 = D$ so $P = V\Sigma V^T$ satisfy $P^2 = V\Sigma V^T V\Sigma V^T = V\Sigma^2 V^T = A^TA$.
We have $\|A\mathbf{v}\|^2 = \langle A\mathbf{v}, A\mathbf{v}\rangle = \langle A^TA\mathbf{v}, \mathbf{v}\rangle = \langle P^2\mathbf{v}, \mathbf{v}\rangle = \langle P\mathbf{v}, P\mathbf{v}\rangle = \|P\mathbf{v}\|^2$.

Given P as in the lemma we show that there is a matrix Q such that $QP = A$ and $Q^TQ = I$.
If P is invertible we define $Q = AP^{-1}$ so $Q^TQ = (P^{-1})^TA^TAP^{-1} = (P^{-1})^TP^2PP^{-1} = I$.
In general we want to define a linear map $Q_1 : \operatorname{Im} P \to \operatorname{Im} A$ by

$$Q_1(P\mathbf{v}) = A\mathbf{v}, \qquad \text{for all } \mathbf{v} \in \mathbb{R}^n.$$

We need to show that Q_1 is well defined, i.e. if $P\mathbf{v}_1 = P\mathbf{v}_2$ then $A\mathbf{v}_1 = A\mathbf{v}_2$. By the lemma

$$\|A\mathbf{v}_1 - A\mathbf{v}_2\| = \|A(\mathbf{v}_1 - \mathbf{v}_2)\| = \|P(\mathbf{v}_1 - \mathbf{v}_2)\| = \|P\mathbf{v}_1 - P\mathbf{v}_2\|.$$

It follows that Q_1 is well defined and by the lemma

$$\|Q_1\mathbf{u}\| = \|\mathbf{u}\|, \qquad \text{for} \quad \mathbf{u} \in \operatorname{Im} P.$$

In particular Q_1 is invertible and by the fundamental theorem of linear algebra

$$\dim \operatorname{Im} P = \dim \operatorname{Im} A.$$

This implies that

$$\dim(\operatorname{Im} P)^\perp = \dim(\operatorname{Im} A)^\perp.$$

We can hence pick an orthonormal basis $\mathbf{e}_1, \dots, \mathbf{e}_m$ for $(\operatorname{Im} P)^\perp$ and an orthonormal basis $\mathbf{f}_1, \dots, \mathbf{f}_m$ for $(\operatorname{Im} A)^\perp$ and define a map by

$$Q_2\big(a_1\mathbf{e}_1 + \cdots + a_m\mathbf{e}_m\big) = a_1\mathbf{f}_1 + \cdots + a_m\mathbf{f}_m.$$

Then $\|Q_2\mathbf{w}\| = \|\mathbf{w}\|$, for $\mathbf{w} \in (\operatorname{Im} P)^\perp$. In general we can uniquely decompose any $\mathbf{x} \in \mathbb{R}^n$:

$$\mathbf{x} = \mathbf{u} + \mathbf{w}, \qquad \text{where} \quad \mathbf{u} \in \operatorname{Im} P, \quad \mathbf{w} \in (\operatorname{Im} P)^\perp,$$

and define

$$Q\mathbf{x} = Q_1\mathbf{u} + Q_2\mathbf{w},$$

For each $\mathbf{v}$ we have

$$Q(P\mathbf{v}) = Q_1(P\mathbf{v}),$$

since $P\mathbf{v} \in \operatorname{Im} P$ and hence has bo component in $(\operatorname{Im} P)^\perp$. By Pythagorean theorem

$$\|Q\mathbf{v}\|^2 = \|Q_1\mathbf{u} + Q_2\mathbf{w}\|^2 = \|Q_1\mathbf{u}\|^2 + \|Q_2\mathbf{w}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{w}\|^2 = \|\mathbf{v}\|^2.$$

Th (Singular Value Decomposition) Any $m \times n$ matrix A can be factored into

$$A = U\Sigma V^T,$$

where U is an orthogonal $m\times m$, V is an orthogonal $n\times n$ and Σ is a diagonal $m\times n$ matrix. The columns of U are eigenvectors of AA^T and the columns of V are eigenvectors of A^TA. The r singular values on the diagonal of Σ are the square roots of the nonzero eigenvalues of AA^T and A^TA.

Remarks For positive definite symmetric $n\times n$ matrices, the decomposition reduces to QDQ^T. In general for $n\times n$ matrices it follows from the polar decomposition and the spectral theorem. On the other hand the polar decomposition follows from the singular value decomposition.

Proof Let $r \leq \min(m,n)$ be the rank of A which is also the rank of A^TA and of AA^T. Since A^TA is symmetric it has a complete set of orthonormal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Number them so the first r corresponds to the nonzero eigenvalues $\sigma_1^2, \ldots, \sigma_r^2$. Set $V = \begin{bmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{bmatrix}$.

Note that $AA^T(A\mathbf{v}_i) = A(A^TA\mathbf{v}_i) = \sigma_i^2 A\mathbf{v}_i$, for $i=1,\ldots,r$. For $i=1,\ldots,r$ let $\mathbf{u}_i$ be equal to $A\mathbf{v}_i$ normalized so $\|\mathbf{u}_i\| = 1$, by the lemma $A\mathbf{v}_i = \sigma_i\mathbf{u}_i$. For $i=r+1,\ldots,m$ let $\mathbf{u}_i$, be other eigenvectors of AA^T so that $\mathbf{u}_1, \ldots, \mathbf{u}_m$ form an orthonormal basis and set $U = \begin{bmatrix} | & & | \\ \mathbf{u}_1 & \cdots & \mathbf{u}_m \\ | & & | \end{bmatrix}$.

Let Σ be the $m\times n$ matrix with $\sigma_1, \ldots, \sigma_r$ in the diagonal and zero elsewhere, i.e. $\Sigma = (\sigma_{ij})$ where $\sigma_{ii} = \sigma_i$, for $i=1,\ldots,r$ and $\sigma_{ij}=0$ if $i\neq j$ and $\sigma_{ii}=0$, if $i>r$; $\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & \\ 0 & \ddots & \ddots & \\ \vdots & \ddots & \sigma_r & \\ & & & \end{bmatrix}$.

It remain to prove that $A = U\Sigma V^T$ or $AV = U\Sigma$.
In fact $A\mathbf{v}_i = \sigma_i\mathbf{u}_i$, for $i = 1, \ldots, r$, and $A\mathbf{v}_i = \mathbf{0}$, for $i = r+1, \ldots, n$, so

$$A \begin{bmatrix} | & & | & | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_r & \mathbf{v}_{r+1} & \cdots & \mathbf{v}_n \\ | & & | & | & & | \end{bmatrix} = \begin{bmatrix} | & & | & | & & | \\ \sigma_1\mathbf{u}_1 & \cdots & \sigma_r\mathbf{u}_r & \mathbf{0} & \cdots & \mathbf{0} \\ | & & | & | & & | \end{bmatrix} = \begin{bmatrix} | & & | & | & & | \\ \mathbf{u}_1 & \cdots & \mathbf{u}_r & \mathbf{u}_{r+1} & \cdots & \mathbf{u}_m \\ | & & | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & \\ 0 & \ddots & \ddots & \\ \vdots & \ddots & \sigma_r & \\ & & & \end{bmatrix}.$$

Applications of SVD

One can approximate large matrices by smaller rank ones by setting small singular values 0. One can construct a **pseudoinverse** $A^+ = V\Sigma^+U^T$, where Σ^+ is obtained from Σ by inverting the nonzero singular values in the diagonal.

Ex Find a singular value decomposition for $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \end{bmatrix}$.

Sol The matrix $A^TA = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 2 & -2 \\ -2 & -2 & 2 \end{bmatrix}$ has rank 1 so it must have a two dimensional nullspace so 0 is a double eigenvalue. It the easy to find that the eigenvalues are $\lambda_1 = 6$, $\lambda_2 = \lambda_3 = 0$. The corresponding orthonormal eigenvectors are $\mathbf{v}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{v}_3 = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$. The last two are found pick two vectors satisfying $x+y-2z = 0$ and applying Gram-Schmidt to get an orthonormal basis. Let $\mathbf{u}_1 = A\mathbf{v}_1/\sqrt{6} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and let $\mathbf{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ be a vector orthogonal to $\mathbf{u}_1$. Set $U = \begin{bmatrix} | & | \\ \mathbf{u}_1 & \mathbf{u}_2 \\ | & | \end{bmatrix}$, $\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. and $V = \begin{bmatrix} | & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ | & | & | \end{bmatrix}$.

Summary

A complex number can be written in polar form $z = r\,e^{i\theta}$. Similarly we can write:
Th (Polar Decomposition) If A is an $n{\times}n$ matrix then there is an orthogonal $n{\times}n$ matrix Q and a symmetric positive definite $n{\times}n$ matrix P such that $A = QP$.
Lemma There is a unique positive definite symmetric matrix P such that $P^2 = A^TA$.
We have $\|P\mathbf{v}\| = \|A\mathbf{v}\|$, for all $\mathbf{v}$.
Pf Since the matrix A^TA is symmetric it can be diagonalized $A^TA = VDV^T$, where V is orthogonal and $D = \mathrm{diag}\{\lambda_1, \dots, \lambda_n\}$. Here $\lambda_1 \ge \cdots \ge \lambda_n \ge 0$ are the eigenvalues of the positive definite matrix A^TA. Hence if we define $\sigma_i = \sqrt{\lambda_i}$, for $i = 1, \dots, n$ and $\Sigma = \mathrm{diag}\{\sigma_1, \dots, \sigma_n\}$ then $\Sigma^2 = D$ so $P = V\Sigma V^T$ satisfy $P^2 = V\Sigma V^T V\Sigma V^T = V\Sigma^2 V^T = A^TA$.
We have $\|A\mathbf{v}\|^2 = \langle A\mathbf{v}, A\mathbf{v}\rangle = \langle A^TA\mathbf{v}, \mathbf{v}\rangle = \langle P^2\mathbf{v}, \mathbf{v}\rangle = \langle P\mathbf{v}, P\mathbf{v}\rangle = \|P\mathbf{v}\|^2$.
Pf For the proof of the polar decomposition we define a linear map $Q_1 : \mathrm{Im}\,P \to \mathrm{Im}\,A$ by

$$Q_1(P\mathbf{v}) = A\mathbf{v}, \qquad \text{for all } \mathbf{v} \in \mathbb{R}^n.$$

so by the lemma $\|Q_1\mathbf{u}\| = \|\mathbf{u}\|$, for $\mathbf{u} \in \mathrm{Im}\,P$, and a linear map $Q_2 : (\mathrm{Im}\,P)^\perp \to (\mathrm{Im}\,A)^\perp$ so $\|Q_2\mathbf{w}\| = \|\mathbf{w}\|$, for $\mathbf{w} \in (\mathrm{Im}\,P)^\perp$. In general we can uniquely write any $\mathbf{x} = \mathbf{u} + \mathbf{w}$, where $\mathbf{u} \in \mathrm{Im}\,P$ and $\mathbf{w} \in (\mathrm{Im}\,P)^\perp$, and we set $Q(\mathbf{x}) = Q_1\mathbf{u} + Q_2\mathbf{w}$. Then $Q(P\mathbf{v}) = Q_1(P\mathbf{v}) = AP\mathbf{v}$.

Th (Singular Value Decomposition) Any $m \times n$ matrix A can be factored into

$$A = U\Sigma V^T,$$

where U is an orthogonal $m{\times}m$, V is an orthogonal $n{\times}n$ and Σ is a diagonal $m{\times}n$ matrix. The columns of U are eigenvectors of AA^T and the columns of V are eigenvectors of A^TA. Let r be the rank of A. The first r entries on the diagonal of Σ are the square roots of the nonzero eigenvalues of AA^T and A^TA, called the singular values of A.
Remarks For positive definite symmetric $n{\times}n$ matrices, the decomposition reduces to QDQ^T. In general for $n{\times}n$ matrices it follows from the polar decomposition and the spectral theorem. On the other hand the polar decomposition follows from the singular value decomposition.
Ex Find a singular value decomposition for $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \end{bmatrix}$.
Sol The matrix $A^TA = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 2 & -2 \\ -2 & -2 & 2 \end{bmatrix}$ has rank 1 so it must have a two dimensional nullspace so 0 is a double eigenvalue. It the easy to find that the eigenvalues are $\lambda_1 = 6$, $\lambda_2 = \lambda_3 = 0$. The corresponding orthonormal eigenvectors are $\mathbf{v}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{v}_3 = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$.
The last two are found pick two vectors satisfying $x + y - 2z = 0$ and applying Gram-Schmidt to get an orthonormal basis. Let $\mathbf{u}_1 = A\mathbf{v}_1/\sqrt{6} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and let $\mathbf{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ be a vector orthogonal to $\mathbf{u}_1$. Set $U = \begin{bmatrix} | & | \\ \mathbf{u}_1 & \mathbf{u}_2 \\ | & | \end{bmatrix}$, $\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. and $V = \begin{bmatrix} | & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ | & | & | \end{bmatrix}$.