3. Symmetric and Positive Matrices

- Eigenvalues
- Eigenvectors
- Example: mechanical systems
- Quadratic forms
- Positive-definite matrices
- Matrix square-roots
- Example: mechanical systems
- Ellipsoids
- Ellipsoids in estimation
- Example: navigation
- Example: column interpretation
- Summary: ellipsoids in estimation

The key points of this section

we need to know how to recognize and avoid the *bad cases*; i.e.

- in control problems, even though $\mathrm{range}(A) = \mathbb{R}^m$, we need huge inputs to generate outputs in certain directions

  e.g., two columns are almost the same, or one column is very small

- in estimation problems, even though $\mathrm{null}(A) = \{0\}$, our sensors are very insensitive in certain directions

  e.g., two rows are almost the same, or one row is very small

to *computationally* detect these problems, we will use

- the *symmetric* matrices associated with the control/estimation problem
- for estimation: $A^T A$ for control: AA^T
- two things in particular: their *eigenvalues* and their *ellipsoids*

Eigenvalues of Symmetric Matrices

a matrix $A \in \mathbb{R}^{n \times n}$ is called *symmetric* if $A = A^T$

eigenvalues

> if A is symmetric, then the eigenvalues of A are real

to show this, suppose x is an eigenvector of A with eigenvalue $\lambda \in \mathbb{C}$. Then $Ax = \lambda x$ and $x \neq 0$

we will show that $\lambda = \bar{\lambda}$. we know that

$$\bar{x}^T A x = \lambda \bar{x}^T x = \lambda \sum_{i=1}^{n} |x_i|^2$$

also

$$\bar{x}^T A x = (\overline{Ax})^T x = \bar{\lambda} \bar{x}^T x = \bar{\lambda} \sum_{i=1}^{n} |x_i|^2$$

since $\|x\| \neq 0$, we have $\lambda = \bar{\lambda}$

Eigenvectors of Symmetric Matrices

for symmetric matrices:

> eigenvectors corresponding to distinct eigenvalues are orthogonal

suppose $Ax_1 = \lambda_1 x_1$ and $Ax_2 = \lambda_2 x_2$. Then

$$\begin{aligned} x_1^T A x_2 &= (Ax_1)^T x_2 \\ &= \lambda_1 x_1^T x_2 \end{aligned}$$

and also

$$x_1^T A x_2 = \lambda_2 x_1^T x_2$$

therefore

$$(\lambda_1 - \lambda_2) x_1^T x_2 = 0$$

and since $\lambda_1 \neq \lambda_2$ we must have $x_1^T x_2 = 0$

- if A is symmetric and has n distinct eigenvalues, then its eigenvectors form an *orthogonal basis* for the vector space $\mathbb{R}^n$

eigenvectors

fact: every $n \times n$ symmetric matrix A has n eigenvectors

$$\{q_1, \ldots, q_n\}$$

which form an *orthonormal basis* for $\mathbb{R}^n$.

in matrix notation

$$U = \begin{bmatrix} q_1 & q_2 & \ldots & q_n \end{bmatrix} \qquad \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

then, since q_i are eigenvectors, we have $AU = U\Lambda$ and so

$$\boxed{U^{-1}AU = \Lambda}$$

i.e. the columns of U are an orthonormal basis which *diagonalizes* A

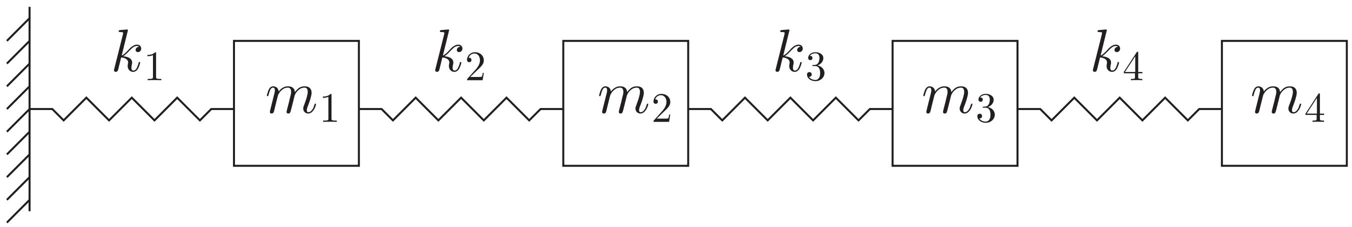
also U is an orthogonal matrix, so $U^{-1} = U^T$.

example: symmetric matrices in mechanical systems

for linear undamped mechanical systems (e.g. elasticity), the equations of motion are

$$M\ddot{x}(t) + Kx(t) = 0$$

where M and K are symmetric matrices. for example:



for this system,

$$M = I \qquad K = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

with $q = U^{-1}x$ and $K = U\Lambda U^{-1}$, we have

$$\ddot{q}(t) + \Lambda q(t) = 0$$

and the motion decomposes into 4 *normal modes*

Quadratic Forms

if A is a symmetric matrix, the quadratic function $f(x) = x^T Ax$ which maps $\mathbb{R}^n$ to $\mathbb{R}$ is called a *quadratic form.* for example,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}^T \begin{bmatrix} 1 & 2 & -1 \\ 2 & -3 & 0 \\ -1 & 0 & 7 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = x^2 + 4xy - 2xz - 3y^2 + 7z^2$$

positive definite matrices

the symmetric matrix A is called *positive definite*, written $A > 0$, if

$$x^T Ax > 0 \qquad \text{for all nonzero } x \in \mathbb{R}^n$$

examples

- cI is positive definite if and only if $c > 0$.
- $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is positive definite if and only if $a > 0$, $c > 0$ and $b^2 < ac$

quadratic forms and eigenvalues

suppose A is symmetric with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. then

$$\lambda_1 \|x\|^2 \geq x^T A x \geq \lambda_n \|x\|^2 \qquad \text{for all } x \in \mathbb{R}^n$$

to see this, let $\{q_1, \dots, q_n\}$ be orthonormal eigenvectors of A with the above eigenvalues. then

$$\begin{aligned} x^T A x &= \sum_{i=1}^n \lambda_i |q_i^T x|^2 \\ &\geq \lambda_n \sum_{i=1}^n |q_i^T x|^2 = \lambda_n \|x\|^2 \end{aligned}$$

notes

- picking $x = q_n$ gives $x^T A x = \lambda_n \|x\|^2$ and so the inequality is *tight*
- this gives the following test for positive definiteness:

$$\boxed{A > 0 \qquad \Longleftrightarrow \qquad \lambda_{\text{min}}(A) > 0}$$

notation for symmetric matrices

- A is called *positive semidefinite* if

$$x^T A x \geq 0 \qquad \text{for all } x \in \mathbb{R}^n$$

- we write $A < 0$ to mean $-A > 0$. In this case, we say A is *negative definite*.
- we write $A > B$ to mean $A - B > 0$.
- similar notation for semidefiniteness.
- a matrix which is neither positive semidefinite nor negative semidefinite is called *indefinite*.

properties of positive definite matrices

- *addition of positive matrices*

$$A > 0 \text{ and } B > 0 \qquad \Longrightarrow \qquad A + B > 0$$

- *block diagonal matrices*

$$A > 0 \text{ and } B > 0 \qquad \Longleftrightarrow \qquad \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} > 0$$

- *invertibility*

$$A > 0 \qquad \Longrightarrow \qquad A \text{ is nonsingular}$$

- *scaling*

$$A, B \geq 0 \quad \text{and} \quad \lambda, \mu \geq 0 \qquad \Longrightarrow \qquad \lambda A + \mu B \geq 0$$

matrix square roots

if A is positive semidefinite, then

$$\begin{aligned} A &= U\Lambda U^T \\ &= U\Lambda^{\frac{1}{2}}U^T U\Lambda^{\frac{1}{2}}U^T \\ &= \left(U\Lambda^{\frac{1}{2}}U^T\right)^2 \end{aligned}$$

where

$$\Lambda^{\frac{1}{2}} = \operatorname{diag}(\sqrt{\lambda}_1, \sqrt{\lambda}_2, \dots, \sqrt{\lambda}_n)$$

the matrix $U\Lambda^{\frac{1}{2}}U^T$ is called the *square root* of A

example: mechanical systems

as before, but drop the assumption that $M = I$

$$M\ddot{x}(t) + Kx(t) = 0$$

usually $M > 0$, so let $z(t) = M^{\frac{1}{2}}x(t)$, so

$$\ddot{z}(t) + M^{-\frac{1}{2}}KM^{-\frac{1}{2}}z(t) = 0$$

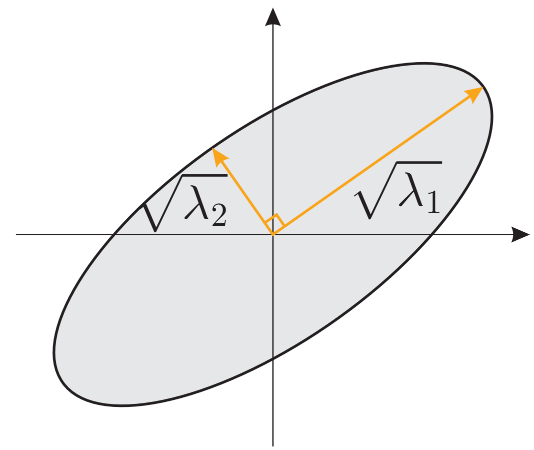
and the eigenvectors of $M^{-\frac{1}{2}}KM^{-\frac{1}{2}}$ give the normal modes.

Ellipsoids

an ellipsoid is a sphere stretched in orthogonal directions called the *principal semiaxes*

every positive definite matrix has a corresponding ellipsoid

$$E = \{\, x \in \mathbb{R}^n \mid x^T B^{-1} x \leq 1 \,\}$$



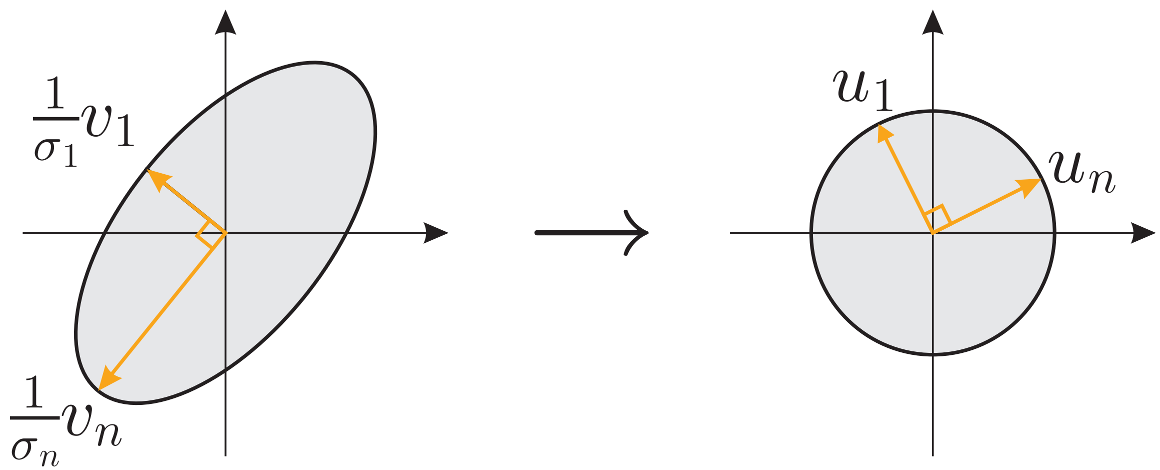
Notes

- $B \in \mathbb{R}^{n\times n}$, $B = B^T$, $B > 0$.
- semiaxis lengths: $\sqrt{\lambda_i}$, where λ_i are eigenvalues of B
- semiaxis directions: eigenvectors of B
- B is positive definite if and only if B^{-1} is, and they have the same eigenvectors. using B instead of B^{-1} replaces $\sqrt{\lambda_i}$ by $\frac{1}{\sqrt{\lambda_i}}$.

ellipsoids in estimation

usual linear equation $y = Ax$ gives $\|y\|^2 = x^T A^T A x$, so

$$\{\, x \in \mathbb{R}^n \mid x^T A^T A x \le 1 \,\} \quad \text{maps into} \quad \{\, y \in \mathbb{R}^m \mid \|y\| \le 1 \,\}$$



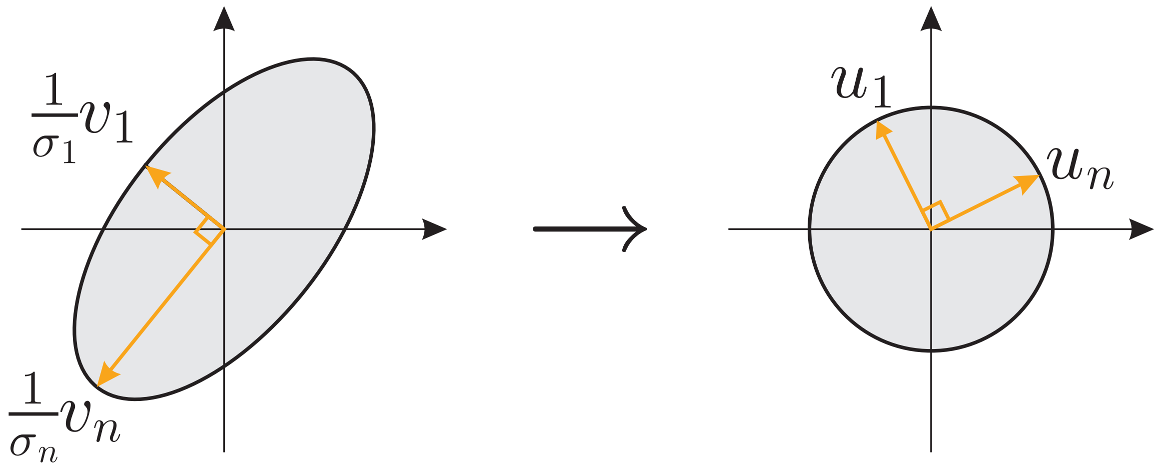
the *unit ball*

here

- $\sigma_i^2 = i$th eigenvalue of $A^T A$; convention $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r > 0$, called *singular values* of A
- v_i is unit eigenvector of $A^T A$ corresponding to σ_i
- $u_i = \frac{1}{\sigma_i} A v_i$, are orthogonal
- assume A is skinny and full rank, then $A^T A$ is positive definite, since

$$x^T A^T A x = (Ax)^T (Ax) \qquad \text{and} \qquad Ax \ne 0 \text{ for all } x \ne 0$$

ellipsoids in estimation



interpretation

- short axis of ellipsoid (eigenvector v_1 corresponding $\lambda_{\text{max}}(A^TA)$) is *stretched most* by sensing.
- long axis of ellipsoid (eigenvector v_n corresponding $\lambda_{\text{min}}(A^TA)$) is *stretched least* by sensing.

therefore

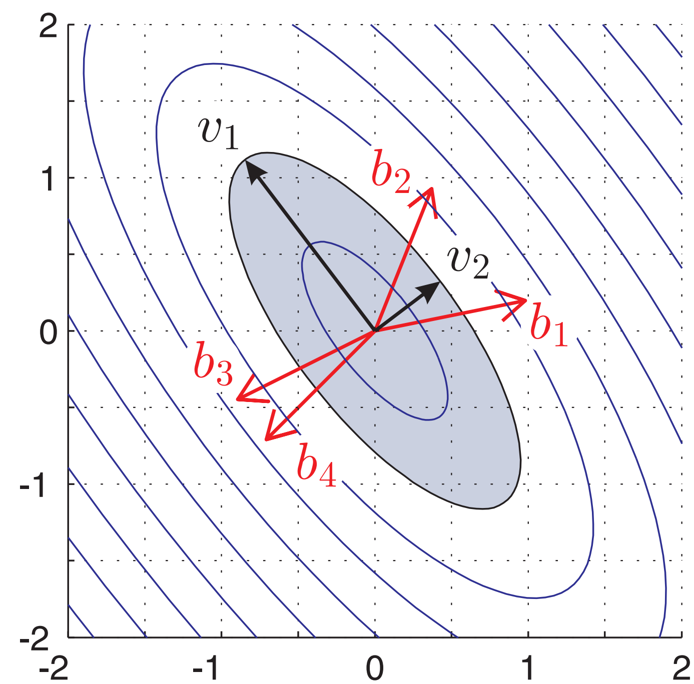
- small changes to x in the direction v_1 cause large changes in sensor readings y; sensors are *highly sensitive*
- small changes to x in the direction v_n cause small changes in sensor readings y; sensors are *insensitive*

example: navigation

here $A \in \mathbb{R}^{4\times 2}$ with

$$A = \begin{bmatrix} b_1^T \\ b_2^T \\ b_3^T \\ b_4^T \end{bmatrix}$$

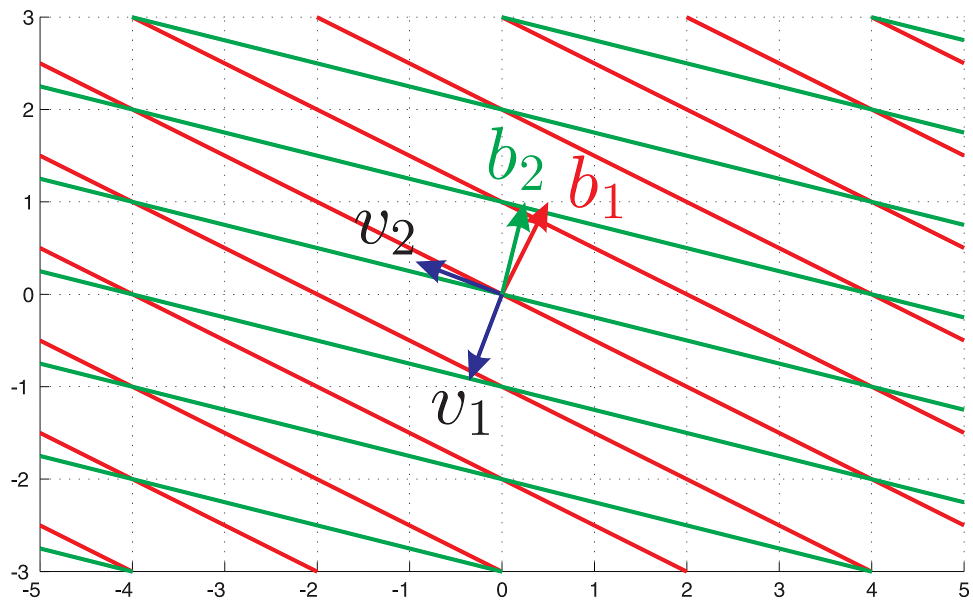
and $y = Ax$. Each b_i is a unit vector.



- x is unknown.
- y is measured, with y_i the component of x in the direction b_i
- the ellipsoid is the set of $x \in \mathbb{R}^2$ which result in $\|y\| \leq 1$
- plot shows contours of $\|y\|$, i.e., contours of $\sqrt{x^T A^T A x}$

the sensors are most sensitive to the component of x along the *short axis* of the estimation ellipsoid

example: row interpretation



$$A = \begin{bmatrix} b_1^T \\ b_2^T \end{bmatrix} = \begin{bmatrix} 0.5 & 1 \\ 0.25 & 1 \end{bmatrix}$$

v_1 and v_2 are eigenvectors of $A^T A$ with corresponding singular values

$$\sigma_1 \approx 1.5117 \qquad \sigma_2 \approx 0.1654$$

sensors are approx 10 times more sensitive to changes in x in the v_1 direction than in v_2 direction.

summary: ellipsoids in estimation

we have $y = Ax$

- the square-roots of the eigenvalues of $A^T A$ are the *singular values*
- small singular values mean we have poor sensor sensitivity in the direction of the corresponding eigenvector
- if a singular value is 0, then we cannot sense anything in the corresponding direction; i.e. the null space is non-empty

  because if x is an eigenvector of $A^T A$ with eigenvalue 0,

$$x^T A^T A x = 0 \quad \Longrightarrow \quad (Ax)^T (Ax) = 0 \quad \Longrightarrow \quad \|Ax\| = 0 \quad \Longrightarrow \quad Ax = 0$$

in practice, we don't compute the eigenvalues of $A^T A$ using Matlab's `eig` routine

we use a more general/powerful tool, which works with A not $A^T A$, (saving numerical accuracy) and works for control as well as estimation problems . . .

. . . the singular value decomposition