<Review with some correction>

quadratic forms in $\mathbb{R}^n$ $\longleftrightarrow$ symmetric $n \times n$ matrix

$$\sum_{i=1}^{n} a_i x_i^2 + \sum_{1 \le i < j \le n} a_{ij} x_i x_j = \mathbf{x}^T A \mathbf{x}.$$

Since every symmetric matrix is orthogonally diagonalizable,

$\exists$ orthogonal P s.t. $P^TAP = D$, diagonal.

Hence $\mathbf{x}^T A \mathbf{x} = \mathbf{x}^T P D P^T \mathbf{x} = (P^T\mathbf{x})^T D (P^T \mathbf{x})$.

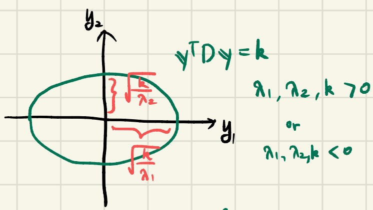
Set $\mathbf{y} = P^T \mathbf{x}$ (i.e., $\mathbf{x} = P\mathbf{y}$). Then $\mathbf{y}^T D \mathbf{y}$ has no cross product terms.

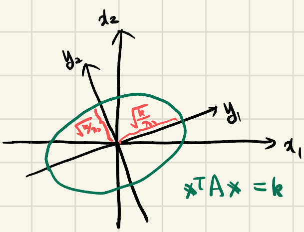
If $n=2$, then $\mathbf{y}^T D \mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2$

- $\lambda_1 = \lambda_2 \Rightarrow \mathbf{y}^T D \mathbf{y} = k$ is a circle if it exists
- $\lambda_1 \neq \lambda_2$ & $\lambda_1, \lambda_2 > 0 \Rightarrow \mathbf{y}^T D \mathbf{y} = k$ is an ellipse if it exists.
- $\lambda_1 \cdot \lambda_2 < 0 \Rightarrow \mathbf{y}^T D \mathbf{y} = k$ is a hyperbola.

If $n=3$, then $\mathbf{y}^T D \mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2$

$\mathbf{y}^T D \mathbf{y} = c$ is sphere, ellipsoid, or hyperboloid.





$\lambda_1 y_1^2 + \lambda_2 y_2^2 = k$

$\frac{\lambda_1}{k} y_1^2 + \frac{\lambda_2}{k} y_2^2 = 1$

Chang of coordinates

$P = [\mathbf{p}_1 \ \mathbf{p}_2]$

$= P_{B \to S}$

$\mathbf{x} = P\mathbf{y}$

$B = \{\mathbf{p}_1, \mathbf{p}_2\}$

$[\mathbf{p}_1]_B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$[\mathbf{p}_2]_B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

(2) What conditions must A satisfy for x^TAx to have positive values for $x \neq 0$?

<u>Def.</u> A quadratic form x^TAx is said to be

<u>positive definite</u> if $x^TAx > 0$ for $\forall x \neq 0$.

<u>negative definite</u> if $x^TAx < 0$ for $\forall x \neq 0$.

<u>indefinite</u> if x^TAx has both pos. & neg. values.

$\Rightarrow$ A is positive definite if x^TAx is pos. def.
negative definite neg. def.
indefinite indefinite.

$\therefore$ Problem (2) $\Longleftrightarrow$ When is a symmetric matrix A is positive definite?

<u>Thm 8.4.3</u> A is symmetric.

(1) x^TAx is pos. def. $\Longleftrightarrow$ all eigenvalues of A are positive.

(2) " " neg. def. $\Longleftrightarrow$ " " " " " negative.

(3) " " indef. $\Longleftrightarrow$ A has at least one pos. eigenvalue and one negative eigenvalue.

<u>Proof.</u> Exercise.

Note that $\begin{cases} x^TAx = y^TDy, & y = P^Tx \\ x \neq 0 \Leftrightarrow y \neq 0. \end{cases}$

<u>Rmk</u> $x^TAx \geq 0 \quad \forall x \Rightarrow$ positive semidefinite

$x^TAx \leq 0 \quad \forall x \Rightarrow$ negative semidefinite.

x^TAx is pos. semi-def. $\rightarrow$ all λ's are nonnegative.

x^TAx is neg. semi-def $\Rightarrow$ all λ's are nonpositive.

Thm 8.4.4 A : symm. 2×2

(a) $x^TAx = 1$ represents an ellipse if A is positive def.

(b) $x^TAx = 1$ has no graph if A is negative def.

(c) $x^TAx = 1$ represents a hyperbola if A is indefinite.

In general, if A is a pos. def. symm. matrix, then $x^TAx = 1$ represents a central ellipsoid. In particular, if $A=I$, then $x^TAx=1$ represents a sphere.

Identifying pos. def. matrix

For an $n\times n$ matrix $A=[a_{ij}]_{n\times n}$, the kth principal submatrix is $[a_{ij}]_{\substack{1\le i\le k\\ 1\le j\le k}}$

Thm 8.4.5 A symm. matrix A is positive def. iff the determinant of every principal submatrix is positive.

e.g.)
$$A = \begin{array}{c} \begin{array}{ccc} 2 & 3 & 1 \end{array} \\ \begin{bmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{bmatrix} \end{array}$$
is pos. def.

Thm 8.4.6 A : symm. T.F.A.E

(1) A is positive def.

(2) $\exists$ symm. pos. def. mx B s.t. $A=B^2$

(3) $\exists$ an inv. mx C s.t. $A=C^TC$.

pf) $(1)\Rightarrow(2)$ $P^TAP = D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$

Since $\forall \lambda_i > 0$, $\exists$ diagonal D_1 s.t. $D=D_1^2$ $\left(\text{i.e., } D_1 = \begin{bmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n} \end{bmatrix}\right)$

$A = PDP^T = PD_1P^TPD_1P^T = (PD_1P^T)^2$

PD_1P^T is pos. def. & symm.

$(2)\Rightarrow(3)$ Take $C=B$. $\Rightarrow$ $A=B^2=C^TC$.

$(3)\Rightarrow(1)$ $x^TAx = x^TC^TCx = (Cx)^T(Cx) = \|Cx\|^2 \ge 0$

Since C is inv., $x^TAx > 0$ $\forall x \ne 0$. $\square$

§ 8.6. Singular Value Decomposition

Let A be an $n\times n$ matrix of rank k.

Then A^TA is symmetric and $\text{rank}(A^TA)=k$.

Since ATA is symmetric, $\exists$ orthogonal matrix $V=[\mathbf{v}_1 \cdots \mathbf{v}_n]$ and diagonal matrix $D=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}$ s.t. $A^TA = VDV^T$ &

$A^TA\mathbf{v}_i = \lambda_i \mathbf{v}_i \quad (i=1,\cdots,n)$.

Since $\|A^TA\|^2 = \lambda_i$, every eigenvalue of ATA are non-negative.

Since $\text{rank}(A^TA)=k$, we may assume that

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_k > 0 = \lambda_{k+1} = \cdots = \lambda_n.$$

If $i\neq j$, then $(A\mathbf{v}_i)\cdot(A\mathbf{v}_j) = \lambda_j \mathbf{v}_i^T\mathbf{v}_j = 0$ ($\because$ V: orthogonal).

Hence $\{A\mathbf{v}_1, \cdots, A\mathbf{v}_n\}$ is orthogonal.

Since $\|A\mathbf{v}_i\|^2 = \lambda_i$, $A\mathbf{v}_1, \cdots, A\mathbf{v}_k$ are nonzero vectors and

$A\mathbf{v}_{k+1} = \cdots = A\mathbf{v}_n = \mathbf{0}$.

Let $\mathbf{u}_i := \dfrac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|}$ for $i=1,\cdots,k$. Then we get an orthonormal basis $\{\mathbf{u}_1,\cdots,\mathbf{u}_k\}$ for $\text{col}(A)$. Hence $\{\mathbf{u}_1,\cdots,\mathbf{u}_k\}$ can be enlarged to an orthonormal basis $\{\mathbf{u}_1,\cdots,\mathbf{u}_k,\mathbf{u}_{k+1},\cdots,\mathbf{u}_n\}$ for $\mathbb{R}^n$.

Let $U=[\mathbf{u}_1\ \mathbf{u}_2 \cdots \mathbf{u}_n]$ and $\Sigma=\begin{bmatrix}\sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n}\end{bmatrix}$.

Then $U\Sigma = AV$. Since V is orthogonal, we have

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

This is called the singular value decomposition of A.

Def.

$\sqrt{\lambda_i}$ is called a singular value of A.

$\mathbf{u}_1,\cdots,\mathbf{u}_k$ are called the left singular vectors.

$\mathbf{v}_1,\cdots,\mathbf{v}_k$ are called the right singular vectors.

Thm 8.6.4 (Singular value decomposition of a general matrix)

A : $m\times n$ matrix of rank k

$$\Rightarrow \quad A = U\Sigma V^T = \left[\mathbf{u}_1 \cdots \mathbf{u}_k \,\middle|\, \mathbf{u}_{k+1} \cdots \mathbf{u}_m\right]_{m\times m} \left[\begin{array}{ccc|c} \sqrt{\lambda_1} & & & \\ & \ddots & & \\ & & \sqrt{\lambda_k} & \\ \hline & & & 0 \end{array}\right]_{m\times n} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix}$$

Note that A^TA is a symmetric $n\times n$ matrix of rank k.

(a) $V = [\mathbf{v}_1 \cdots \mathbf{v}_n]$ orthogonally diagonalize A^TA.

(b) The nonzero diagonal entries of Σ are $\sigma_1 = \sqrt{\lambda_1}, \cdots, \sigma_k = \sqrt{\lambda_k}$, where $\lambda_1, \cdots, \lambda_k$ are eigenvalues of A^TA corr. to the column vectors $\mathbf{v}_1, \cdots, \mathbf{v}_k$ of V.

(c) The column vectors of V are ordered so that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k > 0$.

(d) $\mathbf{u}_i = \dfrac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \dfrac{1}{\sigma_i} A\mathbf{v}_i \quad (i = 1, \cdots, k)$

(e) $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$ is an orthonormal basis for $\mathrm{col}(A)$.

(f) $\{\mathbf{u}_1, \cdots, \mathbf{u}_k, \mathbf{u}_{k+1}, \cdots, \mathbf{u}_m\}$ is an extension of $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$ to an orthonormal basis for $\mathbb{R}^m$.

- $\sigma_1, \cdots, \sigma_k$: the singular values of A
- $\mathbf{u}_1, \cdots, \mathbf{u}_k$: the left singular vectors of A
- $\mathbf{v}_1, \cdots, \mathbf{v}_k$: the right singular vectors of A.
- $A = U\Sigma V^T$: the singular value decomposition of A.

Rmk For a symmetric matrix A, the singular values of A are the absolute values of the nonzero eigenvalues of A.

($\because$ Since $A^TA = A^2$, the eigenvalues of A^TA are λ_i^2.)

Thm 8.6.3 (Polar decomposition)

A : $n \times n$ matrix of rank k

$\Rightarrow$ $A = PQ$, where P is an $n\times n$ positive semidefinite matrix of rank k, and Q is an $n \times n$ orthogonal matrix. Moreover, if A is invertible, then P is positive definite.

pf) $A = U\Sigma V^T = U\Sigma U^T U V^T = \underbrace{(U\Sigma U^T)}_{\substack{P \\ \text{symmetric \&} \\ \text{positive semidefinite}}} \underbrace{(UV^T)}_{\text{orthogonal}}$ $\blacksquare$

Thm 8.6.5 A : $m \times n$ matrix with rank k

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

$\Rightarrow$ (a) $\{u_1, \cdots, u_k\}$: orthonormal basis for col(A)

(b) $\{u_{k+1}, \cdots, u_m\}$: orthonormal basis for null(A^T)

(c) $\{v_1, \cdots, v_k\}$: orthonormal basis for row(A)

(d) $\{v_{k+1}, \cdots, v_n\}$: orthonormal basis for null(A)

$$\Rightarrow A = [u_1 \cdots u_n] \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_k & \\ & & & 0 \ddots \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix} = [u_1 \cdots u_n] \begin{bmatrix} \sigma_1 v_1^T \\ \vdots \\ \sigma_k v_k^T \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

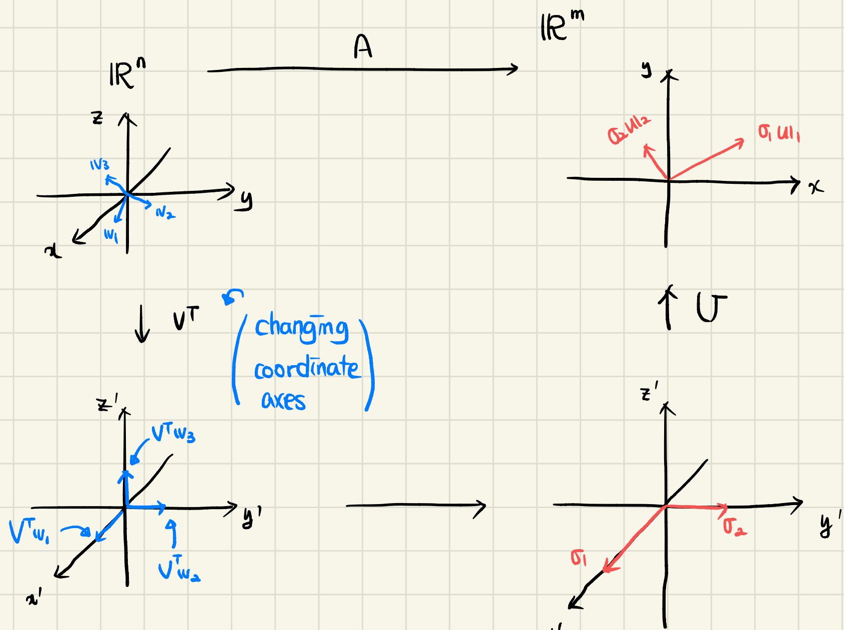
$$= \sigma_1 u_1 v_1^T + \cdots + \sigma_k u_k v_k^T = [u_1 \cdots u_k] \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_k^T \end{bmatrix} \leftarrow \text{reduced SVD}$$

Geometric meaning of SVD.

$A = U\Sigma V^T$, $\quad U$: orthogonal $m \times m$ matrix $\quad \mathcal{B} = \{u_1, \cdots, u_m\}$

$\quad V$: orthogonal $n \times n$ matrix $\quad \mathcal{B}' = \{v_1, \cdots, v_n\}$

e.g.) $A = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ v_3^T \end{bmatrix}$



$[v_i]_{\mathcal{B}'} = e_i$

* $T_A(x) = A(x_{row(A)} + x_{null(A)}) = A x_{row(A)}$

$\therefore$ range(T_A) = col(A) = $T_A($row$(A))$

* $T_A|_{row(A)}$: row$(A) \longrightarrow$ col(A) is one-to-one & onto.

* $\mathcal{B}_1 = \{v_1, \cdots, v_k\}$ is an orthonormal basis for row(A)

$\mathcal{B}_2 = \{u_1, \cdots, u_k\}$ is an orthonormal basis for col(A).

* SVD can be used to data compression.

$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + \sigma_k \mathbf{u}_k \mathbf{v}_k^T.$ ← Required storage space $= km + kn + k$

For $1 \leq r \leq k$,

set $A_r := \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + \sigma_k \mathbf{u}_r \mathbf{v}_r^T.$ ← Required storage space $= rm + rn + r \leq km + kn + k.$

We call A_r the rank r approximation of A.

In § 8.7, $\begin{cases} A^+ := U_1 \Sigma_1^{-1} V_1^T \text{ (the pseudoinverse of } A) \\ \mathbf{x} = A^+ \mathbf{b} \text{ is the least squares solution of } A\mathbf{x} = \mathbf{b}. \end{cases}$

* The nullity of $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ is TWO because the associated null space is a subspace of $\mathbb{R}^3$.