Last time: ORTHOGONAL / UNITARY DIAGONALIZATION

$A = A^T$; $P^{-1} = P^T$

A real symmetric has an orthogonal diagonalization
i.e. there are P orthogonal, D diagonal with $A = PDP^{-1}$.

$A = \bar{A}^T = A^*$; $P^{-1} = \bar{P}^T = P^*$

A complex Hermitian has a unitary diagonalization
i.e. there are P unitary, D diagonal with $A = PDP^{-1}$.

In both cases: A has an orthonormal basis of eigenvectors.

QUADRATIC FORMS

Expressions $q(x_1, \dots, x_n)$ in which all the terms have the form $k x_i x_j$ (where possibly $i=j$)

If $i \neq j$ we call a "cross product term"

e.g. $x_1^2 - 6x_2^2 + 3x_1 x_2$

$2x_1^2 + 3x_2^2 - 5x_3^2 + x_1x_2 + 2x_1x_3 - 4x_3x_2$

$3x^2 + 2y^2 - z^2 + xy - xz + yz$

Typically though write terms $k x_i x_j$ with $i \leq j$ & gather $x_i x_j$ terms with $x_j x_i$ terms

We can write quadratic forms $q(x)$ in matrix notation

$$x^T A x \text{ where}$$

A is symmetric.

e.g. $2x^2 + 5y^2 + 3z^2 - 2xy + 4yz - 7xz$

can be written as $(x, y, z)\begin{pmatrix} & & \\ & & \\ & & \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$

$\uparrow$

Q: how to fill up entries of A?

Well A symmetric so $A = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$

& multiplying out $x^T A x$ we get

$$ax^2 + dy^2 + fz^2 + bxy + byx + cxz + czx$$
$$+ eyz + ezy$$

$$= ax^2 + dy^2 + fz^2 + 2bxy + 2cxz + 2eyz$$

So for our example $2x^2 + 5y^2 + 3z^2 - 2xy + 4yz - 7xz$

this can be written as $x^T \begin{pmatrix} 2 & -1 & 2 \\ -1 & 5 & -7/2 \\ 2 & -7/2 & 3 \end{pmatrix} x$

circled entries should be swapped

We can neaten things up and eliminate the cross product terms by making an <u>ORTHOGONAL change of variables</u>.

A symmetric so we have an orthogonal diagonalization $A = PDP^T$

We set $\underset{(y_1,\dots,y_n)}{y} = P^T \underset{(x_1,\dots,x_n)}{x}$. Then $x^T A x =$

$$(Py)^T A (Py)$$
$$= y^T P^T (PDP^T) Py$$
$$= y^T D y$$

D diagonal so $y^T D y$ has no cross product terms; in fact $y^T D y = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2$ for $\lambda_1, \dots, \lambda_n$ the eigenvalues of A.

Conic Sections in $\mathbb{R}^2$

$$0 = a_1 x^2 + 2a_2 x \cancel{x_2} y + a_3 y^2 + b_1 x + b_2 y + c$$

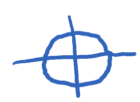
at least one of a_1, a_2, a_3 non-zero

Suppose our conic is centred at 0. Then $b_1 = b_2 = 0$.
(To work out where the centre is if b_1, b_2 not zero: complete the square.)

So we have $\underbrace{a_1 x^2 + 2\cancel{a_1} a_2 xy + a_3 y^2}_{\text{quadratic form } x^T \begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix} x} = -c$

<u>Example</u>

- $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ ← ellipse; circle if $a = b$



- $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$; $\dfrac{y^2}{b^2} - \dfrac{x^2}{a^2} = 1$ ← hyperbola

(• Important example of a conic section not centred at O is $y = kx^2$ or $x = ky^2$)

← parabola

The examples here are in "standard position"
– no cross product term.

If there is a cross product term it means that the ellipse/hyperbola is rotated (about the origin). How much?

Well, make an orthogonal change of variables which will eliminate the cross product term — has effect of rotating the conic back into standard position. — then we really see what we're dealing with!

Example Let's say we have conic section

$$30x_1^2 + 6x_2^2 - 18x_1x_2 - 3 = 0.$$

What kind of conic is this?

Solution First rearrange & set right-hand side to 1:

$$\underbrace{\overset{}{\cancel{30}}_{10}x_1^2 + \cancel{6}_{2}x_2^2 - \cancel{18}_{6}x_1x_2}_{} = \cancel{3}\ 1$$

Then put this quadratic form into matrix notation x^TAx: $A = \begin{pmatrix} 10 & -3 \\ -3 & 2 \end{pmatrix}$.

We want $x^TAx = y^TDy = \lambda_1 y_1^2 + \lambda_2 y_2^2$,

λ_1, λ_2 eigenvalues of A.

So find eigenvalues: solve $0 = \det(A - \lambda I)$

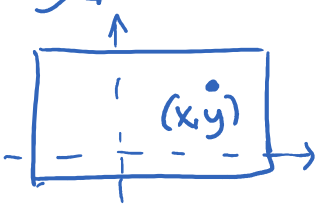
$$= \det\begin{pmatrix} 10-\lambda & -3 \\ -3 & 2-\lambda \end{pmatrix}$$

$$\begin{aligned} &= (10-\lambda)(2-\lambda) - 9 \\ &= 20 - 12\lambda + \lambda^2 - 9 \\ &= \lambda^2 - 12\lambda + 11 \\ &= (\lambda - 11)(\lambda - 1) \end{aligned}$$

So the eigenvalues are 11 and 1 & we can find an orthogonal change of coordinates from $x = (x_1, x_2)$ to $y = (y_1, y_2)$ with $11y_1^2 + y_2^2 = 1$

This is an ellipse. $\longrightarrow$ $\left(\text{i.e. } \dfrac{y_1^2}{(1/\sqrt{11})^2} + y_2^2 = 1 \right)$

Optimization of quadratic forms



Suppose you have a metal plate on which the temperature at (x,y) is given by $q(x,y) = 5x^2 - 5y^2 + 24xy$

The question is, what are the maximum & minimum values of the temperature in some region of the plate?

The answer is given (partly) by:

Constrained Extremum Theorem

Suppose A is a symmetric matrix and $q(x) = x^T A x$.
Then at the point x

$$\underbrace{\lambda_{min}}_{\text{smallest eigenvalue of } A} \|x\|^2 \leq q(x) \leq \underbrace{\lambda_{max}}_{\text{biggest eigenvalue of } A} \|x\|^2 .$$

Proof Idea

$$\begin{aligned} x^T A x &= \lambda_1^{\cancel{2}} y_1^2 + \dots + \lambda_n^{\cancel{2}} y_n^2 \\ &\leq \lambda_{max}^{\cancel{2}} (y_1^2 + \dots + y_n^2) \\ &= \lambda_{max} \|y\|^2 = \lambda_{max} \|x\|^2 \end{aligned}$$

(other direction similar.)

because change of variables is orthogonal. //

Continuing the Theorem:
These extrema ($\lambda_{min} \|x\|^2$ and $\lambda_{max} \|x\|^2$) are achieved when x is an eigenvector corresponding to the relevant eigenvalue.

Back to the heated plate example:

Suppose you want to know the maximum & minimum temperature on the unit circle (i.e. $\|x\|^2 = 1$ or $\|x\| = 1$)

"subject to the constraint that $\|x\|^2 = 1$."

Max. & Min. values are $\lambda_{max} \|x\|^2 = \lambda_{max}$

$$\lambda_{min} \|x\|^2 = \lambda_{min}$$

So find eigenvalues!

We need to find λ_{max} & λ_{min} for A where

$x^T A x = 5x^2 - 5y^2 + 24xy$. i.e. $A = \begin{pmatrix} 5 & 12 \\ 12 & -5 \end{pmatrix}$

Solve $\det(A - \lambda I) = \det\begin{pmatrix} 5-\lambda & 12 \\ 12 & -5-\lambda \end{pmatrix}$

$$= (5-\lambda)(-5-\lambda) - 144$$
$$= -25 + \lambda^2 - 144$$
$$= \lambda^2 - 169 = (\lambda - 13)(\lambda + 13)$$

So eigenvalues are $-13, 13$

λ_{min} ↗ ↖ λ_{max}

So max temp $= 13$
min temp $= -13$

Q: At which points on the unit circle are the max. & min. temperatures achieved?

Solution Find eigenvectors corresponding to λ_{max} and λ_{min} on the unit circle i.e. norm 1 / unit length.

An eigenvector for $\lambda_{min} = -13$ is $\begin{pmatrix} -2/3 \\ 1 \end{pmatrix}$

$\longrightarrow$ Normalize (need unit eigenvector) & we get $\begin{pmatrix} -2/\sqrt{13} \\ 3/\sqrt{13} \end{pmatrix}$.

So min. temperature at $(x,y) = \left(-\frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}}\right)$.

An eigenvector for $\lambda_{max} = 13$ is $\begin{pmatrix} 3/2 \\ 1 \end{pmatrix}$

$\longrightarrow$ Normalize & get $\begin{pmatrix} 3/\sqrt{13} \\ 2/\sqrt{13} \end{pmatrix}$.

So max. temp. at $(x,y) = \left(\frac{3}{\sqrt{13}}, \frac{2}{\sqrt{13}}\right)$.

Definiteness $\quad q(x) = x^T A x$

Q: When is q always positive? (or always negative?)

First a definition:

Def^n $q(x) = x^T A x$ is called (or A itself is called)

(a) positive definite if $x^T A x > 0$ for all $x \neq 0$

(b) negative definite if $x^T A x < 0$ for all $x \neq 0$

(c) indefinite if sometimes $x^T A x > 0$ and sometimes $x^T A x < 0$.

(Note: these are not the only things that can happen but they are the only things we're interested in/worried about)

Answer to the question above is given by:

<u>Theorem</u> If A is symmetric, then $x^T A x$ is

(1) positive definite exactly when all eigenvalues of A are positive

(2) negative definite exactly when all eigenvalues of A are negative

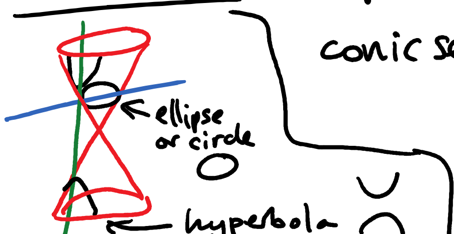
(3) indefinite when at least one eigenvalue of A is positive and at least one eigenvalue of A is negative.

<u>Proof idea</u> Change coordinates to get $x^T A x = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2$

$\rightarrow$ Left as Exerise.

(For (c) look at $e_i^T D e_i$ for each i, and setting $y = e_i$ for each e_i.)

<u>Application</u> Suppose $x^T A x = k$ is the equation of a conic section in $\mathbb{R}^2$.



Do orthogonal change of coordinates to get

$y^T D y = 1$

(In $\mathbb{R}^2$) this is $\lambda_1 y_1^2 + \lambda_2 y_2^2 = 1$

So $x^TAx = k$ represents

- an ellipse if $\lambda_1, \lambda_2 > 0$ i.e. if A positive definite
- an hyperbola if λ_1, λ_2 opposite signs i.e. A is indefinite

(- no graph at all if $\lambda_1, \lambda_2 < 0$ i.e. A is negative definite).

Here is a different way to determine if x^TAx (or A) is positive/negative definite or indefinite without using eigenvalues.

__Def^n__ For a $n \times n$ matrix

$P_{A,1}$ $P_{A,2}$ $P_{A,3}$ $P_{A,n}$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ & & & & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}$$

we call the "top left" $k \times k$ submatrix of A (1st k rows $\times$ 1st k columns) the __kth principal__ submatrix of A, called $P_{A,k}$.

<u>Theorem</u> If A symmetric, then $x^T A x$ is

(1) positive definite if every principal submatrix of A has positive determinant ($\det P_{A,k} > 0$)

(2) negative definite if the signs of the determinants of the principal submatrices of A alternate, starting negative i.e.

$\det(P_{A,1}) < 0$, $\det(P_{A,2}) > 0$, $\det(P_{A,3}) < 0$ and so on ...

(3) indefinite if some ~~of the~~ determinant of a principal submatrix of A is negative and some determinants is positive, & we're not in the case that $x^T A x$ is negative definite (i.e. situation in (2)).

Application to Derivative Test

Let $f: \mathbb{R}^2 \to \mathbb{R}$ be a function with all 2nd partial derivatives.

Recall critical point of f: $(x_0, y_0) \in \mathbb{R}^2$ with

$$f_x(x_0, y_0) = 0 = f_y(x_0, y_0).$$

Theorem from Calculus If (x_0, y_0) is a critical point of f (with f having continuous second derivatives in an open region containing (x_0, y_0). Then

(a) f has a local minimum at (x_0, y_0)

$$\text{if } \underbrace{f_{xx}(x_0,y_0)f_{yy}(x_0,y_0) - f_{xy}^2(x_0,y_0)}_{\text{call this } g(x_0,y_0)} > 0$$

and $f_{xx}(x_0, y_0) > 0$

(b) f has a local max. at (x_0, y_0) if
$g(x_0, y_0) > 0$ & $f_{xx}(x_0, y_0) < 0$

(c) f has a saddle point if $g(x_0, y_0) < 0$.

(Test inconclusive if $g(x_0, y_0) = 0$.)

We can put this in terms of definiteness of quadratic forms.

Definition The **Hessian** of f is the matrix

$$H(x,y) = \begin{pmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{yx}(x,y) & f_{yy}(x,y) \end{pmatrix}$$

Symmetric!!

So the **Theorem** in terms of matrices says:

For f and (x_0,y_0) as above,

(a) f has a local minimum at (x_0,y_0) if $H(x_0,y_0)$ is positive definite

(b) f has a local maximum at (x_0,y_0) if $H(x_0,y_0)$ is negative definite

(c) f has a saddle point at (x_0,y_0) if $H(x_0,y_0)$ is indefinite.

This idea can be generalized to functions of n variables with the appropriate Hessian (symmetric) $n \times n$ matrix. Thus the "second derivative test" for determining the nature of critical points comes down to determining if a certain matrix is positive definite, negative definite or indefinite — and this we know we can do by finding the eigenvalues