Matrices and polynomials

Fact #1. One has $f(A)=0 \iff f(J)=0$.

This is because $J=B^{-1}AB$ and so $f(J)=\sum_{k=0}^{n} c_k J^k = \sum_{k=0}^{n} c_k B^{-1}A^k B = B^{-1}f(A)B$.

Fact #2. One has $f(A)=0 \implies f(\lambda)=0 \quad \forall$ eigenvalue λ.

This is because $A\vec{v}=\lambda\vec{v}$ and so $0=f(A)\vec{v}=\underbrace{\sum c_k A^k}\vec{v}=\underbrace{\sum c_k\lambda^k}\vec{v}=f(\lambda)\vec{v}$.

Theorem (Cayley – Hamilton) We have $f(A)=0$ when $f(\lambda)=\det(A-\lambda I)$ is the char. polynomial of A.

Proof. ⊙ If $A=\begin{bmatrix}\lambda & & \\ 1 & \ddots & \\ & 1 & \lambda\end{bmatrix}$ is a $k\times k$ block,

then $f(x)=\det(A-xI)=(\lambda-x)^k$ is the char. polynomial

and $(A-\lambda I)^k=0$ so A satisfies $(x-\lambda)^k=\pm(\lambda-x)^k$.

⊙ If $A=\left[\begin{array}{c|c} J_1 & \\ \hline & J_2\end{array}\right]$ consists of two blocks,

then
$$f(x)=\det(A-xI)=\det\left[\begin{array}{c|c} J_1-xI & \\ \hline & J_2-xI\end{array}\right]$$
$$=\det\left[\begin{array}{c|c} J_1-xI & \\ \hline & I\end{array}\right]\left[\begin{array}{c|c} I & \\ \hline & J_2-xI\end{array}\right]$$
$$=\det(J_1-xI)\cdot\det(J_2-xI).$$

This implies $f(x)=f_1(x)\cdot f_2(x)$ and we also have
$$f(A)=\sum c_k A^k=\sum c_k\left[\begin{array}{c|c} J_1^k & \\ \hline & J_2^k\end{array}\right]=\left[\begin{array}{c|c} f(J_1) & \\ \hline & f(J_2)\end{array}\right]=\left[\begin{array}{c|c} \cancel{f_1(J_1)}f_2(J_1) & \\ \hline & f_1(J_2)\cancel{f_2(J_2)}\end{array}\right]$$
$$=0.$$

⊙ If $A=\begin{bmatrix}J_1 & & \\ & \ddots & \\ & & J_m\end{bmatrix}$ consists of m blocks, then $f(A)=0$ by induction. This proves the statement for Jordan forms.

⊙ If A is an arbitrary matrix, then $f(J)=0$ by above so $f(A)=0$ as well. ▨

$\boxed{\text{Minimal polynomial}}$ We know $f(A)=0$ for <u>some</u> polynomial f.

We define $m(\lambda)$ = the polynomial of <u>lowest</u> degree satisfied by f.

We require $m(\lambda)$ to be <u>monic</u>, namely highest coeff $= 1$.

Then $m(\lambda)$ is easily seen to be unique.

<u>Example 1.</u> Let $A = \begin{bmatrix} 2 & & \\ & 2 & \\ & 1 & 2 \end{bmatrix}$.

Char polynomial $= \det(A-\lambda I) = (2-\lambda)^3$ is cubic.

Minimal polynomial $= (\lambda-2)^2 = (2-\lambda)^2$ is quadratic. In fact,

$$A-2I = \begin{bmatrix} 0 & & \\ & 0 & \\ & 1 & 0 \end{bmatrix} \Rightarrow (A-2I)^2 = \begin{bmatrix} 0 & & \\ & 0 & \\ & 0 & 0 \end{bmatrix} = 0.$$

The min. polynomial is related to the sizes of the blocks.

<u>Example 2</u> (Computing inverses) Let A be as before. We know $m(\lambda) = (\lambda-2)^2 = \lambda^2 - 4\lambda + 4$ is the minimal polynomial.

Thus $A^2 - 4A + 4I = 0 \Rightarrow 4I = 4A - A^2$

$\Rightarrow 4I = A(4I - A)$

and $A^{-1} = \frac{1}{4}(4I - A) = \frac{1}{4}\begin{bmatrix} 2 & & \\ & 2 & \\ & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1/2 & & \\ & 1/2 & \\ & -1/4 & 1/2 \end{bmatrix}$.

<u>Example 3</u> (Computing powers) Suppose A is $n \times n$ with $m(\lambda) = \lambda(\lambda-2)$.

Then $A(A-2I) = 0 \Rightarrow \boxed{A^2 = 2A}$

$\Rightarrow A^3 = 2A^2 = 2(2A) = 2^2 A$

$\Rightarrow A^4 = 2^2 A^2 = 2^2(2A) = 2^3 A$ etc.

so that $A^k = 2^{k-1} A$ for all k.

<u>Fact #1.</u> Let $m(\lambda)$ = min poly of A, $f(\lambda)$ = char. poly of A.

Then $m(\lambda)$ divides $f(\lambda)$ and $m(\lambda_i) = 0$ $\forall$ eigenvalue λ_i.

In particular, $m(\lambda) = (\lambda-\lambda_1)^{k_1} \cdots (\lambda-\lambda_p)^{k_p}$ for some $k_i \geq 1$.

Proof: We know A satisfies $f(\lambda)$ = char. polynomial. It also satisfies $m(\lambda)$ by definition. Divide these two.

$$\begin{array}{r|l} & Q(\lambda) \\ \hline m(\lambda) & f(\lambda) \\ & \vdots \\ & R(\lambda) \end{array}$$

We get $f(\lambda) = m(\lambda)\,Q(\lambda) + R(\lambda)$

with the remainder R either 0

or of $\deg R < \deg m$.

We know $f(A) = 0$ and $m(A) = 0$, so $R(A) = 0$.

This gives $R = 0$ by minimality.

This proves the first part, that $m(\lambda)$ divides $f(\lambda)$. For the second part $m(\lambda_i) = 0$, note that

$$A\vec{v} = \lambda\vec{v} \Rightarrow A^k\vec{v} = \lambda^k\vec{v} \Rightarrow m(A)\vec{v} = m(\lambda)\vec{v}$$
$$\Rightarrow m(\lambda) = 0.$$

<u>Fact #2.</u> One has $m(\lambda) = (\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\cdots(\lambda-\lambda_p)^{k_p}$ with k_i the size of the *largest* Jordan block with eigenvalue λ_i.

<u>Proof.</u> Min poly of A = Min poly of J. Suppose $J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_p \end{bmatrix}$ with J_i being $k_i \times k_i$ with eigenvalue λ_i. Then

$$m(J) = \sum c_k J^k = \sum c_k \begin{bmatrix} J_1^k & & \\ & \ddots & \\ & & J_p^k \end{bmatrix} = \begin{bmatrix} m(J_1) & & \\ & \ddots & \\ & & m(J_p) \end{bmatrix}$$

so $m(J) = 0$ if and only if <u>$m(J_i) = 0$</u> for all i.

On the other hand, $J_i = \begin{bmatrix} \lambda_i & & \\ 1 & \ddots & \\ & 1 & \lambda_i \end{bmatrix} \Rightarrow (J_i - \lambda_i I)^{k_i} = 0$

with $(J_i - \lambda_i I)^j \neq 0$ if $j < k$.

This implies exponent k_i = size of *largest* Jordan block.

<u>Example 1.</u> Suppose A is 3×3 with $\lambda = 1, 2, 3$. Then $J = \begin{bmatrix} \boxed{1} & & \\ & \boxed{2} & \\ & & \boxed{3} \end{bmatrix}$.

Then $k_i = 1$ for all i and $m(\lambda) = (\lambda-1)(\lambda-2)(\lambda-3)$.

<u>Example 2.</u> Suppose A is 4×4 with $J = \begin{bmatrix} 2 & & & \\ 1 & 2 & & \\ & & \boxed{2} & \\ & & & \boxed{1} \end{bmatrix}$.

Then $m(\lambda) = (\lambda-2)^2(\lambda-1)$ and $f(\lambda) = \det(A-\lambda I) = \det(J-\lambda I) = (2-\lambda)^3(1-\lambda)$.

Example 3. Suppose $f(\lambda) = -(\lambda-1)^2(\lambda-3)^3$ and $m(\lambda) = (\lambda-1)(\lambda-3)^2$.
What is the Jordan form?

$\boxed{\lambda=1} \rightsquigarrow$ double eigenvalue that contributes 1×1 blocks

$\boxed{\lambda=3} \rightsquigarrow$ triple eigenvalue that contributes 1×1 or 2×2 blocks (largest being 2×2).

Thus $$J = \begin{bmatrix} \boxed{1} & & & & \\ & \boxed{1} & & & \\ & & 3 & & \\ & & 1 & 3 & \\ & & & & \boxed{3} \end{bmatrix}.$$

Bilinear forms

A bilinear form on a real vector space V is a function $f: V\times V \to \mathbb{R}$ that is linear in each variable. We require:

$$f(\vec{v}_1+\vec{v}_2, \vec{w}) = f(\vec{v}_1,\vec{w}) + f(\vec{v}_2,\vec{w})$$

and $f(c\vec{v},\vec{w}) = c\cdot f(\vec{v},\vec{w})$ for any $c\in\mathbb{R}$

and similar conditions for the second variable.

- Standard example is the dot product in $\mathbb{R}^n$,

$$f(\vec{x},\vec{y}) = \vec{x}\bullet\vec{y} = \sum_{i=1}^{n} x_i y_i$$

This is easily seen to be bilinear. Note that $\bullet$ does not represent matrix multipl. since $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ is $n\times1$ and $\vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ is $n\times1$.

We could write

$$\vec{x}^t\cdot\vec{y} = \underbrace{[x_1 \cdots x_n]}_{1\times n}\cdot\underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}}_{n\times1} = \underbrace{\left[\sum x_i y_i\right]}_{1\times1} = \sum x_i y_i$$

by identifying 1×1 matrices with their single entry.

Thus $f(\vec{x},\vec{y}) = \vec{x}^t\vec{y}$ is a bilinear form on $\mathbb{R}^n$.

Theorem. Every bilinear form on $\mathbb{R}^n$ has the form
$$f(\vec{x}, \vec{y}) = \sum_{i,j} a_{ij} x_i y_j \quad \text{for some scalars } a_{ij}.$$

Proof. One has
$$f(\vec{x}, \vec{y}) = f\left(\sum_i x_i \vec{e}_i, \sum_j y_j \vec{e}_j\right)$$
$$= \sum_{i,j} f(x_i \vec{e}_i, y_j \vec{e}_j)$$
$$= \sum_{i,j} x_i y_j f(\vec{e}_i, \vec{e}_j)$$
and the result follows by letting $a_{ij} = f(\vec{e}_i, \vec{e}_j)$. $\blacksquare$

Matrix of a bilinear form Suppose $f: V \times V \to \mathbb{R}$ is a bilinear form on a real vector space V and let $v_1, v_2, \dots, v_n$ be a basis of V.
Then
$$f(\vec{v}, \vec{w}) = f\left(\sum_i x_i \vec{v}_i, \sum_j y_j \vec{v}_j\right)$$
$$= \sum_{i,j} f(x_i \vec{v}_i, y_j \vec{v}_j) = \sum_{i,j} x_i y_j \underline{f(\vec{v}_i, \vec{v}_j)}$$
$$= \sum_{i,j} a_{ij} x_i y_j,$$
where $a_{ij} = f(v_i, v_j)$. We call the matrix A the matrix of the bilinear form w.r.t. the basis $\vec{v}_1, \dots, \vec{v}_n$.

Example 1. Let P_2 = polynomials of degree at most 2.
We define $f(p(x), q(x)) = \int_0^1 p(x) q(x)\, dx$ ... for all $p, q \in P_2$.
We claim this is bilinear. In fact,
$$f(p_1(x) + p_2(x), q(x)) = \int_0^1 (p_1 + p_2) q \, dx = \int_0^1 p_1 q + \int_0^1 p_2 q$$
$$= f(p_1, q) + f(p_2, q)$$
and
$$f(cp(x), q(x)) = \int_0^1 c p(x) q(x)\, dx = c \int_0^1 p(x) q(x)\, dx = c f(p, q).$$

Note. For ease of notation, we'll write $<p,q>$ for bilinear forms instead of $f(p,q)$ for simplicity. In this case

$$<p,q> = \int_0^1 p(x)\, q(x)\, dx.$$

Let's compute the matrix w.r.t. the standard basis of P_2, namely $\vec{v}_1 = 1$, $\vec{v}_2 = x$, $\vec{v}_3 = x^2$. By definition $\boxed{a_{ij} = <\vec{v}_i, \vec{v}_j>.}$

We can compute all the entries by writing $\vec{v}_i = x^{i-1}$ $\forall i$.

Then $a_{ij} = <\vec{v}_i, \vec{v}_j> = <x^{i-1}, x^{j-1}> = \int_0^1 x^{i-1} \cdot x^{j-1}\, dx$

$$= \left[\frac{x^{i+j-1}}{i+j-1} \right]_{x=0}^{1} = \frac{1}{i+j-1} \quad \text{for } i,j = 1,2,3.$$

The matrix is thus

$$A = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}.$$

Example 2. Consider $\mathbb{R}^2$ with $<\vec{x}, \vec{y}> = 2x_1y_1 + x_1y_2 + 3x_2y_1 + 4x_2y_2$.

This is bilinear $<x,y> = \sum a_{ij} x_i y_j$. We take the standard basis $\vec{v}_1 = \vec{e}_1$, $v_2 = \vec{e}_2$. What is the matrix A of the form?

We have

$$a_{11} = <\vec{v}_1, \vec{v}_1> = <\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}> = 2 \quad \text{----- } x_1y_1 \text{ coefficient}$$

$$a_{12} = <\vec{v}_1, \vec{v}_2> = <\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}> = 1 \quad \text{---- } x_1y_2 \text{ coeff}$$

$$a_{21} = <\vec{v}_2, \vec{v}_1> = <\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}> = 3 \quad \text{---- } x_2y_1 \text{ coeff}$$

$$a_{22} = <\vec{v}_2, \vec{v}_2> = <\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}> = 4 \quad \text{--- } x_2y_2 \text{ coeff.}$$

More generally, $<\vec{x}, \vec{y}> = \sum a_{ij} x_i y_j$ has matrix A with respect to the standard basis because $<\vec{e}_i, \vec{e}_j> = a_{ij}$.

Notation: We could also write $\langle \vec{x}, \vec{y} \rangle = \sum a_{ij} x_i y_j$ in terms of matrices. In fact,

$$\underset{1\times n}{\vec{x}^{\,t}} \underset{n\times n}{A} \underset{n\times 1}{\vec{y}} = \left(\sum x_i \vec{e}_i\right)^t A \left(\sum y_j \vec{e}_j\right)$$

$$= \sum_i x_i \vec{e}_i^{\,t} \cdot A \cdot \sum_j y_j \vec{e}_j$$

$$= \sum x_i y_j \left(\vec{e}_i^{\,t} A \vec{e}_j\right).$$

Here, $A\vec{e}_j$ = jth column of A because $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}$

$\vec{e}_i^{\,t} A$ = ith row of A because $[1\ 0 \ldots 0] \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = [a_{11} \cdots a_{1n}]$

$\therefore$ $\vec{e}_i^{\,t} A \vec{e}_j$ = ith row of jth column of A = $[a_{ij}] = a_{ij}$.

$$\boxed{\text{Conclusion} \ldots \langle \vec{x}, \vec{y} \rangle = \sum a_{ij} x_i y_j \text{ becomes } \langle \vec{x}, \vec{y} \rangle = \vec{x}^{\,t} A \vec{y}.}$$

- Standard basis $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$. Then $\langle \vec{e}_i, \vec{e}_j \rangle = \vec{e}_i^{\,t} A \vec{e}_j = a_{ij}$ so the matrix of the form is just A.
- Basis of eigenvectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$. Then $\langle \vec{v}_i, \vec{v}_j \rangle = \vec{v}_i^{\,t} \underline{A \vec{v}_j}$

$$= \lambda_j \left(\vec{v}_i^{\,t} \vec{v}_j\right)$$

$$= \lambda_j \left(\vec{v}_i \cdot \vec{v}_j\right).$$

One may thus compute $\langle \vec{v}_i, \vec{v}_j \rangle$ easily for all i, j.

Dot/ Inner product We say $\langle \vec{x}, \vec{y} \rangle$ is an inner product on a real vector space V, if it is ① bilinear, ② symmetric in the sense that $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$ for all $\vec{x}, \vec{y} \in V$ and ③ positive definite $\langle \vec{x}, \vec{x} \rangle$ is positive for all $\vec{x} \neq 0$.

Example A. Consider $\mathbb{R}^n$ with $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^{n} x_i y_i = \vec{x}^t \vec{y}$.
This is bilinear, symmetric and pos. def. since $\langle \vec{x}, \vec{x} \rangle = \sum x_i^2 \geq 0$ with equality if and only if $\vec{x} = 0$.

Example B. Consider P_n = polynomials of degree $\leq n$. We define
$$\langle f, g \rangle = \int_a^b f(x)\, g(x)\, dx$$
with $a<b$ fixed. This is bilinear, symmetric with $\langle f, f \rangle = \int_a^b f(x)^2\, dx \geq 0$. Moreover,
$$\int_a^b f(x)^2\, dx = 0 \iff f(x) = 0$$
for all x ..... [by continuity].

Example C. (Checking symmetry) Consider $\mathbb{R}^n$ with $\langle \vec{x}, \vec{y} \rangle = \sum a_{ij} x_i y_j = \vec{x}^t A \vec{y}$.

Symmetric means ---- $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$ for all $\vec{x}, \vec{y}$

namely ----- $\sum a_{ij} x_i y_j = \sum a_{ij} y_i x_j$.

This is true precisely when $\underline{a_{ij} = a_{ji} \text{ for all } i, j}$

or simply $\underline{A = A^t}$

For instance, $A = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$ gives a symmetric form

and so does $A = \begin{bmatrix} 32 & 3 & 2 \\ 3 & 64 & 4 \\ 2 & 4 & 51 \end{bmatrix}$.

However, $A = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$ is not symmetric.

Example D. (Checking form is pos. def.) This is generally hard.

Consider $\langle \vec{x}, \vec{y} \rangle = x_1 y_1 + x_1 y_2 + 3 x_2 y_1 + 6 x_2 y_2$ on $\mathbb{R}^2$.

Then
$$\begin{aligned} \langle \vec{x}, \vec{x} \rangle &= x_1^2 + x_1 x_2 + 3 x_2 x_1 + 6 x_2^2 \\ &= \underbrace{x_1^2 + 4 x_1 x_2}_{} + 6 x_2^2 \\ &= (x_1 + 2x_2)^2 + 2 x_2^2 \quad \text{by completing the square.} \end{aligned}$$

This implies $\langle \vec{x}, \vec{x} \rangle$ is non-negative and
$\langle \vec{x}, \vec{x} \rangle = 0 \iff x_1 + 2x_2 = x_2 = 0 \iff x_1 = x_2 = 0$.

Example E. Consider $\langle \vec{x}, \vec{y} \rangle = x_1 y_1 + x_1 y_2 + 3 x_2 y_1 + 3 x_2 y_2$.

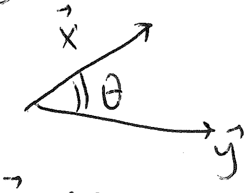
Then
$$\begin{aligned} \langle \vec{x}, \vec{x} \rangle &= x_1^2 + 4 x_1 x_2 + 3 x_2^2 \\ &= (x_1 + 2x_2)^2 - x_2^2 \end{aligned}$$

and this is not pos. def. For instance, take $\left\{ \begin{matrix} x_1 = -2x_2 \\ x_2 = 1 \end{matrix} \right\}$.

Then $\langle \vec{x}, \vec{x} \rangle = -1$ but $\vec{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ is nonzero.

$\boxed{\text{Angles / Length}}$ Once we have $\langle \vec{x}, \vec{y} \rangle$, a dot/inner product, we can define length $\langle \vec{x}, \vec{x} \rangle = \|\vec{x}\|^2$, namely $\|\vec{x}\| = \sqrt{\langle \vec{x}, \vec{x} \rangle}$.
We can also define angles between $\vec{x}$ and $\vec{y}$ by

$$\langle \vec{x}, \vec{y} \rangle = \|\vec{x}\| \cdot \|\vec{y}\| \cdot \cos\theta,$$



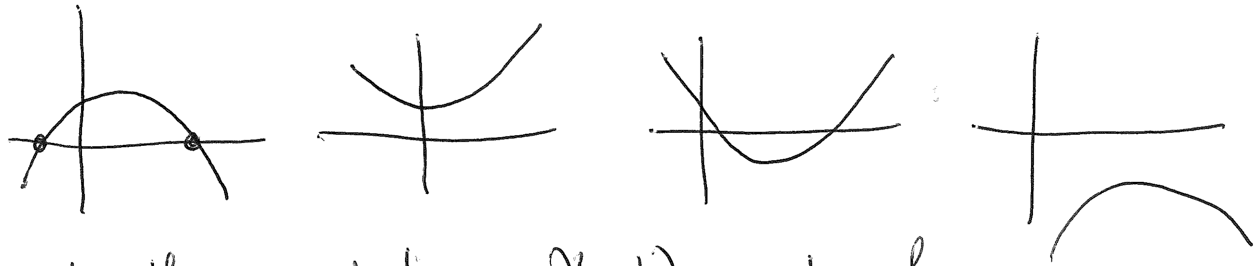
namely
$$\cos\theta = \frac{\langle \vec{x}, \vec{y} \rangle}{\|x\| \cdot \|y\|} \quad \text{whenever } \vec{x}, \vec{y} \neq 0.$$

Theorem (Cauchy-Schwarz inequality) One has $|\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \cdot \|y\|$.

Proof. Consider $0 \leq \|\vec{x} + \lambda \vec{y}\|^2 = \langle \vec{x} + \lambda \vec{y}, \vec{x} + \lambda \vec{y} \rangle$
for any $\lambda \in \mathbb{R}$. We get $0 \leq \langle \vec{x}, \vec{x} \rangle + \langle \vec{x}, \lambda \vec{y} \rangle + \langle \lambda \vec{y}, \vec{x} \rangle + \langle \lambda \vec{y}, \lambda \vec{y} \rangle$

Thus $\quad 0 \leq \|\vec{x}\|^2 + 2\lambda \langle \vec{x}, \vec{y} \rangle + \lambda^2 \|\vec{y}\|^2 \; \dots \; \forall \lambda \in \mathbb{R}.$

When is $\quad a\lambda^2 + b\lambda + c \geq 0 \quad \forall \lambda \in \mathbb{R}$???



We need the quadratic $a\lambda^2 + b\lambda + c$ to have

$a \geq 0$ and $b^2 - 4ac \leq 0$ (to avoid two real roots).

We have $a = \|\vec{y}\|^2 \geq 0$. We get $b^2 \leq 4ac$.

$$\Rightarrow \quad 2^2 \langle x, y \rangle^2 \leq 4 \|\vec{y}\|^2 \|\vec{x}\|^2$$

$$\Rightarrow \quad |\langle x, y \rangle| \leq \|\vec{x}\| \cdot \|\vec{y}\|. \qquad \blacksquare$$

Orthogonality

- $\vec{x}$ is orthogonal to $\vec{y}$ means $\langle \vec{x}, \vec{y} \rangle = 0$
- $\vec{v}_1, \dots, \vec{v}_n$ orthogonal basis of V means $\langle \vec{v}_i, \vec{v}_j \rangle = 0 \quad \forall i \neq j$.
- $\vec{v}_1, \dots, \vec{v}_n$ orthonormal basis means orthogonal with $\|\vec{v}_i\| = 1$.

Theorem (Orthogonal coordinates) If $v_1, \dots, v_n$ is an orthogonal basis and $\vec{v} \in V$ is arbitrary, then $\vec{v} = \sum_{i=1}^{n} c_i \vec{v}_i = \sum_{i=1}^{n} \frac{\langle v_i, v \rangle}{\langle v_i, v_i \rangle} \vec{v}_i$.

Proof. Since $\vec{v}_1, \dots, \vec{v}_n$ form a basis, $\vec{v} = \sum c_i \vec{v}_i$ for some scalars c_i.
Taking the dot product with $\vec{v}_j$ gives

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_n\vec{v}_n \Rightarrow \langle \vec{v}, \vec{v}_j \rangle = c_1 \langle \vec{v}_1, \vec{v}_j \rangle + \dots + c_n \langle \vec{v}_n, \vec{v}_j \rangle$$

$$\Rightarrow \langle \vec{v}, \vec{v}_j \rangle = c_j \langle \vec{v}_j, \vec{v}_j \rangle \text{ by orthog.}$$

and thus $c_j = \dfrac{\langle \vec{v}_j, \vec{v} \rangle}{\langle \vec{v}_j, \vec{v}_j \rangle}$. This determines c_j for all j. $\blacksquare$

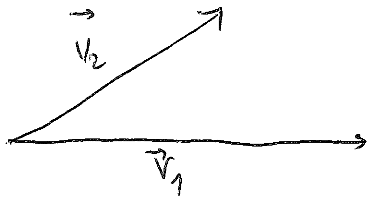
Remark. If the basis is orthonormal, we get $\vec{v} = \sum_{i=1}^{n} \langle v_i, v \rangle \vec{v}_i$.

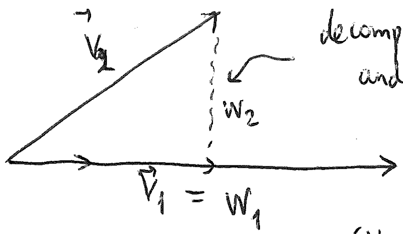
Example. Let $\vec{v}_1 = \frac{1}{\sqrt{58}}\begin{bmatrix} 3 \\ 7 \end{bmatrix}$ and $\vec{v}_2 = \frac{1}{\sqrt{58}}\begin{bmatrix} -7 \\ 3 \end{bmatrix}$.
These form an orthonormal basis of $\mathbb{R}^2$. Given some other vector like $\vec{v} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$, we get $\vec{v} = \frac{41}{\sqrt{58}}\vec{v}_1 + \frac{1}{\sqrt{58}}\vec{v}_2$.

Gram-Schmidt procedure for dot/inner products

- Starting with an arbitrary basis $\vec{v}_1, \dots, \vec{v}_n$ of V one may construct an orthogonal basis $\vec{w}_1, \dots, \vec{w}_n$ of V as follows.

Two vectors. Consider $\vec{v}_1, \vec{v}_2$ first. We need to replace them by $\vec{w}_1, \vec{w}_2$.
We keep $\vec{w}_1 = \vec{v}_1$, the first vector.





decompose $\vec{v}_2$ into two parts and keep the orthogonal part

Claim: $$\vec{v}_2 = \underbrace{\frac{\langle \vec{v}_2, \vec{w}_1\rangle}{\langle w_1, w_1\rangle}\vec{w}_1}_{\text{parallel part}} + \underbrace{\vec{v}_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1}_{\text{perpendicular part}}.$$

Check: if $u = \vec{v}_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1$

then $\langle u, w_1\rangle = \langle v_2, w_1\rangle - \frac{\langle v_2, w_1\rangle}{\cancel{\langle w_1, w_1\rangle}}\cancel{\langle w_1, w_1\rangle} = 0.$ ✓

Any number of vectors. Suppose we have $\vec{w}_1, \vec{w}_2, \dots, \vec{w}_k, \vec{u}_{k+1}, \dots, \vec{u}_n$ with the first k being orthogonal. We decompose the next vector into two parts
$$\vec{u}_{k+1} = \sum_{i=1}^{k} \frac{\langle u_{k+1}, w_i\rangle}{\langle w_i, w_i\rangle} w_i + \underbrace{\left[\vec{u}_{k+1} - \sum_{i=1}^{k}\frac{\langle u_{k+1}, w_i\rangle}{\langle w_i, w_i\rangle} w_i\right]}.$$

We claim that
$$\boxed{\vec{w}_{k+1} = \vec{u}_{k+1} - \sum_{i=1}^{k}\frac{\langle u_{k+1}, w_i\rangle}{\langle w_i, w_i\rangle} w_i}$$
is perpendicular to $\vec{w}_1, \dots, \vec{w}_k$. In fact,
$$\langle w_{k+1}, \vec{w}_j\rangle = \langle u_{k+1}, \vec{w}_j\rangle - \sum_{i=1}^{k}\frac{\langle u_{k+1}, w_i\rangle}{\langle w_i, w_i\rangle}\langle w_i, w_j\rangle$$

$= \langle u_{k+1}, w_j \rangle - \frac{\langle u_{k+1}, w_j \rangle}{\cancel{\langle w_j, w_j \rangle}} \cancel{\langle w_j, w_j \rangle}$ by orthog.

$= 0$, as needed.

Example. Let $\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\vec{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $\vec{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

This is a basis of $\mathbb{R}^3$. We obtain an orthogonal basis by letting

$$\vec{w}_1 = \vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\vec{w}_2 = \vec{u}_2 - \frac{\langle u_2, w_1 \rangle}{\langle w_1, w_1 \rangle} \vec{w}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

$$\vec{w}_3 = \vec{u}_3 - \frac{\langle u_3, w_1 \rangle}{\langle w_1, w_1 \rangle} \vec{w}_1 - \frac{\langle u_3, w_2 \rangle}{\langle w_2, w_2 \rangle} \vec{w}_2$$

$$= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1/\cancel{3}}{2/\cancel{3}} \begin{bmatrix} -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ -1/2 \\ 1/2 \end{bmatrix}$$

Example. Consider P_1 = polynomials of degree ≤ 1.

We let $\langle f, g \rangle = \int_0^1 f(x) g(x)\, dx$. The standard basis $u_1 = 1$, $u_2 = x$.

This is not orthogonal since $\langle 1, x \rangle = \int_0^1 x\, dx = \frac{1}{2} \neq 0$.

To get an orthogonal basis, we define

$$w_1 = u_1 = 1$$

$$w_2 = u_2 - \frac{\langle u_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = x - \frac{\int_0^1 x\, dx}{\int_0^1 1\, dx} = x - 1/2.$$

Remark: Gram-Schmidt ensures $\mathrm{Span}\{w_1, \ldots, w_k\} = \mathrm{Span}\{\vec{u}_1, \ldots, \vec{u}_k\}$ for each k, so the vectors $\vec{w}_i$ are nonzero.

Orthogonal bases for symmetric forms

- Gram-Schmidt works for inner products (symmetric + positive definite). In fact, one needs to divide by $\langle \vec{w}_i, \vec{w}_i \rangle$ which is positive whenever $\vec{w}_i$ is nonzero. We want to deal with symmetric forms, more generally.

Theorem (Symmetric forms) Suppose $\langle \vec{x}, \vec{y} \rangle$ is bilinear & symmetric on V. Then there exist $\vec{v}_1, \dots, \vec{v}_n$ a basis of V with $\langle \vec{v}_i, \vec{v}_j \rangle = 0 \quad \forall i \neq j$.

Proof. We use induction. If V is one-dimensional, this is fine. Suppose true for n-dimensional spaces and $\dim V = n+1$.

- We seek a vector v_1 with $\langle \vec{v}_1, \vec{v}_1 \rangle$ nonzero.

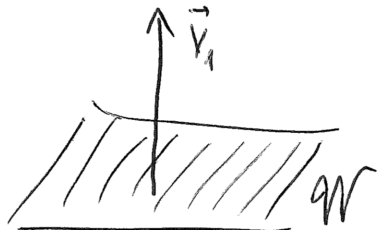
Case 1. We could have $\langle \vec{u}, \vec{v} \rangle = 0 \quad \forall \vec{u}, \vec{v}$. In that case any basis is fine!

Case 2. Suppose $\langle \vec{u}, \vec{v} \rangle \neq 0$ for some $\vec{u}, \vec{v}$. We look at

$$\underbrace{\langle \vec{u}+\vec{v}, \vec{u}+\vec{v} \rangle}_{①} = \underbrace{\langle \vec{u}, \vec{u} \rangle}_{②} + 2\langle \vec{u}, \vec{v} \rangle + \underbrace{\langle \vec{v}, \vec{v} \rangle}_{③}.$$

One of ①, ②, ③ is nonzero $\Rightarrow \exists \vec{v}_1$ with $\langle \vec{v}_1, \vec{v}_1 \rangle \neq 0$.

- Consider $\mathcal{U} = \operatorname{Span}\{\vec{v}_1\}$
and $\mathcal{W} = \mathcal{U}^\perp = \{\vec{w} \in V : \langle \vec{w}, \vec{v}_1 \rangle = 0\}$



We claim that $V = \mathcal{U} \oplus \mathcal{W}$.

If this is true, then a basis of V is $\vec{v}_1, \dots, \vec{v}_{n+1}$ with $\vec{v}_2, \dots, \vec{v}_{n+1}$ being orthogonal by induction. To check the sum is direct, note that

$$\vec{v} = \underbrace{\frac{\langle \vec{v}, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \vec{v}_1}_{\text{in } \mathcal{U}} + \underbrace{\left[\vec{v} - \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} \vec{v}_1 \right]}_{\text{in } \mathcal{W} \text{ since it's orthog. to } \vec{v}_1}.$$

Moreover, $\mathcal{U} \cap \mathcal{W} = \{0\}$ because

$$\left\{ \begin{matrix} \vec{v} \in \mathcal{U} \\ \vec{v} \in \mathcal{W} \end{matrix} \right\} \Rightarrow \left\{ \begin{matrix} \vec{v} = c\vec{v}_1 \\ \langle v, \vec{v}_1 \rangle = 0 \end{matrix} \right\} \Rightarrow c\langle v_1, v_1 \rangle = 0$$

$$\Rightarrow c = 0$$

$$\Rightarrow \vec{v} = 0. \quad \blacksquare$$

Matrix of a form

Theorem (Matrix of a form) Suppose $\langle \vec{x}, \vec{y} \rangle = \vec{x}^t A \vec{y}$ on $\mathbb{R}^n$.
With respect to the standard basis ..... matrix is A
With respect to basis $B = [\vec{v}_1 \dots \vec{v}_n]$ .... matrix is $B^t A B$.

Proof. Standard basis.... (i,j)th entry $= \langle \vec{e}_i, \vec{e}_j \rangle = \vec{e}_i^{\,t} A \vec{e}_j$
$= (i,j)$th entry of A.

Other basis .... (i,j)th entry $= \langle \vec{v}_i, \vec{v}_j \rangle = \langle B\vec{e}_i, B\vec{e}_j \rangle$

$$= (B\vec{e}_i)^t A (B\vec{e}_j)$$
$$= e_i^t \underbrace{B^t A B}\, \vec{e}_j$$
$$= (i,j)\text{th entry of } B^t A B. \quad \blacksquare$$

Corollary. Let A = real symmetric matrix. Then there exists an invertible matrix B such that $B^t A B$ = diagonal.

Proof. Look at $\langle \vec{x}, \vec{y} \rangle = \vec{x}^t A \vec{y}$. This is bilinear & symmetric.
We know $\exists$ an orthogonal basis with $\langle \vec{v}_i, \vec{v}_j \rangle = 0 \ \forall i \neq j$.
Then $(B^t A B)_{ij} = 0 \ \forall i \neq j$ so $B^t A B$ is diagonal.

Example. Let $A = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix}$. We look for eigenvectors of A.

Eigenvalues are ---- $\lambda_1 = 2$ and $\lambda_2 = 7$
Eigenvectors are ---- $v_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

These eigenvectors are orthogonal to one another!! Since $\vec{v}_1, \vec{v}_2$ are eigenvectors, we know $B = \begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix} \Rightarrow B^{-1} A B = \begin{bmatrix} 2 & \\ & 7 \end{bmatrix}$.
Since those are orthogonal

$$\langle \vec{v}_i, \vec{v}_j \rangle = \vec{v}_i^{\,t} \underbrace{A \vec{v}_j} = \vec{v}_i^{\,t} \lambda_j \vec{v}_j = 0 \quad \text{when } i \neq j.$$

In particular, $B^t A B$ = diagonal as well.

Spectral theorem If A is real & symmetric, then there exists an invertible matrix B with $B^t A B = B^{-1} A B =$ diagonal.

We'll need to know/show that the eigenvalues are real. When $\vec{x}$ is complex, the dot product $\langle \vec{x}, \vec{x} \rangle = \vec{x}^t \vec{x} = \sum_{k=1}^{n} x_k^2$ could be negative. We can obtain a dot product for $\mathbb{C}^n$ by noting that $z = a+bi \Rightarrow \bar{z} = a-bi \Rightarrow z\bar{z} = (a+bi)(a-bi) = a^2+b^2 \geq 0$. We define $\langle \vec{x}, \vec{x} \rangle = \vec{x}^* \vec{x}$ with $\vec{x}^* = \overline{x^t}$ = transpose conjugate of $\vec{x}$.

Bilinear forms

Over real vector spaces

① Linear in ~~2nd~~ 1st variable

② Linear in 2nd variable

$$\langle \vec{v}, \vec{w}_1 + \vec{w}_2 \rangle = \langle \vec{v}, \vec{w}_1 \rangle + \langle \vec{v}, \vec{w}_2 \rangle$$
$$\langle \vec{v}, c\vec{w} \rangle = c \cdot \langle \vec{v}, \vec{w} \rangle$$

③ Bilinear forms on $\mathbb{R}^n$

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^t A \vec{y}$$

④ Symmetric forms .... $\vec{x}^t A \vec{y}$

Symmetry means $A^t = A$ or $a_{ij} = a_{ji}$.

A typical example is $A = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}$.

⑤ Standard dot product on $\mathbb{R}^n$

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^t \vec{y} = \sum x_k y_k .$$

Main property:

$$\langle \vec{x}, A\vec{y} \rangle = \langle A^t \vec{x}, \vec{y} \rangle$$

when A = matrix & $\vec{x}, \vec{y}$ vectors

$$\langle A\vec{x}, \vec{y} \rangle = \langle \vec{x}, A^t \vec{y} \rangle .$$

Over complex vector spaces

① Conjugate linear in 1st variable

$$\langle \vec{v}_1 + \vec{v}_2, \vec{w} \rangle = \langle \vec{v}_1, \vec{w} \rangle + \langle \vec{v}_2, \vec{w} \rangle$$
$$\langle c\vec{v}, \vec{w} \rangle = \bar{c} \langle \vec{v}, \vec{w} \rangle$$

② Linear in 2nd variable

③ Bilinear forms on $\mathbb{C}^n$

$$\langle \vec{x}, \vec{y} \rangle = \overline{x^t} A y = \vec{x}^* A \vec{y} .$$

④ Hermitian forms ... $\vec{x}^* A \vec{y}$

Here we require $A^* = A$ or $\overline{A^t} = A$ or $\overline{a_{ij}} = a_{ji}$.

A typical example is $A = \begin{bmatrix} 2 & 3+i \\ 3-i & 4 \end{bmatrix}$.

⑤ Standard dot product on $\mathbb{C}^n$

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^* \vec{y} = \sum \bar{x}_k y_k .$$

Main property:

$$\langle \vec{x}, A\vec{y} \rangle = \langle A^* \vec{x}, \vec{y} \rangle$$

and

$$\langle A\vec{x}, \vec{y} \rangle = \langle \vec{x}, A^* \vec{y} \rangle$$

Proof of ⑤. $\langle x, Ay\rangle = \vec{x}^t A \vec{y} = \underbrace{\vec{x}^t A^{tt}} \vec{y} = (A^t\vec{x})^t \vec{y}$

$= \langle A^t\vec{x}, \vec{y}\rangle$

and similarly $\langle Ax, y\rangle = (Ax)^t y = x^t A^t y = \langle x, A^t\vec{y}\rangle$. ▨

Theorem 1 (Real eigenvalues). Suppose $A^* = A$ .... A is Hermitian.
Then all eigenvalues are real.
[This is $A^t = A$ when A is real.]

Proof. Suppose λ = eigenvalue and $\vec{v}$ = eigenvector.
We need λ = real, namely $\lambda = \bar{\lambda}$. In fact,

$$\lambda\langle\vec{v},\vec{v}\rangle = \langle v, \lambda v\rangle = \langle v, Av\rangle = \langle A^* v, v\rangle$$
$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \parallel$$
$$\bar{\lambda}\langle v,v\rangle = \langle \lambda v, v\rangle = \langle Av, v\rangle$$

so $\lambda\langle v,v\rangle = \bar{\lambda}\langle v,v\rangle$ so $\lambda = \bar{\lambda}$ and λ is real. ▨

Example Take $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$. The eigenvalues are $\lambda_1 = 0$, $\lambda_2 = 5$.
The eigenvectors are $v_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. These are orthogonal!

Theorem 2 (Orthogonal eigenvectors) Suppose $A^* = A$ once again.
Suppose $A\vec{v}_1 = \lambda_1\vec{v}_1$ and $A\vec{v}_2 = \lambda_2\vec{v}_2$ with distinct $\lambda_1 \neq \lambda_2$.
Then $\langle\vec{v}_1,\vec{v}_2\rangle = 0$, namely v_1, v_2 are orthogonal.

Proof. $\lambda_2\langle v_1, v_2\rangle = \langle v_1, \lambda_2 v_2\rangle = \langle v_1, Av_2\rangle$

$= \langle A^* v_1, v_2\rangle = \langle Av_1, v_2\rangle$

$= \langle \lambda_1 v_1, v_2\rangle = \bar{\lambda}_1\langle v_1, v_2\rangle$

$= \lambda_1\langle v_1, v_2\rangle$ by Theorem 1.

This gives $\langle v_1, v_2\rangle = 0$ or $\lambda_2 = \lambda_1$. ▨

Example. Take $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 2 \\ 1 & 2 & 1 \end{bmatrix}$. This is real & symmetric.

The eigenvalues are ---- $\lambda_1 = 0$, $\lambda_2 = 4$, $\lambda_3 = -2$

The eigenvectors are ---- $\vec{v}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$.

These form an orthogonal basis of $\mathbb{R}^3$ and we can obtain an orthonormal basis

$$w_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad w_2 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad w_3 = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}.$$

Theorem 3. (Orthonormal basis) Suppose $\vec{v}_1, \dots, \vec{v}_n$ is an orthonormal basis of $\mathbb{R}^n$. Then $B = [\vec{v}_1 \cdots \vec{v}_n]$ satisfies $B^t = B^{-1}$.

Proof. We need to check $B^t B = I_n$. In fact,

$$(B^t B)_{ij} = \vec{e}_i^{\,t} (B^t B) \vec{e}_j = \underbrace{\vec{e}_i^{\,t} B^t}\, \underbrace{B\vec{e}_j} = (B\vec{e}_i)^t (B\vec{e}_j)$$

so $(B^t B)_{ij} = v_i^t \vec{v}_j$ = dot product of $\vec{v}_i$ & $\vec{v}_j$.

Since $\vec{v}_i$ are orthonormal, we get

$$(B^t B)_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \quad \text{----} \quad \text{and so } B^t B = I_n. \quad \blacksquare$$

Example 1. Let $A = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}$. The eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = 7$.

The eigenvectors are $\vec{v}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. An orthonormal basis is

$$\vec{w}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ -1 \end{bmatrix} \quad \text{and} \quad w_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

So $B = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \Rightarrow B^t A B = B^{-1} A B = \begin{bmatrix} 2 & \\ & 7 \end{bmatrix}$

Example 2. Let $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. This is symmetric with

eigenvalues ---- $\lambda_1 = \lambda_2 = -1$ and $\lambda_3 = 2$

eigenvectors ---- $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ and $\vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Note that v_1, v_3 are orth. and $\vec{v}_2, v_3$ are orth. but $\vec{v}_1 \cdot \vec{v}_2 = 1$.

We apply Gram-Schmidt to v_1, v_2 only:

$$\vec{w}_1 = \vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \vec{w}_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1 \\ 1/2 \end{bmatrix}$$

These are still eigenvectors with eigenvalue $\lambda = -1$ since $N(A - \lambda I)$ is a subspace. We get an orthonormal basis

$$\vec{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \vec{u}_2 = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \quad \vec{u}_3 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Then $B = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & -2/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \end{bmatrix} \Rightarrow B^t A B = B^{-1} A B = \begin{bmatrix} -1 & & \\ & -1 & \\ & & 2 \end{bmatrix}$.

Orthogonal matrices

- We say B is orthogonal if $B^t = B^{-1}$ or $B^t B = I_n$.

This is true $\iff$ the columns of B are orthonormal.

Orthogonal 2×2 matrices have the form

$$B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{or} \quad B = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

The former represents rotation and the latter represents reflection.

- Multiplication by orthogonal matrices preserves dot product, angles and length. Namely, if $\vec{x}, \vec{y}$ are vectors and B is orthogonal $(n \times n)$, then $\langle B\vec{x}, B\vec{y} \rangle = (B\vec{x})^t B\vec{y} = x^t \underbrace{B^t B}\vec{y} = \langle \vec{x}, \vec{y} \rangle$

and $\|B\vec{x}\| = \sqrt{\langle Bx, Bx \rangle} = \sqrt{\langle \vec{x}, \vec{x} \rangle} = \|x\|$ and $\cos\theta = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}$ remains the same.

Application (Quadratic forms) Consider a quadratic function in n variables

$$f(x,y) = ax^2 + bxy + cy^2$$

or even $f(x_1, x_2, \dots, x_n) = \sum_{i \le j} a_{ij} x_i x_j$. This is closely related to bilinear forms

$$\langle \vec{x}, \vec{y} \rangle = \sum_{i,j} a_{ij} x_i y_j \qquad \text{and the case } \vec{x} = \vec{y}.$$

The bilinear form is $\langle \vec{x}, \vec{y} \rangle = x^t A \vec{y}$ and $\langle \vec{x}, \vec{x} \rangle = \vec{x}^t A \vec{x}$.

If we change variables $\vec{x} = B\vec{z}$ with B orthogonal, then

$$\vec{x}^t \underset{\uparrow \text{ original matrix } A}{A} \vec{x} = (B\vec{z})^t A (B\vec{z}) = \vec{z}^t \big(\underset{\uparrow \text{ final matrix } B^tAB}{B^t A B}\big) \vec{z}$$

If $B^t A B$ is diagonal, we end up with

$$\vec{z}^t \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \vec{z} = \sum_{i=1}^{2} \lambda_i z_i^2 .$$