Needham, VDGF, Chapter 20, Exercise 12, pp. 221-222.
Palmer, June 17, 2023, 7 am PT

12. Visual Linear Algebra*. Do all the following by reasoning directly and geometrically about the linear transformations themselves, not by applying the "Devil's machine"....*

(i) Verify that the geometric interpretation [15.4] [p. 155] *of the SVD generalizes to* $\mathbb{R}^3$*, and therefore the interpretation [15.5]* [p. 157] *of* M^T *does too.*
Note: Although visualization becomes harder, the same is true in $\mathbb{R}^n$.

One can imagine various ways in which there might be a generalization of SVD to $\mathbb{R}^3$. The following seems the most general:

Instead of an origin-centred circle, we construct an origin-centred sphere on three orthogonal axes. Preservation of midpoints enables one to prove that M, now a three dimensional transformation corresponding to [M] with three columns and three rows, is equivalent to three perpendicular stretches (by the "singular values" σ_1, σ_2, and σ_3), followed by a twist about each of the three axes: $\tau_i = \varphi_i - \theta_i$, where i,j stands for X, Y, or Z. One can think of this as combining three versions of [15.4] [p. 155], one each for the xy, xz, and yz planes. M^T is still interpreted with the same stretch with the rotations applied in reverse direction.

Then one could write

$$M = R_{\varphi_X} \circ R_{\varphi_Y} \circ R_{\varphi_Z} \circ \Sigma \circ R_{-\theta_Z} \circ R_{-\theta_Y} \circ R_{-\theta_X}$$

and

$$M^T = R_{\theta_Z} \circ R_{\theta_Y} \circ R_{\theta_X} \circ \Sigma \circ R_{-\varphi_X} \circ R_{-\varphi_Y} \circ R_{-\varphi_Z}$$

(ii) Recall that an ***orthogonal*** *linear transform is one that preserves lengths, and therefore angles (in general), and orthogonality (in particular)—examples include rotations and reflections. In* $\mathbb{R}^3$*, use (i) to explain why any orthogonal transformation R has the property that* $R^{-1} = R^T$.

In part (i) and (15.12), one sees that twists effected by M^T are reversed by sign reversals on θ in the orthogonal transformation R_θ. We learn, too, on p. 156, that twists effected by M via the orthogonal transformations are undone by the inverse transformation M^{-1}. Hence, reversal of a rotation or reflection R can be written R^{-1}. Since both the transpose R^T and the inverse R^{-1} reverse the direction of R transformation as shown in [15.5] [p. 157], R has the property that $R^{-1} = R^T$.

(iii) Recall that a matrix M is called ***skew-symmetric*** *if* $[M]^T = -[M]$*. In* $\mathbb{R}^2$ *use (15.12)* [p. 156] *to give a geometric characterization (in terms of twist, τ) of the underlying skew-symmetric linear transformation M.*

$[M]^T = -[M] \Longrightarrow -[M] = [R_\theta][\Sigma][R_{-\varphi}]$ (p. 156)

$M = -R_\theta \Sigma R_{-\varphi} = R_{\theta \pm \pi} \Sigma R_{-\varphi}$

So for skew-symmetric M, $\tau = \theta \pm \pi - \varphi$. The transpose rotates M by $\pm\pi$.

(iv) A symmetric linear transformation P is called ***positive definite*** *if* $\mathbf{x} \cdot (P\mathbf{x}) > 0$ *for all* $\mathbf{x}$*. Use (15.13)* [p. 156] *to show that P is positive definite if (and only if) all of its eigenvalues are positive.*

If $\mathbf{x}$ is an eigenvector of P, then $P\mathbf{x} \equiv \lambda\mathbf{x}$. Let λ be any eigenvalue of P. Then $\mathbf{x} \cdot (\lambda \mathbf{x}) > 0 \Longrightarrow \lambda(\mathbf{x} \cdot \mathbf{x}) > 0$. Since $\mathbf{x} \cdot \mathbf{x} > 0$ for all $\mathbf{x}$, P is positive definite iff all of its eigenvalues are positive.

(v) Arguing directly, or alternatively, building on (iv), show that a symmetric, positive, definite linear transformation must have a positive determinant (and hence be invertible).

From (iv) one can conclude that the effect of positive definite P is to stretch $\mathbf{x}$ *in the same direction*. If P is symmetric and positive, it can be diagonalized:

$$P = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix},$$

The stretch in the same direction is only possible if both σ_1 and σ_2 are positive, in which case the determinant $\sigma_1 \sigma_2$ is also positive, which supports (v). We can make the same argument using the linear algebra of (iv):

$$\mathbf{x} \cdot P\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} \sigma_1 x \\ \sigma_2 y \end{pmatrix} = \sigma_1 x^2 + \sigma_2 y^2 > 0, \text{ where } \sigma_1 > 0 \ \&\& \ \sigma_2 > 0$$

This shows that $\mathbf{x} \cdot P\mathbf{x} > 0$ is generally true iff both σ_1 and σ_2 are positive, in which case the determinant $\sigma_1 \sigma_2$ is also positive and P is invertible. The reasoning is the same for a three column vector. [Note: In Poole, David. *Linear Algebra*. Brookes/Cole, 2003, p. 410, the quadratic form is written $\mathbf{x}^T A\mathbf{x}$, where $\mathbf{x}$ is a column vector.]

Show that the converse is false, by finding a simple counterexample in $\mathbb{R}^3$*. (Hint: Geometrically, the determinant is the (signed) volume-expansion factor of the linear transformation, the sign being positive (+) or negative (-) according to whether orientation is preserved or reversed, respectively.)*

Converse: A linear transformation with a positive determinant (hence invertible) must be symmetric and positive-definite.

All that is needed to prove the falsity of the converse is an asymmetric matrix with a positive determinant.

Let $[M] = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ -2 & 1 & 1 \end{pmatrix}$ with determinant 3. $[M]$ has a positive determinant, but it is asymmetric, so the converse of (v) is false. It has complex eigenvalues $\left\{1 + i\sqrt{2}, 1 - i\sqrt{2}, 1\right\}$.

(vi) Given (iv) above, it follows from (15.15) that $M^T M$ and MM^T are always symmetric and positive-definite. Conversely, show that any symmetric, positive-definite linear transformation P can be factorized as $M^T M$, in infinitely many ways.

Let $M = R_\varphi \Sigma \mathrm{R}_{-\theta}$. Then

$$M^T M = \left(R_\theta \Sigma R_{-\varphi}\right)\left(R_\varphi \Sigma \mathrm{R}_{-\theta}\right) = \Sigma^2$$

$$MM^T = \left(R_\varphi \Sigma R_{-\theta}\right)\left(R_\theta \Sigma R_{-\varphi}\right) = \Sigma^2$$

The equations show that $M^T M$ does not depend on the values of θ and φ. Since θ and φ are arbitrary real numbers, the number of combinations $\{\theta, \varphi\}$ is infinite (where θ and φ either bounded by 0 and 2π or unbounded), so the number of factorizations of $M^T M$ is infinite.

(vii) Building on (vi), show that just as any positive number has a real square root, so any positive-definite linear transformation P has a square root q, such that $P = Q^2$. In greater detail, show that $Q = R^{-1} DR$, where R is orthogonal, and [D] is the diagonal matrix whose entries are the square roots of the eigenvalues of P.

Since P can be factored into $M^T M$, where M is a positive-definite linear transformation such that $M^T = M$, it follows that $P = M^T M = M M = M^2$, and P has a square root q = Q = M. Since P is positive definite, we can diagonalize it

$$[P] = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}, \quad [Q] = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}$$

Where σ_1^2 and σ_2^2 are the eigenvalues of [P].

Under zero twist where R is orthogonal, $[Q] = [R]^{-1}[D][R] \longrightarrow [Q] = [D]$. Hence, $Q = D$. For orthogonal matrices, we use

$$R = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \quad \text{and} \quad R^{-1} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}$$

$$[R]^{-1}[D][R] = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}\begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}\begin{pmatrix} c & -s \\ s & c \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}\begin{pmatrix} c\,\sigma_1 & -s\,\sigma_1 \\ s\,\sigma_2 & c\,\sigma_2 \end{pmatrix} = \begin{pmatrix} c^2\,\sigma_1 + s^2\,\sigma_2 & -c\,s\,\sigma_1 + c\,s\,\sigma_2 \\ -c\,s\,\sigma_1 + c\,s\,\sigma_2 & s^2\,\sigma_1 + c^2\,\sigma_2 \end{pmatrix}$$

When $\theta = \pi/2$, this reduces to $\begin{pmatrix} \sigma_2 & 0 \\ 0 & \sigma_1 \end{pmatrix} = D$, which is Q. This shows that $Q = R^{-1}\,DR =$ D where R is orthogonal.

(viii) A symmetric linear transformation S is called positive semidefinite if $\boldsymbol{x}\bullet S\boldsymbol{x} \geq 0$*, for all* $\boldsymbol{x}$*. If M is any linear transformation of* $\mathbb{R}^3$*, show that M = RS, where R is orthogonal, and S is symmetric and positive-semidefinite. This is called the* ***Polar Decomposition*** *of M.*

The statement says that any linear transformation of $\mathbb{R}^3$ can be factored as a rotation or reflection applied to a positive-semidefinite linear transformation. We established in (i) that in $\mathbb{R}^3$, the SVD transformation, which applies to every linear transformation in that domain, generalizes to

$$M = R_{\varphi_X} \circ R_{\varphi_Y} \circ R_{\varphi_Z} \circ \Sigma \circ R_{-\theta_Z} \circ R_{-\theta_Y} \circ R_{-\theta_X}$$

where Σ is symmetric. Because each of the six rotations $R_{\{\theta_{(X,Y,Z)},\phi_{(X,Y,Z)}\}}$ corresponds to an orthogonal matrix $[R_\Theta]$, this can be written as $M = R_\varphi(\Sigma R_{-\theta}) \equiv \left(R_{\varphi_X} \circ R_{\varphi_Y} \circ R_{\varphi_Z}\right) \circ (\Sigma \circ R_{-\theta_Z} \circ R_{-\theta_Y} \circ R_{-\theta_X})$, where $\left[R_\varphi\right] = \left[R_{\varphi_X}\,R_{\varphi_Y}\,R_{\varphi_Z}\right]$ and $[R_{-\theta}] = [\Sigma][R_{-\theta_Z}\,R_{-\theta_Y}\,R_{-\theta_X}]$ are also orthogonal matrices. Since Σ is positive definite, $[S] = [\Sigma][R_{-\theta}]$ is also positive definite. Then we can let $R \equiv \left[R_\phi\right]$, $S \equiv [S]$, so that $M =$ RS as required.