SIS Scuola Interdipartimentale delle Scienze, dell'Ingegneria e della Salute

DiST Dipartimento di Scienze e Tecnologie

L. Magistrale in IA (ML&BD)

Scientific Computing (part 2 – 6 credits)

prof. Mariarosaria Rizzardi

Centro Direzionale di Napoli – Bldg. C4
room: n. 423 – North Side, 4th floor
phone: 081 547 6545
email: mariarosaria.rizzardi@uniparthenope.it

Contents

- Induced matrix norms.
- Geometrical interpretation of matrix norms.
- Condition number of a matrix.
- Orthogonality and linear independence.
- Another definition of the standard scalar product.

Induced matrix norm

By definition, a ***matrix norm*** is said ***induced by a vector norm*** if

$$\|A\| \overset{\text{def}}{=} \max_{v \neq \underline{0}} \frac{\|A\,v\|}{\|v\|}$$

the same as $\|A\| \overset{\text{def}}{=} \max_{\|v\|=1} \|A\,v\|$

matrix norms

vector norms

where A is a rectangular matrix of size (m×n).

An *induced matrix norm* satisfies all the properties of a vector norm and, in addition, the following:

5. $\|A v\| \le \|A\|\,\|v\|$

6. $\|AB\| \le \|A\|\,\|B\|$

Examples of induced matrix norms

Euclidean norm (or 2-norm)

$$\|A\|_2 = \sqrt{\max\left|eigenvalue\right| \; of \; A^H A}$$

max singular value of A

```
A=rand(3);
disp([norm(A) max(sqrt(eig(A'*A)))])
     1.4465     1.4465
```

Uniform norm (or maximum norm or ∞-norm)

$$\|A\|_\infty = \max_{1\le i\le m} \sum_{j=1}^{n} \left|a_{ij}\right|$$

sum over cols
max over rows

```
disp([norm(A,inf) max(sum(abs(A),2))])
1.9389    1.9389
```

Taxicab norm (or Manhattan norm or 1-norm)

$$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{m} \left|a_{ij}\right|$$

sum on rows
max over cols

```
disp([norm(A,1) max(sum(abs(A)))])
1.8795    1.8795
```

What is measured by a matrix norm?

A vector norm gives the length of a vector.

by Propr. 5 ⇒ $\|Av\| \le \|A\| \cdot \|v\|, \quad v \neq \underline{0} \quad \Leftrightarrow \quad \dfrac{\|Av\|}{\|v\|} \le \|A\|$

scale factor

a matrix norm $\|A\|$ gives the maximum amplification of $\|x\|$ in Ax.

Example

$$v = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \|v\|_2 \cong 2.24, \qquad v_A = Av, \quad v_B = Bv, \quad v_C = Cv$$

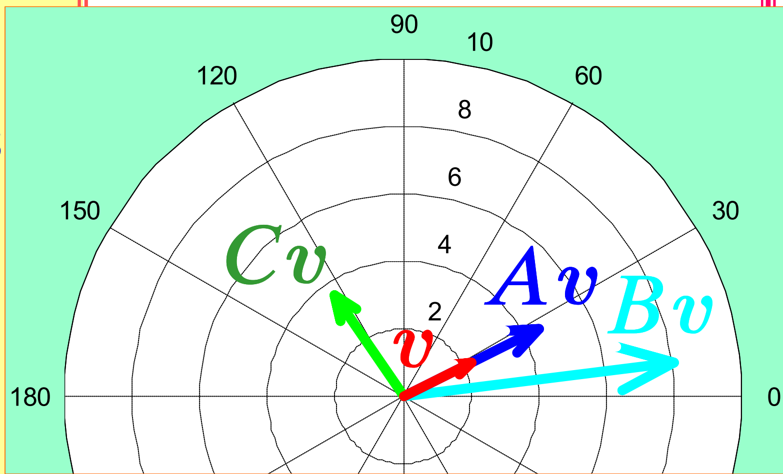
$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \Rightarrow \|A\|_2 = 2, \quad \frac{\|Av\|_2}{\|v\|_2} = 2$$

$$B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow \|B\|_2 = 4, \quad \frac{\|Bv\|_2}{\|v\|_2} \cong 3.6$$

$$C = \begin{pmatrix} -2 & 2 \\ 1 & 1 \end{pmatrix} \Rightarrow \|C\|_2 \cong 2.8, \quad \frac{\|Cv\|_2}{\|v\|_2} \cong 1.6$$

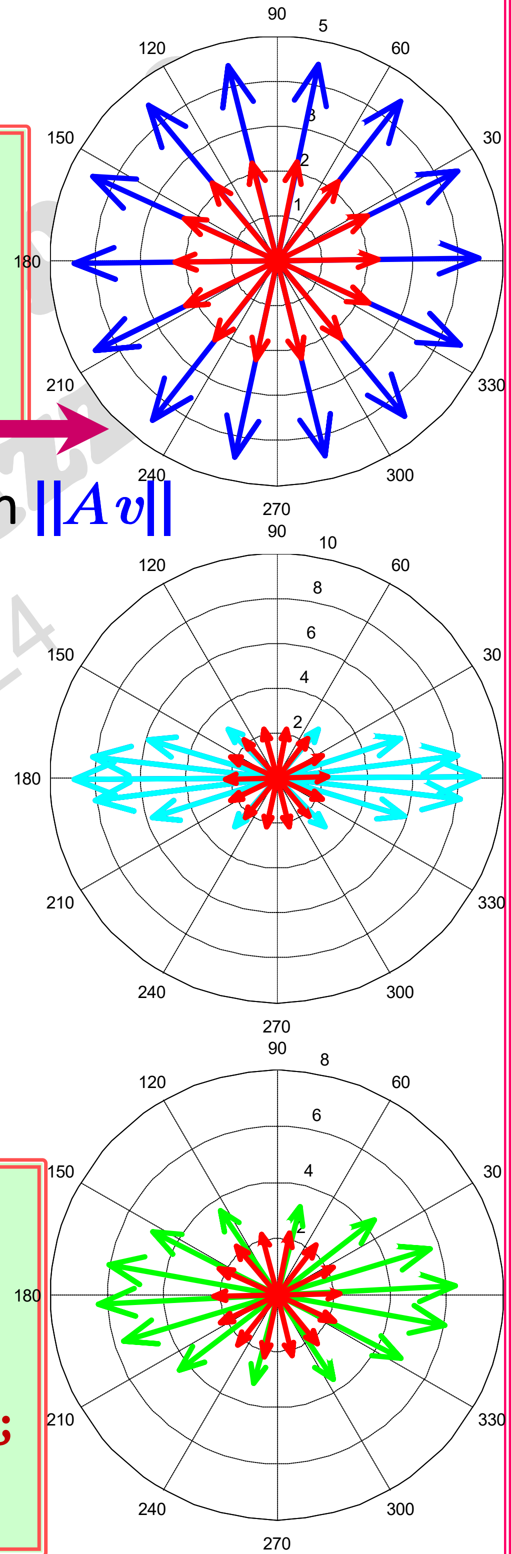
The matrix with maximum norm gives the longest transformed vector

```
A=2*eye(2); B=diag([4 1]); C=[-2 2;1 1];
v=[2 1]'; vA=A*v; vB=B*v; vC=C*v;
compass(vB(1),vB(2),'c'); hold on
compass(vA(1),vA(2),'b'); compass(vC(1),vC(2),'g');
compass(v(1),v(2),'r'); disp(norm(v))
    2.2361
disp([norm(A)  norm(B)  norm(C); ...
      norm(vA) norm(vB) norm(vC)])
      2.0000    4.0000    2.8284
      4.4721    8.0623    3.6056
```

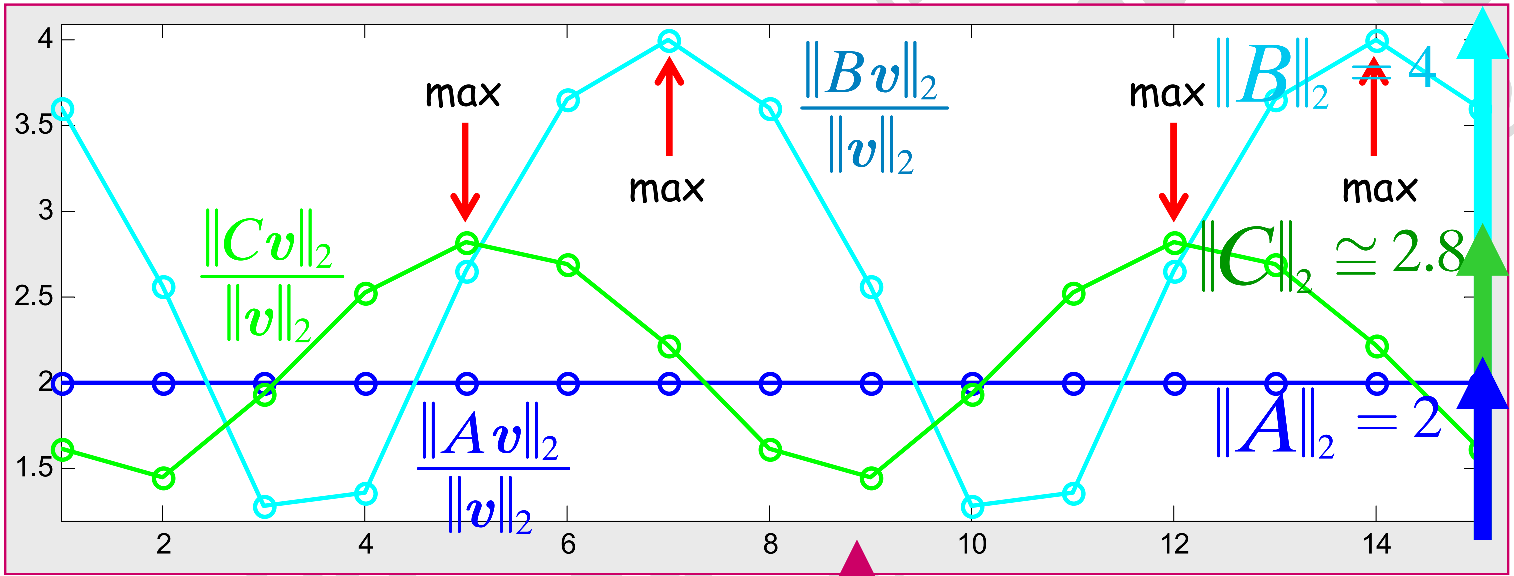


Example (contd.)

```
   ...
N=15;  t=linspace(-pi,pi,N);  z=(2+i)*exp(i*t);
v=[real(z);imag(z)]; vA=A*v; vB=B*v; vC=C*v;
figure; compass(vA(1,:),vA(2,:),'b'); hold on; compass(z,'r')
figure; compass(vB(1,:),vB(2,:),'c'); hold on; compass(z,'r')
figure; compass(vC(1,:),vC(2,:),'g'); hold on; compass(z,'r')
```



the matrix norm $\|A\|$ gives the maximum elongation of $\|v\|$ in $\|Av\|$




```
   ...
nv=vecnorm(v);
nvA=vecnorm(vA);
nvB=vecnorm(vB);
nvC=vecnorm(vC);
figure; h=plot((1:N),[(nvA./nv)' (nvB./nv)' (nvC./nv)'],'o-');
set(h(1),'Color','b');set(h(2),'Color','c');set(h(3),'Color','g')
axis tight
```

compute the $\| \|_2$ for each column in the matrix

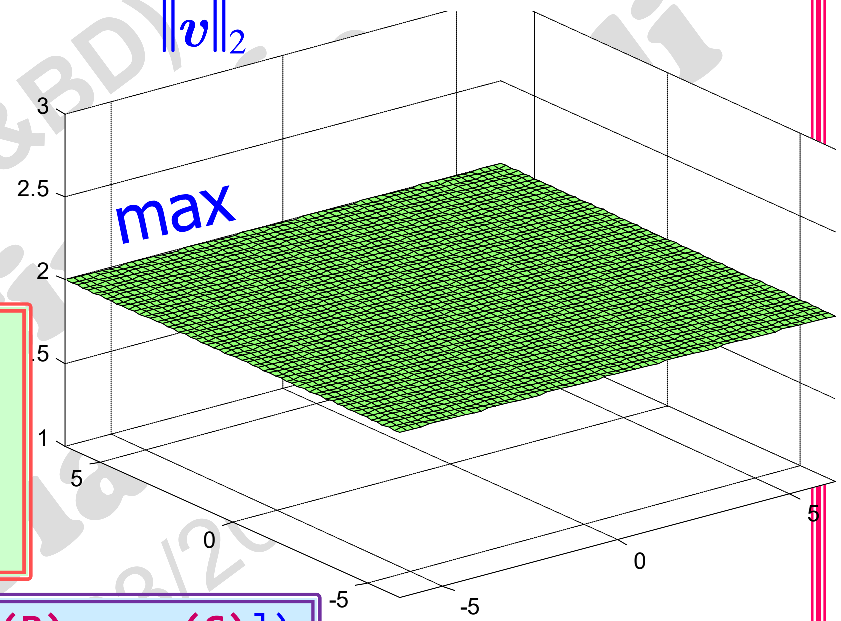
Lab: estimate $\|A\|$ using the Symbolic Math Toolbox

$$\|A\| \overset{\text{def}}{=} \max_{v \neq 0} \frac{\|A\,v\|}{\|v\|}$$

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} -2 & 2 \\ 1 & 1 \end{pmatrix}$$

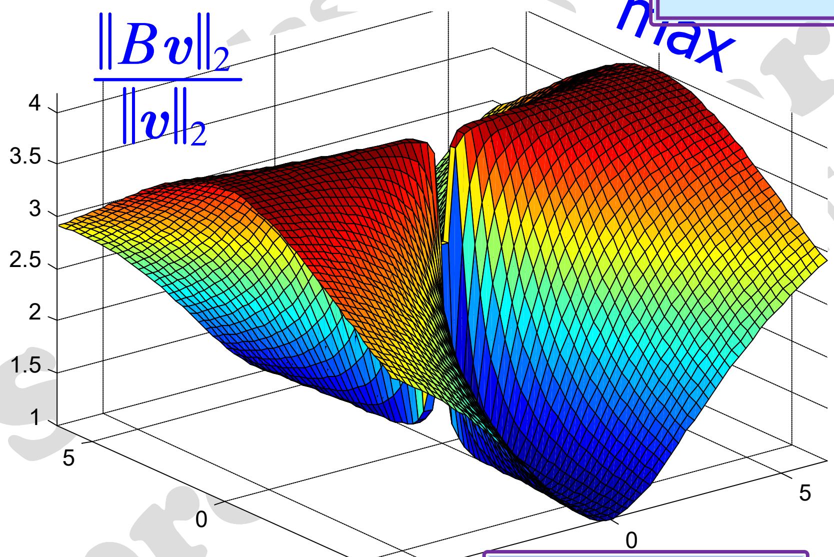
```
A=2*eye(2); B=[4 0;0 1]; C=[-2 2;1 1];
syms x y real;  v=[x y]';
Anorm=simplify(norm(A*v)/norm(v)); ezsurf(Anorm)
Bnorm=simplify(norm(B*v)/norm(v)); ezsurf(Bnorm)
Cnorm=simplify(norm(C*v)/norm(v)); ezsurf(Cnorm)
```

$\frac{\|A\,v\|_2}{\|v\|_2}$ $\quad \|A\|_2=2$




```
disp([norm(A) norm(B) norm(C)])
         2         4    2.8284
```

$\|B\|_2=4$ $\quad \frac{\|B\,v\|_2}{\|v\|_2}$

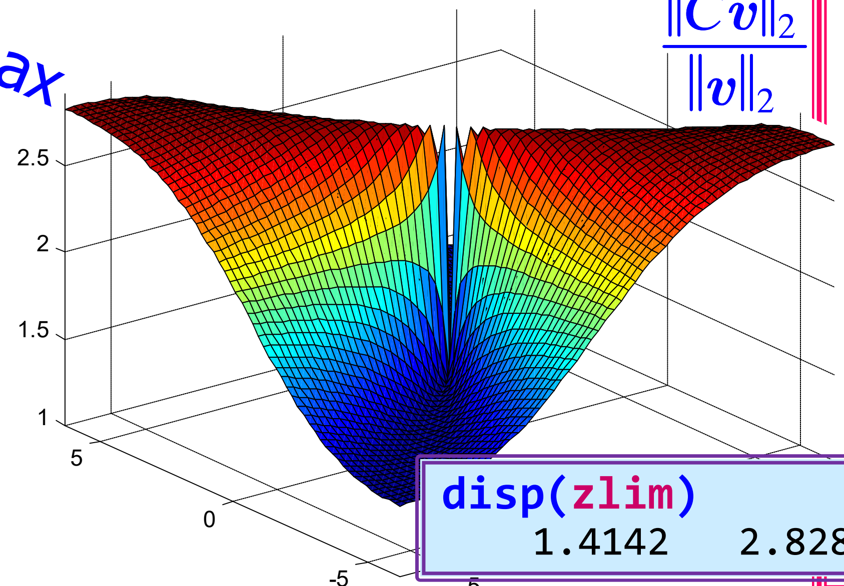



```
AX=axis; disp(AX(5:6))
    1.0022      3.9995
```

```
disp(zlim)
```

approx of $\|B\|_2$

$\|C\|_2=2.83$ $\quad \frac{\|C\,v\|_2}{\|v\|_2}$




```
disp(zlim)
    1.4142    2.8284
```

$$\|B\| \stackrel{\text{def}}{=} \max_{v\neq 0} \frac{\|B\,v\|}{\|v\|}$$

Lab. (contd.)

```
syms x y real; v=[x y]'; B=diag([4 1]);
Bnorm=simplify(norm(B*v)/norm(v));
ezsurf(Bnorm); colormap('jet')
G=simplify(gradient(Bnorm));
S=solve(G(1)==0, G(2)==0, ...
      'ReturnConditions',true) % search stationary pts
S = struct with fields:
    x: [2×1 sym]
    y: [2×1 sym]
parameters: u
conditions: [2×1 sym]
S.parameters
ans = u
S.conditions
ans = in(u, 'real')
      in(u, 'real')
S.x
ans = u
      0
S.y
ans = 0
      u
V1=simplify(subs(Bnorm,{x,y},{S.x(1),S.y(1)}))
V1  = 4      max
V2=simplify(subs(Bnorm,{x,y},{S.x(2),S.y(2)}))
V2  = 1      min
```

$$B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$$

gradient of φ: $\nabla\varphi(x_1^*, x_2^*) = \begin{pmatrix} \frac{\partial\varphi}{\partial x_1}(x_1^*, x_2^*) \\ \frac{\partial\varphi}{\partial x_2}(x_1^*, x_2^*) \end{pmatrix}$

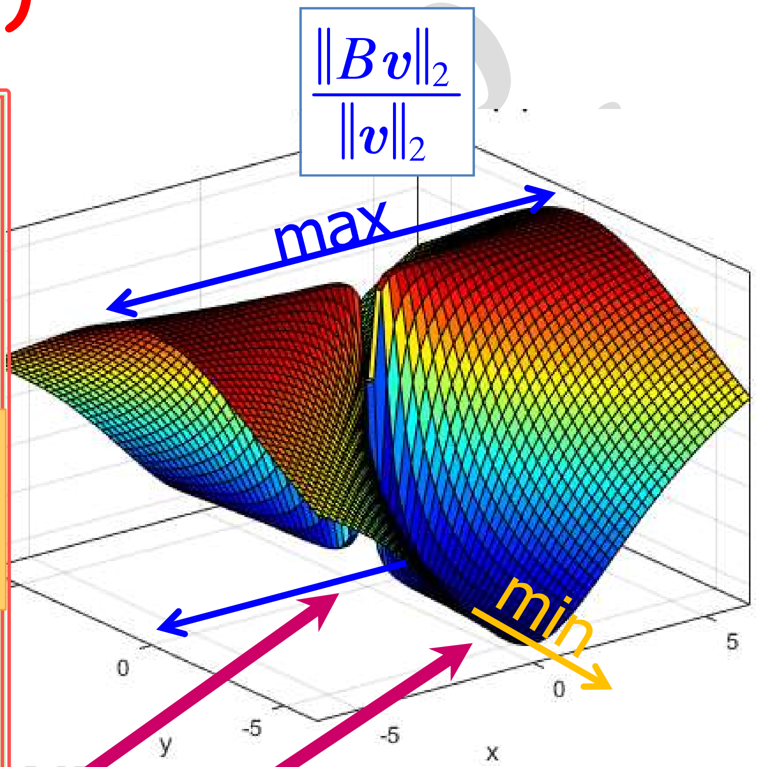
stationary points:

$[S.x(1);S.y(1)] = u\begin{pmatrix}1\\0\end{pmatrix}$ x-axis

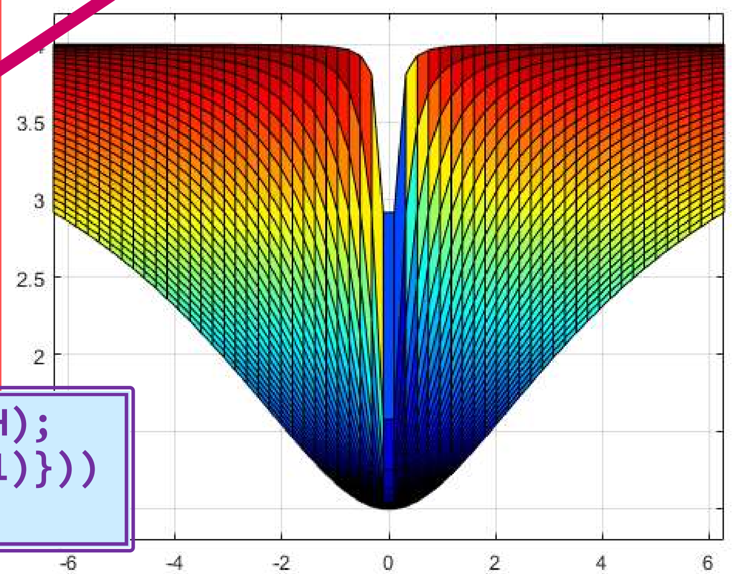
$[S.x(2);S.y(2)] = u\begin{pmatrix}0\\1\end{pmatrix}$ y-axis

⇒ $\|B\| = 4$

```
H=hessian(Bnorm,[x,y]); detH=det(H);
disp(subs(detH,{x,y},{S.x(1),S.y(1)}))
0
```

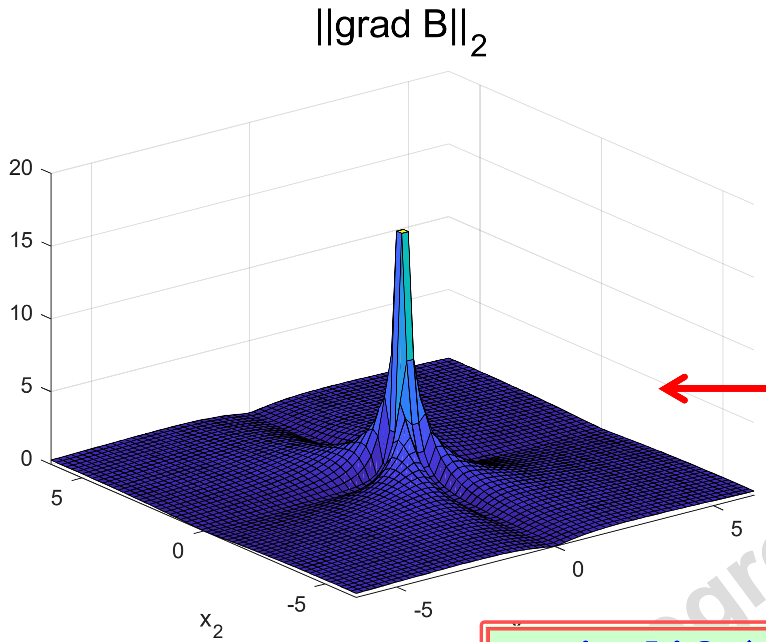


Go to X-Z view:
the surface is all between 1 and 4

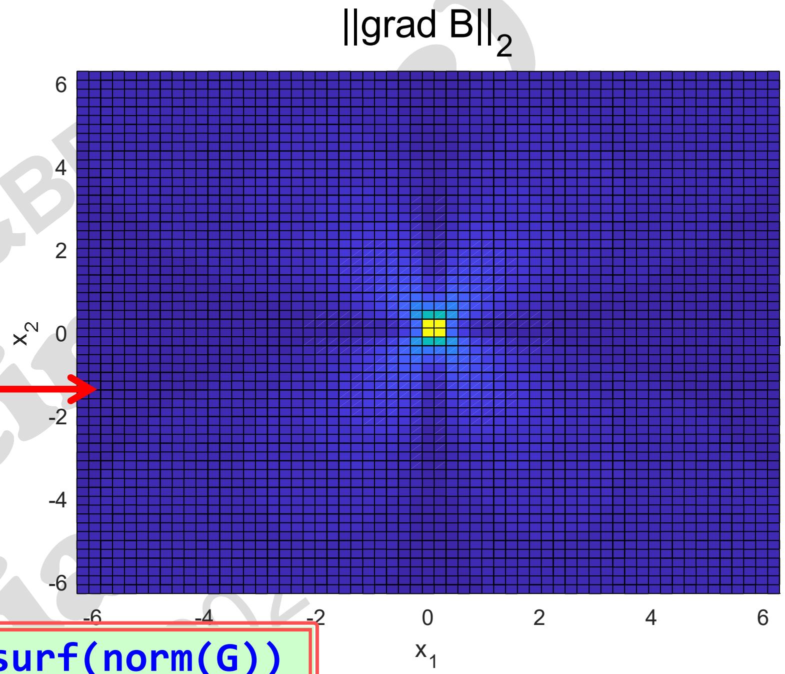


$(\mathbf{H}_f)_{i,j} = \frac{\partial^2 f}{\partial x_i\,\partial x_j}$ The Hessian matrix **H** can locate the local extrema. But if det(**H**)=0 nothing can be said.

How to highlight where it is $\left\|\nabla\varphi\left(x_1^*,x_2^*\right)\right\|_2 = 0$ top view



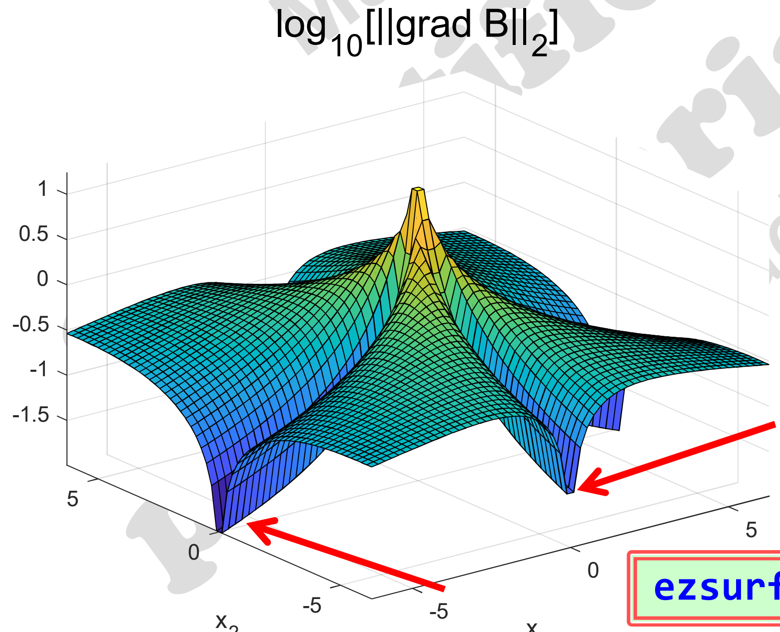
unclear!



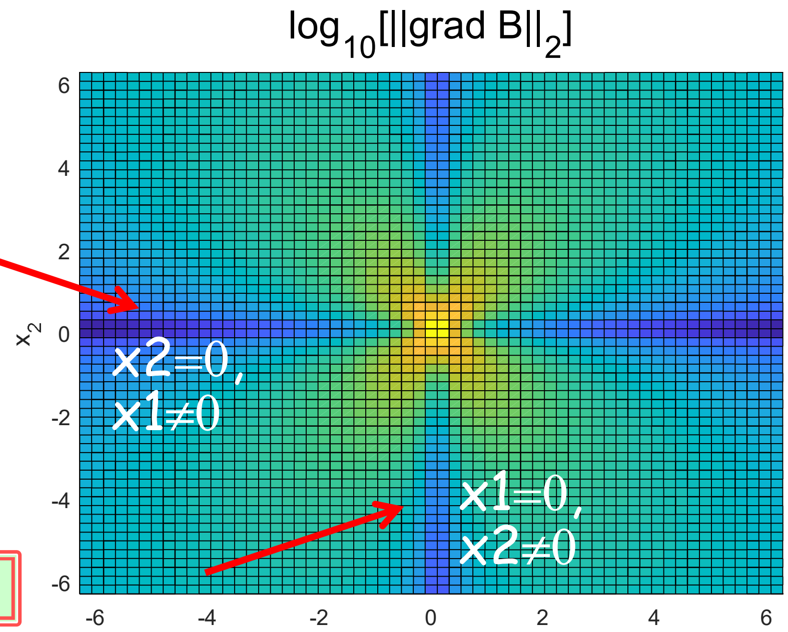

```
G=simplify(gradient(Bnorm)); ezsurf(norm(G))
```

If the surface is not negative, it may be preferable to draw its **log$_{10}$**

$$\left\|\nabla\varphi\left(x_1^*,x_2^*\right)\right\|_2 = 0 \iff \log_{10}\left\|\nabla\varphi\left(x_1^*,x_2^*\right)\right\|_2 = -\infty$$



clear!




```
ezsurf(log10(norm(G)))
```

Lab. (contd.)

$$C=\begin{pmatrix}-2 & 2\\ 1 & 1\end{pmatrix}$$

```
C=[-2 2;1 1]; syms x y real;  v=[x y]';
Cnorm=simplify(norm(C*v)/norm(v));
ezsurf(Cnorm); colormap('jet')
G=simplify(gradient(Cnorm));
S=solve(G(1)==0, G(2)==0, ...
      'ReturnConditions',true) % search stationary pts
S = struct with fields:
    x: [2×1 sym]
    y: [2×1 sym]
parameters: u
conditions: [2×1 sym]
S.parameters
ans = u
S.conditions
ans = in(u, 'real')
      in(u, 'real')
S.x
ans = -u
       u
S.y
ans = u
      u
V1=simplify(subs(Cnorm,{x,y},{S.x(1),S.y(1)}))
V1  = 2*2^(1/2) ≈2.8284   max
V2=simplify(subs(Cnorm,{x,y},{S.x(2),S.y(2)}))
V2  = 2^(1/2) ≈1.4142     min
```

```
S=solve(G, ...
   'ReturnConditions',true)
```

stationary points:

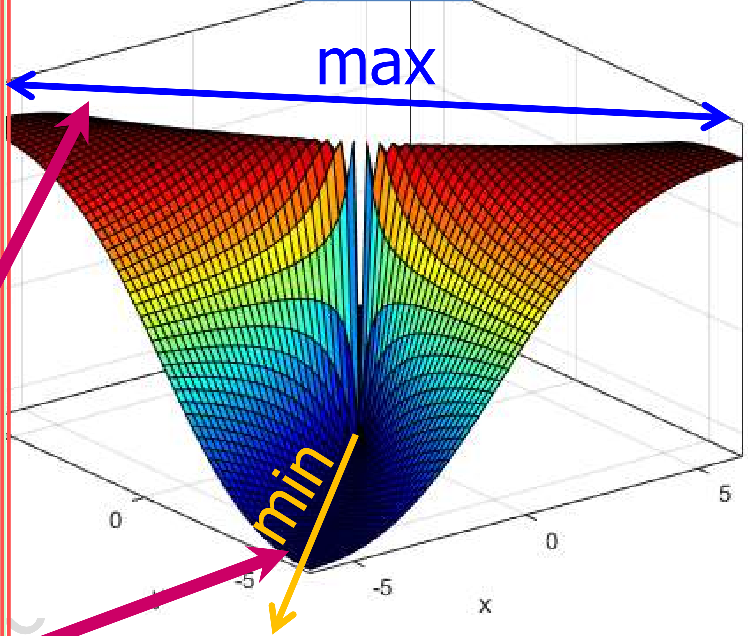
[S.x(1);S.y(1)] $= u\begin{pmatrix}-1\\ 1\end{pmatrix}$

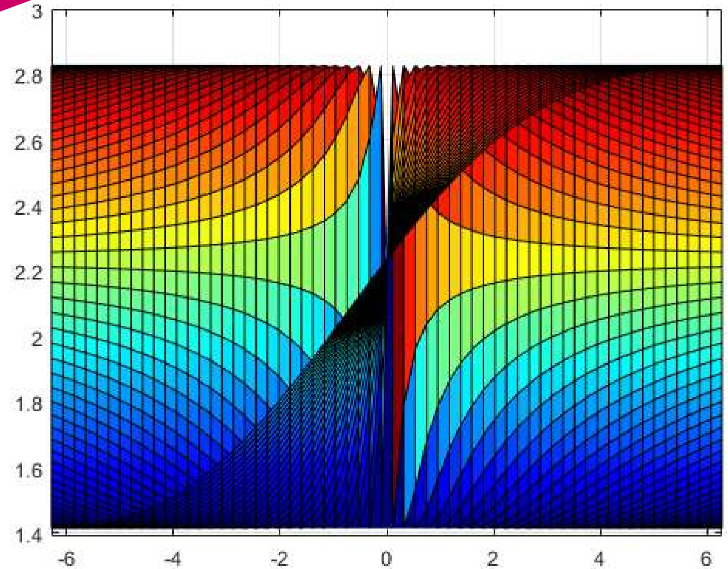
[S.x(2);S.y(2)] $= u\begin{pmatrix}1\\ 1\end{pmatrix}$

```
detH=det(hessian(Cnorm,[x,y]));
disp(subs(detH,{x,y},{S.x(1),S.y(1)}))
0     nothing can be said about this point
```

⇨ ||C|| = 2*sqrt(2)

$$\frac{\|Cv\|_2}{\|v\|_2}$$

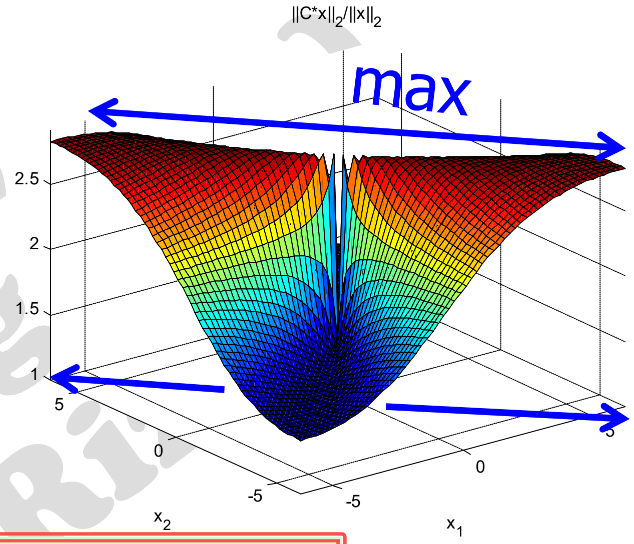




Go to X-Z view:
the surface is all between 1.4.. and 2.8..

$||C||_2$

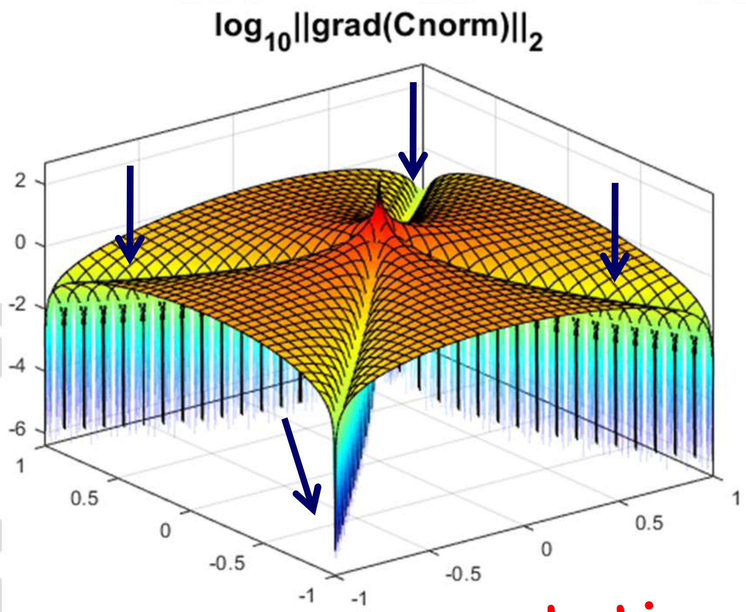
```
C=[-2 2;1 1]; syms x y real; v=[x y]';
Cnorm=simplify(norm(C*v)/norm(v));
ezsurf(Cnorm); colormap('jet')
```

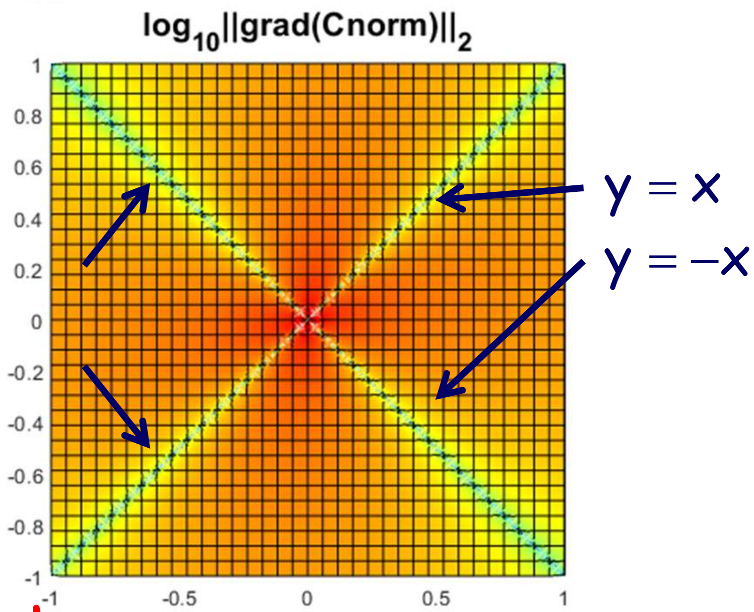


How to highlight where it is $\left\|\nabla\varphi\left(x_1^*, x_2^*\right)\right\|_2 = 0$

```
G=simplify(gradient(Cnorm)); % gradient
Gnorm=simplify(norm(G))        % 2-norma of gradient
fsurf(log10(Gnorm),[-1 1]); box on; view(2); axis equal
```

top view





stationary points

(prof. M. Rizzardi)

Exercise: estimate the condition number using MATLAB Symbolic Math Toolbox

previously estimated →

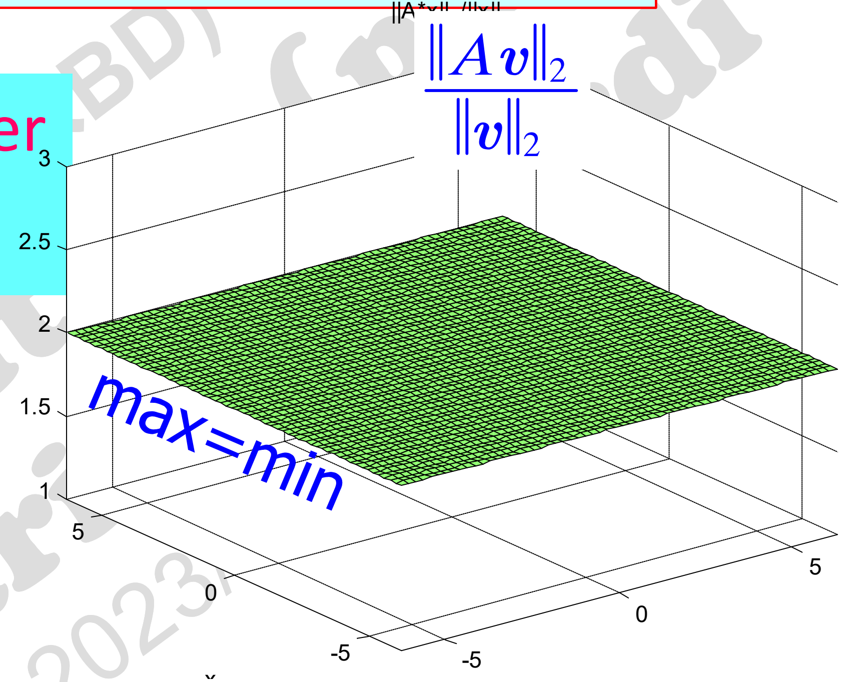
$$M = \max_{v \neq \underline{0}} \frac{\|Av\|_2}{\|v\|_2}$$

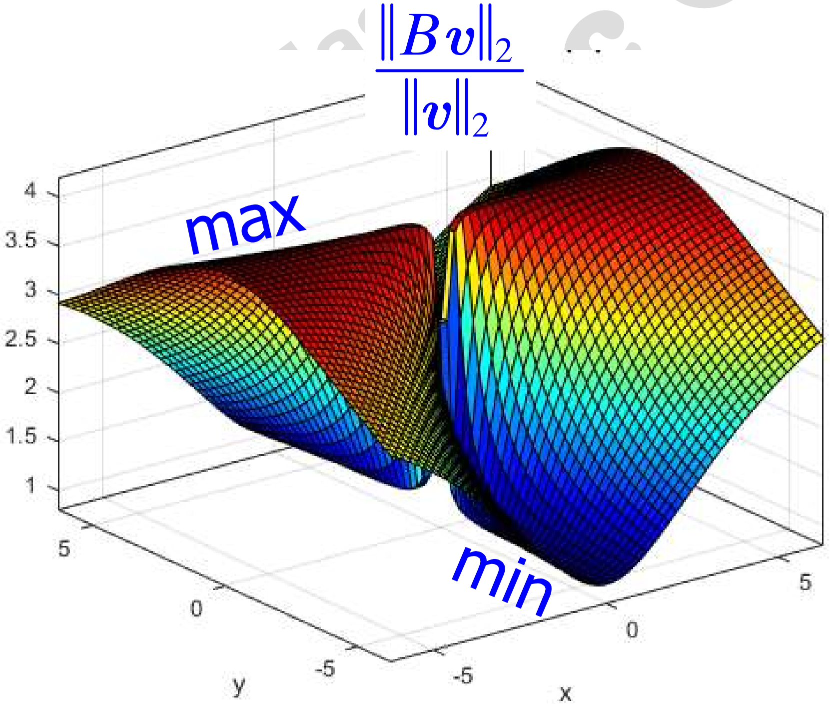
$$m = \min_{v \neq \underline{0}} \frac{\|Av\|_2}{\|v\|_2}$$

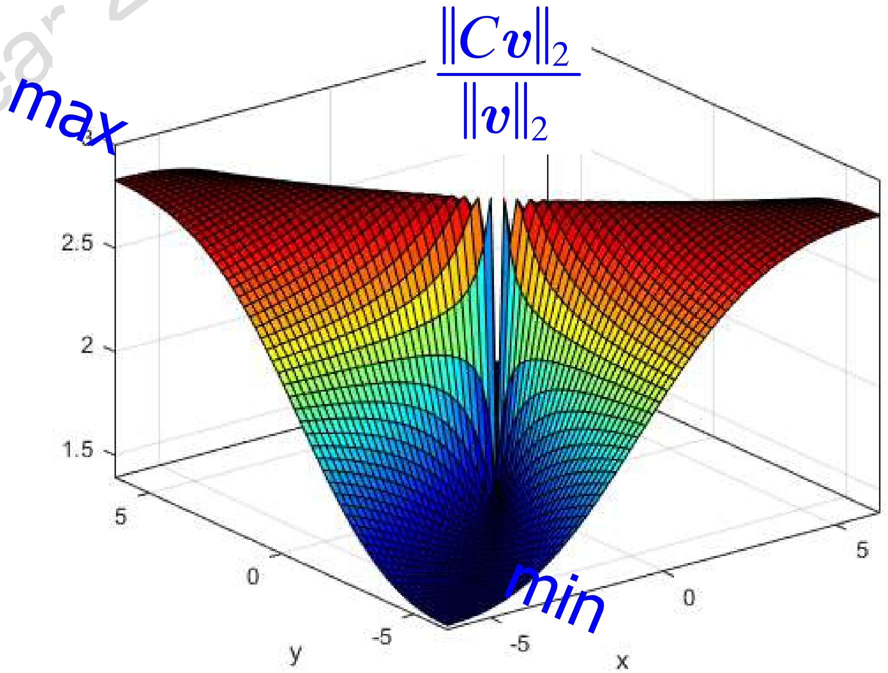
condition number

$$\kappa(A) = \frac{M}{m}$$

```
disp([cond(A) cond(B) cond(C)])
 1.0000    4.0000    2.0000
```







More on the "condition number" of a matrix

Let A be **any matrix** (also rectangular)

A matrix norm measures ...

how much a matrix can stretch vectors: $M = \max_{v \neq \underline{0}} \frac{\|Av\|_2}{\|v\|_2}$

but it may be important to know ...

how much a matrix can shorten vectors: $m = \min_{v \neq \underline{0}} \frac{\|Av\|_2}{\|v\|_2}$

$$\kappa(A) = \frac{M}{m}$$

Let A be a **square matrix**

A singular square matrix A is such that: $m = \min_{v \neq \underline{0}} \frac{\|Av\|_2}{\|v\|_2} = 0$

and the condition number is infinity.

Why?

Let A be now a **square** and **non-singular matrix**: $w = Av \iff v = A^{-1}w$

$$m = \min_{v \neq \underline{0}} \frac{\|Av\|_2}{\|v\|_2} = \min \frac{\|w\|_2}{\|A^{-1}w\|_2} = \frac{1}{\max \frac{\|A^{-1}w\|}{\|w\|}} = \frac{1}{\|A^{-1}\|}$$

condition number for inversion

more general

$$\kappa(A) = \frac{M}{m}$$

only for square invertible matrices

$$\kappa(A) = \|A\| \times \|A^{-1}\|$$

Orthogonal vectors

u*,*v are **orthogonal**

$\boldsymbol{u} \perp \boldsymbol{v}$ $\quad \overset{\text{def}}{\Longleftrightarrow} \quad$ $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$

Connection between independence and orthogonality

n (non-zero) vectors are **mutually orthogonal**

$\Longrightarrow$ they are **linearly independent**

$\not\Longleftarrow$ the reverse is not true

The vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are **orthogonal**, therefore **independent**.

The vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ are **independent**, but **not orthogonal**.

RECALL

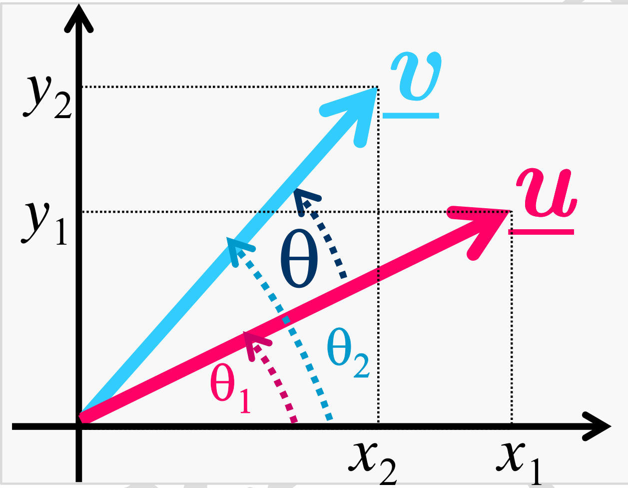
standard scalar product in $\mathbb{R}^n$: $\langle \boldsymbol{u},\boldsymbol{v}\rangle \overset{\text{def}}{=} \sum_{k=1}^{n} u_k v_k$

Another definition of standard scalar product in $\mathbb{R}^n$:

Is its definition equal to: $\langle \boldsymbol{u},\boldsymbol{v}\rangle = \|\boldsymbol{u}\|\times\|\boldsymbol{v}\|\times\cos\theta$?

$\Rightarrow$ $c = \cos(\theta) = \dfrac{\langle u,v\rangle}{\|u\|\cdot\|v\|}$

Since two vectors lie on a plane, the proof is in $\mathbb{R}^2$



By the properties of right triangles we get:

$$x_1 = \|\boldsymbol{u}\|\cos\theta_1, \quad y_1 = \|\boldsymbol{u}\|\sin\theta_1$$
$$x_2 = \|\boldsymbol{v}\|\cos\theta_2, \quad y_2 = \|\boldsymbol{v}\|\sin\theta_2$$

By def: $\langle \boldsymbol{u},\boldsymbol{v}\rangle = x_1x_2 + y_1y_2 = \|\boldsymbol{u}\|\|\boldsymbol{v}\|\underbrace{[\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2]}_{\cos(\theta_2-\theta_1)=\cos\theta}$

Why is it sufficient to prove the formula in $\mathbb{R}^2$?

(prof. M. Rizzardi)