Unit V - Inner Product Spaces

- real and complex inner products and norms
- applications in R^3
  - equations of planes and distance
  - cross product
- orthonormal sets and bases
  - orthogonal matrices
  - orthogonal diagonalization
- miscellaneous useful stuff
  - p-norms in R^n
  - positive-definite matrices
  - quadratic forms etc.

The main definition

- V is a real [finite-dimensional] vector space
- for each u,v∈V define a scalar $\langle u,v\rangle$ with the following properties:
  1. *linear:* $\langle au_1+bu_2,v\rangle = a\langle u_1,v\rangle + b\langle u_2,v\rangle$
  2. *symmetric:* $\langle u,v\rangle = \langle v,u\rangle$
  3. *positive definite:* $\langle u,u\rangle \geq 0, \text{ and } \langle u,u\rangle = 0 \text{ only if } u=0$
- $\langle u,v\rangle$ is called an *inner product* of u and v
- the vector space V with the inner product defined is called a (real) *inner product space*
  - real inner product spaces are sometimes called *Euclidean spaces*

Examples: Real inner product spaces

- the usual dot product of vectors in R^3:
  $\langle u,v\rangle = u\cdot v = (x_1,y_1,z_1)\cdot(x_2,y_2,z_2) = x_1x_2+y_1y_2+z_1z_2$
- the same dot product can be generalized to R^n
  $u=(u^1,u^2,\cdots,u^n), v=(v^1,v^2,\cdots,v^n)$
  $\langle u,v\rangle = u^1v^1+u^2v^2+\cdots+u^nv^n$
- in $R_{m,n}$ we can define ⟨A,B⟩ = Tr(B^TA)
  - the trace Tr ( ) of a square matrix is the sum of its diagonal entries

Simple examples: Real inner products

[Problem 7.2] u = (1,2,5), v = (2,-3,5), w = (4,2,-3) in R^3. Find u•v, u•w, v•w, (u+v)•w, ||u||, ||v||.

A variation: complex inner product

- V is a complex [finite-dimensional] vector space
- for each u,v∈V define a scalar $\langle u,v\rangle$ with the following properties:
  1. linear $\langle au_1+bu_2,v\rangle = a\langle u_1,v\rangle + b\langle u_2,v\rangle$
  2. *conjugate symmetric:* $\langle u,v\rangle = \overline{\langle v,u\rangle}$
  3. positive definite $\langle u,u\rangle \geq 0, \text{ and } \langle u,u\rangle = 0 \text{ only if } u=0$
- $\langle u,v\rangle$ is called a complex *inner product* of u and v
- the vector space V with the inner product defined is called a (complex) *inner product space*
  - complex inner product spaces are sometimes called *unitary spaces*

Examples: Complex inner product spaces

- in C^3 we have to use the conjugate in the definition of the usual dot product:
  $\langle u,v\rangle = u\cdot v = (x_1,y_1,z_1)\cdot(x_2,y_2,z_2) = x_1\overline{x_2}+y_1\overline{y_2}+z_1\overline{z_2}$
- the same dot product can be generalized to C^n
  $u=(u^1,u^2,\cdots,u^n), v=(v^1,v^2,\cdots,v^n)$
  $\langle u,v\rangle = u^1\overline{v^1}+u^2\overline{v^2}+\cdots+u^n\overline{v^n}$
- in $C_{m,n}$ we can define ⟨A,B⟩ = Tr(B^*A)
  - the conjugate transpose of a complex matrix is $B^* = \overline{B}^T$

Inner products

- in the second position < , > is linear for real spaces

$$\langle u, av_1 + bv_2 \rangle = a\langle u, v_1 \rangle + b\langle u, v_2 \rangle$$

- .....and conjugate linear for complex spaces

$$\langle u, av_1 + bv_2 \rangle = \overline{a}\langle u, v_1 \rangle + \overline{b}\langle u, v_2 \rangle$$

- if u and v are column vectors in R^n the inner product can be written as a matrix product $\langle u,v \rangle = u^T v$
  - the matrix product AB can actually be defined this way as inner products of rows of A with columns of B
- the complex congugate is necessary for ⟨u,u⟩ to be a real number so we can define the length of a vector......

Definition: Norm of a vector

- this definition applies to both real and complex inner product spaces
- ⟨u,u⟩ is a non-negative real number so we can define $\|u\| = \sqrt{\langle u, u \rangle}$
- ||u|| is the *norm* of the vector u
  - this norm is associated with the given inner product
- a vector for which ||u|| = 1 is called a *unit vector*
- any non-zero vector u can be *normalized* to a unit vector 'in the same direction' $\hat{u} = (1/\|u\|)u$

Simple examples: Complex inner products

[Problem 7.48] Suppose ⟨u,v⟩ = 3+2i. Find ⟨(2-4i)u, v⟩, ⟨u, (3+4i)v⟩, and ⟨(3-6i)u, (5-2i)v⟩.

Examples: Function spaces

- the *standard* inner product on the space of continuous real-valued functions on [a,b]:

$$\langle f, g \rangle = \int_a^b f(t)g(t)dt$$

- if the functions are complex-valued you have to use the conjugate as usual for things to work:

$$\langle f, g \rangle = \int_a^b f(t)\overline{g(t)}dt$$

Examples: Function spaces

[Problem 7.5] Find ⟨f,g⟩, ⟨f,h⟩, ||f||, ||g|| and normalize g, with $f(t)=t+2$, $g(t)=3t-2$, and $h(t)= t^2-2t-3$. The inner product and norm are defined on the interval [0,1].

Norms

- the zero vector is the only vector with norm 0

  ||0|| = ⟨0,0⟩ = ⟨0v,0⟩ = 0⟨v,0⟩ = 0
- ||u-v|| = d(u,v) ≥ 0 is called the *distance* between vectors u and v
- ||ku|| = |k| ||u||
  - note the |k| means the modulus in complex spaces, or absolute value in real spaces
- triangle inequality ||u+v|| ≤ ||u|| + ||v||
  - the proof of this last one uses a very important result....

The Cauchy-Schwartz inequality

- for two vectors in a complex inner product space

$$|\langle u, v\rangle| \leq ||u||\ ||v||$$

- we can show this by evaluating $||u - \langle u, \hat{v}\rangle \hat{v}||^2$ where $\hat{v} = v/||v||$ is a unit vector along v
  - this is a trick motivated by geometry (see projections later)
  - evaluate this norm using the inner product
  - simplify using facts about complex numbers
  - apply the fact that the norm above is non-negative to get the desired inequality

Cauchy-Schwartz inequality proof

Angle

- the C-S inequality allows us to define the *angle* between any two vectors u,v using

$$\cos\theta = \frac{\langle u, v\rangle}{||u||\,||v||}$$

- for instance in R^4 with u=(1,2,3,4), v = (-1,0,-2,2)
  $||u||^2 = 1+4+9+16 = 30$, $||v||^2 = 1+4+4 = 9$, and
  $\langle u,v\rangle = -1-6+8 = 1$ so $\cos\theta = 1/(3\sqrt{30})$
- or... in function spaces, e.g. f and g from problem 7.5 on slide 11
  $\langle f,g \rangle = -1$, $|| f || = \sqrt{57}$, $|| g || = 1$ so $\cos\theta = -1/(\sqrt{57})$
- the case $\cos\theta = 0$ is particularly important.....

Orthogonality

- two vectors u,v in an inner product space are *orthogonal* if $\langle u,v\rangle = 0$
- only the zero vector is orthogonal to all vectors v:
  $\langle 0,v\rangle = \langle 0v,v\rangle = 0\langle v,v\rangle = 0$
  $\langle u,v\rangle = 0$ all u implies $\langle u,u\rangle = 0$ so u = 0
- orthogonality is symmetric...it's a property of a pair of vectors
  - since $\langle u,v\rangle = \langle v,u\rangle = 0$
- the concept agrees with the geometric idea of 'perpendicularity' since $\cos\theta = 0$ so $\theta = \pi/2$
- extending these geometric concepts to <u>any</u> inner product space is a powerful technique
  - e.g. allows consideration of orthogonal functions

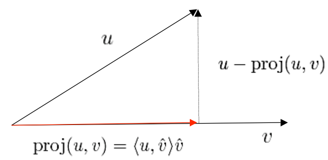
Examples: orthogonality

[example 7.6] (a) u=(1,1,1), v=(1,2,-3) (b) $\langle \sin t, \cos t\rangle$ in $C[-\pi,\pi]$.

Examples: orthogonality

[example 7.7] Find a non-zero unit vector orthogonal to both u=(1,2,3) and v=(2,5,4).

Orthogonal projection



- the *orthogonal projection* of u on v is the vector

$$\langle u, \hat{v} \rangle \hat{v} = \langle u, \frac{v}{\|v\|} \rangle \frac{v}{\|v\|}$$

- C-S says that ||proj(u,v)|| ≤ ||u||
  - the proof is based on using ||u-proj(u,v)|| ≥ 0

Orthogonal projection

- the scalar $\langle u, \hat{v} \rangle$ is called the *component* of u along v
  - in real ips this may be a positive or negative value
  - in complex ips this may have any complex value
- two vectors u,v are
  - orthogonal if and only if proj(u,v) = 0
  - linearly dependent if and only if proj(u,v) = u
- we will generalize these concepts to projections onto a subspace, but for now an illustration with a single vector...

Example: projections

[problem 1.12] For u=(1,-3,4), v=(3,4,7) find (a) cos θ (b) proj(u,v) and (c) d(u,v).

Orthogonal complement

- V is an ips, S is a subset of vectors
- the *orthogonal complement* of S ['S perp'] is

$$S^{\perp} = \{v \in V | \langle v, w \rangle = 0 \text{ for all } w \in S\}$$

- $S^{\perp}$ is a subspace of V:
  - w∈S, u,v∈$S^{\perp}$, and a,b scalars then
  - ⟨au+bv,w⟩ = a⟨u,w⟩ + b⟨v,w⟩ =a(0) + b(0) = 0
  - so au+bv∈$S^{\perp}$
- if S = {w} we can write $w^{\perp}$ instead of $S^{\perp}$ for simplicity
- if W is supspace of V then
  - $W \cap W^{\perp} = \{0\}$
  - $(W^{\perp})^{\perp} = W$

Example: orthogonal complement

[problem 7.11] Find a basis for $W^{\perp}$ if W is the subspace of R^5 spanned by u=(1,2,3,-1,2), v=(2,4,7,2,-1).

Rowspace and nullspace again

The orthogonal complement of the row space R of a matrix is the nullspace N

- the previous example illustrates this
- given a basis for R you can find the nullspace N by finding the orthogonal complement $R^{\perp}$
- given a basis for the nullspace N you can find R by finding the orthogonal complement $N^{\perp}$

Applications in R^3: equation of a plane

- $w \in R^3$ is a nonzero vector
- $w^\perp$ is the plane P through the origin and perpendicular to w
- w is called a *normal* vector of P
- in general we can find the equation of a plane P in R^3 if we have
  - a normal vector n and...
  - any point (x_0, y_0, z_0) on P
- let (x,y,z) be the position vector of a point in P
- then the vector $(x-x_0, y-y_0, z-z_0)$ lies in the plane
- so the required equation is

$$n \cdot (x-x_0, y-y_0, z-z_0) = 0$$

Example: hyperplane in R^n

The previous result applies just the same to hyperplanes in R^n. [example 1.6b] Find an equation of the hyperplane in R^4 with normal vector n = (4,-2,5,6) and passing through (1,3,-4,2).

Digression: Cross product in R^3

- how do you get a normal vector if you are given two vectors lying in a plane in R^3, or equivalently the position vectors of three points on the plane?
- use the cross product...but this only applies to R^3
- $u,v \in R^3$ and define the vector $u \times v \in R^3$ by

$$u \times v = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ u^1 & u^2 & u^3 \\ v^1 & v^2 & v^3 \end{vmatrix}$$

  - **u×v is orthogonal to both u,v** [easy to show]
  - the direction of u×v is uniquely defined by the 'right hand thread rule'
  - $||u \times v|| = ||u||\ ||v|| \sin \theta$

Example: cross product

Show that u×v is orthogonal to both u and v by evaluating u·u×v and v·u×v [is this notation ambiguous?].

Some properties of cross product

- u×v = - v×u
- u×(v+w) = u×v + u×w
- k(u×v) = (ku)×v = u×(kv)
- u×**0** = **0**
- u×u = **0**
- u×(v×w) ≠ (u×v)×w
- u×(v×w) = (u·w)v - (u·v)w
  - the vector lies in the plane determined by w and v

Examples: cross product

[problem1.25] for u=(2,-3,4), v=(3,1,-2), w=(1,5,3). Find u×v and u×w. Extra: find u·v×w, u×(v×w) and (u×v)×w.

Equation of a plane in R^3

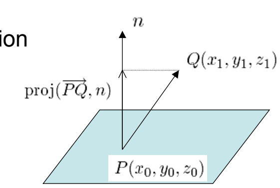
- to return to the equation of a plane P in R^3....
- we had the basic point (r_0) normal vector (n) form, in vector notation: $n\cdot (r-r_0) = 0$
- given instead (1) two vectors lying in P you can find the normal vector using the cross product or....
- given (2) the position vectors of three points on P you can subtract in pairs to get two vectors lying in P and proceed as in (1)
- the equation of a plane in R^3 is always of the form ax+by+cz+d = 0, where n = (a,b,c) is the normal vector

Example: equation of a plane in R^3

Find the equation of the plane through the points u=(2,-3,4), v=(3,1,-2), w=(1,5,3).

Application: point to plane distance in R^3

- let point P with position vector $r_0=(x_o,y_0,z_0)$ & vector n=(a,b,c) define the equation of a plane
- choose a point Q with position vector $r_1=(x_1,y_1,z_1)$
- we want to calculate the distance D from Q to the plane
- as shown D can be found using a projection



Point to plane distance in R^3

- write the plane equation as ax+by+cz+d = 0
- calculate the required length D:

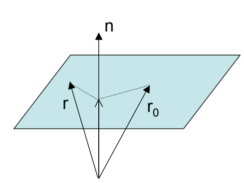
$$\begin{aligned} D &= \left|\overrightarrow{PQ}\cdot \hat{n}\right| \\ &= \frac{|(x_1-x_0, y_1-y_0, z_1-z_0)\cdot(a,b,c)|}{\sqrt{a^2+b^2+c^2}} \\ &= \frac{|a(x_1-x_0)+b(y_1-y_0)+c(z_1-z_0)|}{\sqrt{a^2+b^2+c^2}} \\ &= \frac{|ax_1+by_1+cz_1-(ax_0+by_0+cz_0)|}{\sqrt{a^2+b^2+c^2}} \\ &= \frac{|ax_1+by_1+cz_1+d|}{\sqrt{a^2+b^2+c^2}} \end{aligned}$$

Example: point to plane distance in R^3

Find the distances: (a) from the point (1,-4,-3) to the plane with equation 2x-3y+6z = -1 and (b) between this plane and the plane with equation 4x-6y+12z = 13.

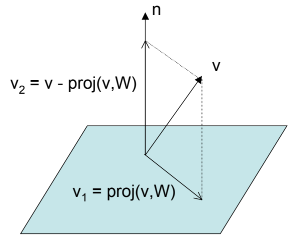
Another geometric view of planes

- another way of viewing a vector equation of a plane is $n\cdot r = n\cdot r_0$ or...
- $r\cdot n = r_0\cdot n$ or...
- $\text{proj}(r,n) = \text{proj}(r_0,n)$ which makes sense geometrically...



Orthogonal projection onto a subspace

- we know how to find the orthogonal projection of a vector onto a line
- what about onto a general subspace? say a plane...
- v is decomposed into a sum of two orthogonal vectors:
  $v = v_1 + v_2$
- $v_1 \in W$ and $v_2 \in W^\perp$
- geometrically we can see that the decomposition is unique



Orthogonal projection onto a subspace

- in general let W is a subspace of an ips V
- **then V can be decomposed as the direct sum $V = W \oplus W^\perp$**
  - this is an important result [see problem 7.28]
  - for instance if A is a matrix transformation V→U the decomposition is V = rowspace A ⊕ nullspace A
- any vector $v \in V$ can be decomposed <u>uniquely</u> as $v = v_1 + v_2$ with $v_1 \in W$ and $v_2 \in W^\perp$
- question: for a general subspace W how do we calculate proj(v,W)?
- to answer this we need the concept of orthogonal sets and bases....

Orthogonal sets

- a set of non-zero vectors $\{u_1, u_2, ..., u_n\}$ is an *orthogonal set* if $\langle u_i,u_j\rangle = 0$ for $i \neq j$
- if, in addition, $\langle u_i,u_i\rangle = 1$ it is an *orthonormal set* of vectors
- **orthogonal sets are linearly independent**
  - an important result [see problem 7.15]
- examples:
  - the standard basis vectors are an orthonormal basis for R^n
  - the vectors {1, cos t, cos 2t, ..., sin t, sin 2t, ...} are an orthogonal set in $C[-\pi,\pi]$....used in Fourier analysis
  - the basis {(1,2), (1,-1)} for R^2 is <u>not</u> an orthogonal basis

Orthogonal bases

- let $\{u_1, u_1, ..., u_n\}$ be an <u>orthogonal</u> basis for V
- express a vector v in terms of this basis:
  $v = a_1u_1 + a_2u_2 + ... + a_nu_n$
- now calculate $\langle v,u_i\rangle = \langle a_iu_i,u_i\rangle = a_i\langle u_i,u_i\rangle$
- so we can solve for the scalars a_i called *Fourier coefficients* of v with respect to the $\{u_i\}$ basis
  $$a_i = \frac{\langle v, u_i\rangle}{\langle u_i, u_i\rangle}$$
- then the decomposition of v is
  $$v = \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle}u_1 + \frac{\langle v, u_2\rangle}{\langle u_2, u_2\rangle}u_2 + \cdots + \frac{\langle v, u_n\rangle}{\langle u_n, u_n\rangle}u_n$$
- or
  $$v = \text{proj}(v, u_1) + \text{proj}(v, u_2) + \cdots + \text{proj}(v, u_n)$$

Example: orthogonal bases

[problem 7.13] $S=\{u_1,u_2,u_3,u_4\}$ where $u_1=(1,1,0,-1)$, $u_2=(1,2,1,3)$, $u_3=(1,1,-9,2)$, $u_4=(16,-13,1,3)$ is an orthogonal basis. Find the coordinates of a general vector (a,b,c,d) with respect to S.

....Example: orthogonal bases

Orthogonal projection onto a subspace

- we can now answer the question asked on slide 38...how do you calculate proj(v,W)?

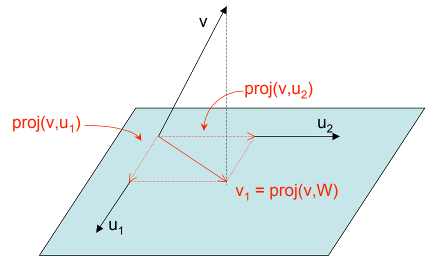
- let $\{u_1, u_2, \ldots, u_r\}$ be an orthogonal basis for W
- then we can define proj(v,W) by

$$\mathrm{proj}(v,W) = \mathrm{proj}(v,u_1) + \mathrm{proj}(v,u_2) + \cdots + \mathrm{proj}(v,u_r)$$

- this **o n l y** works because the basis is orthogonal

Projection onto a subspace

- illlustrating for the plane...



Example: projection onto a subspace

[problem 7.23a] Find the projection of v = (1,3,5,7) onto the subspace W spanned by $\{u_1,u_2\}$ where u_1=(1,1,1,1), u_2=(1,-3,4,-2).

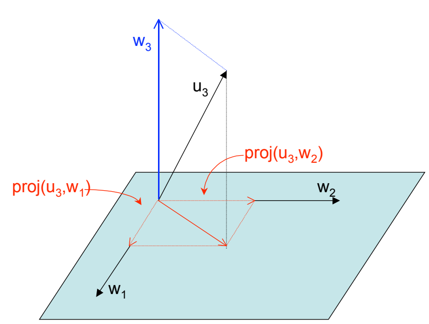
Orthonormal bases

- if $\{\hat{u}_1, \hat{u}_2, \cdots, \hat{u}_n\}$ is an orthonormal basis things can be written more neatly....
- we get simply $a_i = \langle v, \hat{u}_i \rangle$
- and $v = \langle v, \hat{u}_1 \rangle \hat{u}_1 + \langle v, \hat{u}_2 \rangle \hat{u}_2 + \cdots + \langle v, \hat{u}_n \rangle \hat{u}_n$
  - which, of course, gives exactly the same formula when written as projections
- the last question to ask is how to proceed if you have a basis but it isn't orthogonal?
- there is a simple procedure to make it orthogonal....

Gramm-Schmidt algorithm

- this procedure converts any old basis $\{u_1,u_2,\ldots,u_n\}$ of V into an orthogonal basis
- define
  - $w_1 = u_1$
  - $w_2 = u_2 - \mathrm{proj}(u_2,w_1)$
  - $w_3 = u_3 - \mathrm{proj}(u_3,w_1) - \mathrm{proj}(u_3,w_2)$
  - .....and so on
- then $\{w_1,w_2,\ldots,w_n\}$ is an orthogonal basis of V
- the vectors w_i can easily be normalized to form an orthonormal basis of V

Why Gramm-Schmidt works: geometrically



Why Gramm-Schmidt works: algebraically

- let $\{u_1,u_2,...,u_r\}$ be an orthogonal set, v a vector
- use the coefficients on slide 40: $a_i = \frac{\langle v, u_i\rangle}{\langle u_i, u_i\rangle}$
- define $v' = v - a_1u_1 - a_2u_2 - ... - a_ru_r$
- then

$$\begin{aligned} \langle v', u_i\rangle &= \langle v - a_1u_1 - a_2u_2 - \cdots - a_ru_r, u_i\rangle \\ &= \langle v, u_i\rangle - a_i\langle u_i, u_i\rangle \\ &= \langle v, u_i\rangle - \frac{\langle v, u_i\rangle}{\langle u_i, u_i\rangle}\langle u_i, u_i\rangle \\ &= \langle v, u_i\rangle - \langle v, u_i\rangle \\ &= 0 \end{aligned}$$

- so v' is orthogonal to ALL the u_i vectors

Example: Gramm-Schmidt algorithm

[example 7.10] Find an orthonormal basis for the subspace W of R^4 spanned by $\{u_1,u_2,u_3\}$ where $u_1=(1,1,1,1)$, $u_2=(1,2,4,5)$ $u_3=(1,-3,-4,-2)$.

....Example: Gramm-Schmidt algorithm

Gramm-Schmidt in practice

- G-S is not a very robust numerical algorithm, but it's valuable for important theoretical reasons
  - e.g. construction of the Legendre polynomials in the next example
- you won't usually have more than three basis vectors because the calculations can be tedious
- to avoid messy arithmetic clear all fractions as you make your choice for each w_i vector
  - any multiple of each w_i will do the job just as well
- normalize all the w_i vectors at the very last step once you have the orthogonal basis
- the procedure works on any linearly independent set
  - it gives an orthogonal basis for the subspace it spans

Example: constructing Legendre polynomials

[example 7.11] $P_3(t)$ polynomial space with $\langle f,g\rangle$ defined on the interval [-1,1]. Apply G-S to the mononomial basis to find an orthogonal basis.

.... Example: constructing Legendre polynomials

Example: projection onto a subspace

[problem 7.23b] Find the projection of v = (1,3,5,7) onto the subspace W spanned by $\{u_1,u_2\}$ where u_1=(1,1,1,1), u_2=(1,2,3,2).

....Example: projection onto a subspace

Example: orthogonal basis

[problem 7.12] Find an orthogonal basis for $w^\perp$ where w=(1,2,3,1).
...an alternative approach without using G-S.

....Example: orthogonal basis

Orthogonal matrices

- a real n×n matrix P is *orthogonal* if $P^{-1} = P^T$
- this gives $P^TP = PP^T = I$
- if (and only if) we use the <u>standard</u> i.p. ⟨ , ⟩ on R^n the following are equivalent
  - P is orthogonal
  - the rows of P are an orthonormal basis of R^n
  - the columns of P are an orthonormal basis of R^n
- examples:

$$\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \end{bmatrix} \qquad \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

Example: orthogonal matrices

[problem 7.32] Find an orthogonal matrix whose first row is u_1=(1/3,2/3,2/3).

Orthogonal matrices

- a matrix P is a change of basis matrix between two orthonormal bases if and only if it is orthogonal [see problem 7.37]
- a more accurate [better] term for 'orthogonal matrix' would be 'orthonormal matrix' but the other is traditional
- an important application involves orthogonal diagonalization of a matrix [or operator]
- when can we find an orthogonal P so that P^TAP is diagonal?

Diagonalization of real symmetric matrices

- if A is a real symmetric matrix then:
  - A has only real eigenvalues
  - the eigenvectors of A corresponding to distinct eigenvalues are orthogonal
- these results give a method to orthogonally diagonalize any real symmetric matrix
  1. find the eigenvalues and an eigenvector for each non-repeated eigenvalue
  2. for any repeated e.val. find an orthogonal basis of eigenvectors for its eigenspace
  3. normalize the whole set of eigenvectors
  4. write the normalized eigenvectors as the columns of P

Example: orthogonal diagonalization

[problem 9.25] Orthogonally diagonalize $B = \begin{bmatrix} 11 & -8 & 4 \\ -8 & -1 & -2 \\ 4 & -2 & -4 \end{bmatrix}$

....Example: orthogonal diagonalization

Quadratic forms

- a *quadratic form* in n real variables $x_1, \ldots, x_n$ is an expression of type $q(x_1, \ldots, x_n) = \sum a_{ij}x_ix_j$
  - e.g. $2x^2 - 4y^2 + 3xy + yz$ is a quadratic form in three variables x,y,z
- we'll write $u = (x_1, \ldots, x_n)$ when needed so the quadratic form is q(u)
  - or sometimes just q(x,y,z)
- the terms of type $a_{ij}x_ix_j$ with $i \neq j$ are called *cross-product* terms
- a quadratic form can be represented using a matrix product u^TAu with a symmetric matrix A....

Quadratic forms

- $$\begin{aligned} q(x,y,z) &= a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + a_{12}xy + a_{13}xz + a_{23}yz \\ &= \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a_{11} & a_{12}/2 & a_{13}/2 \\ a_{12}/2 & a_{22} & a_{23}/2 \\ a_{13}/2 & a_{23}/2 & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \end{aligned}$$
- a quadratic form is *diagonal* if it is a sum of squares
  - i.e. there are no cross-terms so....
  - the matrix A is diagonal
- recall [slide 7] that an inner product ⟨u,v⟩ can be written using column vectors as u^Tv
- so a quadratic form q(u) can also be associated with the standard inner product in R^n
  - $q(u) = u^TAu = u^T(Au) = \langle u,Au \rangle = \langle Au,u \rangle$
  - A must be a real symmetric matrix

Diagonalizing a quadratic form

- the quadratic form $q(u) = u^TAu$ can be orthogonally diagonalized by a change of variables
  - find a P which orthogonally diagonalizes A to D
  - put $u = Pv$
  - then $u^TAu = (Pv)^TA(Pv) = v^TP^TAPv = v^TDv$

Example: diagonalizing a quadratic form

[problem 9.26] Find an orthogonal substitution which diagonalizes $q(x,y) = x^2 + 6xy - 7y^2$.

...Example: quadratic form

Transforming a conic to standard form

- a quadratic equation in R^2 is an expression of the form $ax^2 + 2bxy + cy^2 + dx + ey + f = 0$
  - the associated quadratic form is $q(x,y) = ax^2 + 2bxy + cy^2$
  - the cross-term represents a rotation away from standard form...
  - the principal axes are not the xy axes
  - principal axes are given by the e.vecs of the A matrix
  - the linear terms represent a translation away from the origin
- orthogonal diagonalization of the form provides a change of variables that rotates the conic to standard axes
  - a translation (complete the square) can then put it into standard position
- the same can be done for *quadric surfaces* in R^3

Positive definite matrix

- if A is a symmetric $n \times n$ matrix then
  $\lambda_1 \geq u^TAu \geq \lambda_n$ where
  - λ_1 is the largest e.val. of A
  - λ_n is the smallest e.val. of A
  - u is constrained so $||u|| = 1$ in the usual inner product of R^n
- $u^TAu = \lambda_k$ if u is an e.vec. of A associated with λ_k
- a symmetric matrix A is *positive-definite* if $q(u) = u^TAu > 0$ for all non-zero vectors u
- so.....a symmetric matrix is positive definite if and only if all the e.vals. are positive
  - there are other tests for positive definiteness
  - **the property is very important in both physical and numerical applications**

A positive definite matrix has ...

- all positive entries on the main diagonal
  - to show: apply v^TAv with the standard basis vectors
- the largest entry (in abs val.) on the main diagonal
- det (A) > 0 so it is always invertible
- a unique square root matrix B so that $B^2 = A$

Diagonally dominant matrices

- A is *diagonally dominant* if:
  $|a_{ii}| > \sum |a_{ij}|$, $i \neq j$, $i = 1, \ldots n$
- a diagonally dominant matrix is positive definite if it is:
  - symmetric and ....
  - has all main diagonal entries positive
- ...but the converse is false
  - there are positive definite matrices that are not diagonally dominant [find one]
  - there are also positive definite matrices that are diagonally dominant and not symmetric [any one with all positive eigenvalues]

Symmetric positive definite matrices

- symmetric positive definite matrices appear in many applications:
  - solution of partial differential equations ... heat conduction, mass diffusion etc (Poisson and Laplace equations)
  - analysis of stress
  - linear regression models
  - optimization problems

Norms

- the axioms required to define a useful *norm* (length) in a vector space V are (u,v∈V, k scalar):
  1. $||v|| \geq 0$ and $||v|| = 0$ if an only if $v = 0$
  2. $||kv|| = |k|\ ||v||$
  3. $||u + v|| \leq ||u|| + ||v||$
- $d(u,v) = ||u - v||$ is then defined as the *distance* between u and v
- you've seen how a norm can be associated with an inner product by defining $||u||^2 = \langle u,u \rangle$
- but not all useful norms are obtained this way.....

p-norms in R^n

- for $u=(a_1,\ldots,a_n) \in R^n$ [or C^n] we can define the *p-norm* $\|u\|_p = (|a_1|^p + \cdots + |a_n|^p)^{1/p}$
- interesting special cases:

$$\begin{aligned} \|u\|_1 &= |a_1| + \cdots + |a_n| \\ \|u\|_2 &= \sqrt{|a_1|^2 + \cdots + |a_n|^2} \\ \|u\|_\infty &= \max(|a_1|, \cdots, |a_n|) \end{aligned}$$

- the 2-norm is associated with the standard innner product in R^n
- distance in R^n can be measured using any of these p-norms
  - distances and lengths of vectors will, of course, be different for each p-norm

Unit p-norm sets in R^2

- consider R^2 and define S_p to be all vectors of unit p-length i.e. $||u||_p = 1$