Lecture 26/28—Positive Definite Matrices

Linear Algebra
MATH 124, Fall, 2010

Prof. Peter Dodds

Department of Mathematics & Statistics
Center for Complex Systems
Vermont Advanced Computing Center
University of Vermont

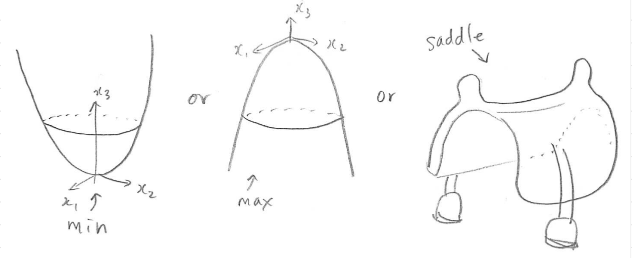
Outline

Simple example problem 1 of 2:

What does this function look like?:

$$f(x_1, x_2) = 2x_1^2 - 2x_1x_2 + 2x_2^2.$$

- Three main categories:



- Standard approach for determining type of extremum involves calculus, derivatives, horrible things...
- Obviously, we should be using linear algebra...

Simple example problem 1 of 2:

Linear Algebra-ization...

- We can rewrite

$$f(x_1, x_2) = 2x_1^2 - 2x_1x_2 + 2x_2^2.$$

as

$$f(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \boxed{\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x}}$$

- Note: A is symmetric as $A = A^{\mathrm{T}}$ (delicious).
- Interesting and sneaky...

Simple example problem 2 of 2:

What about this curve?:

$$2x_1^2 + 2x_1x_2 + 2x_2^2 = 1.$$

Linear Algebra-ization...

Again, we'll see we can rewrite as

$$1 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \boxed{\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x}}$$

Goal:

- Understand how $\mathbb{A}$ governs the form $\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x}$.
- Somehow, this understanding will involve almost everything we've learnt so far: row reduction, pivots, eigenthings, symmetry, ...

General 2×2 example:

- Write $\mathbb{A} = \mathbb{A}^{\mathrm{T}} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$.

- $$\vec{x}^{\mathrm{T}} \mathbb{A} \, \vec{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} ax_1 + bx_2 \\ bx_1 + cx_2 \end{bmatrix}$$

$$= x_1(ax_1 + bx_2) + x_2(bx_1 + cx_2) = ax_1^2 + bx_1 x_2 + bx_1 x_2 + cx_2^2$$

$$= ax_1^2 + 2bx_1 x_2 + cx_2^2.$$

- See how a, b, and c end up in the quadratic form.

General 2×2 example—creating $\mathbb{A}$:

We have: $\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x} = a x_1^2 + 2 b x_1 x_2 + c x_2^2 = f(x_1, x_2)$

- Back to our first example:
  $f(x_1, x_2) = 2x_1^2 - 2x_1 x_2 + 2x_2^2$.
- Identify $a = 2$, $b = -1$, and $c = 2$.
- $$:f(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$
- Second example: $2x_1^2 + 2x_1 x_2 + 2x_2^2 = 1$.
- Identify $a = 2$, $b = 1$, and $c = 2$.
- $$: \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 1$$

General 3×3 example:

- $$\text{Write } \mathbb{A} = \mathbb{A}^{\mathrm{T}} = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}.$$

- $$\vec{x}^{\mathrm{T}} \mathbb{A}\, \vec{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

  $$= a x_1^2 + d x_2^2 + f x_3^2 + 2 b x_1 x_2 + 2 c x_1 x_3 + 2 e x_2 x_3.$$

- Again: see how the terms in $\mathbb{A}$ distribute into the quadratic form.

General story:

- Using the definition of matrix multiplication,

$$\vec{x}^{\mathrm{T}} \mathbb{A}\, \vec{x} = \sum_{i=1}^{n} [\vec{x}^{\mathrm{T}}]_i [\mathbb{A}\, \vec{x}]_i = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ij} x_j = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

- We see the $x_i x_j$ term is attached to a_{ij}.
- On-diagonal terms look like this: $a_{77} x_7^2$ and $a_{33} x_3^2$.
- Off-diagonal terms combine, e.g., $(a_{13} + a_{31}) x_1 x_3$.
- Given some f with a term $23 x_1 x_3$, we could divide the 23 between a_{13} and a_{31} however we like.
- e.g., $a_{13} = 36$ and $a_{31} = -13$ would work.
- But we choose to make $\mathbb{A}$ symmetric because symmetry is great.

A little abstraction:

A few observations:

1. The construction $\vec{x}^{\mathrm{T}}\mathbb{A}\,\vec{x}$ appears naturally.
2. Dimensions of $\vec{x}^{\mathrm{T}}$, $\mathbb{A}$, and $\vec{x}$:
   1 by n, n by n, and n by 1.
3. $\vec{x}^{\mathrm{T}}\mathbb{A}\,\vec{x}$ is a 1 by 1.
4. If $\mathbb{A}\vec{v} = \lambda\vec{v}$ then

$$\vec{v}^{\mathrm{T}}\mathbb{A}\vec{v} = \vec{v}^{\mathrm{T}}(\mathbb{A}\vec{v}) = \vec{v}^{\mathrm{T}}(\lambda\vec{v}) = \lambda\vec{v}^{\mathrm{T}}\vec{v} = \lambda||\vec{v}||^2.$$

5. If $\lambda > 0$, then $\vec{v}^{\mathrm{T}}\mathbb{A}\vec{v} > 0$ always (given $\vec{v} \neq \vec{0}$).
6. Suggests we can build up to saying something about $\vec{x}^{\mathrm{T}}\mathbb{A}\,\vec{x}$ starting from eigenvalues...

Lecture 26

Definitions:

Positive Definite Matrices (PDMs):

- Real, symmetric matrices with positive eigenvalues.
- Math version:

$$\mathbb{A} = \mathbb{A}^{\mathrm{T}},$$

$$a_{ij} \in R \; \forall \; i, j = 1, 2, \cdots n,$$

$$\text{and } \lambda_i > 0, \; \forall \; i = 1, 2, \cdots n.$$

Semi-Positive Definite Matrices (SPDMs):

- Same as for PDMs but now eigenvalues may now be 0:

$$\lambda_i \geq 0, \; \forall \; i = 1, 2, \cdots, n.$$

- Note: If some eigenvalues are < 0 we have a sneaky matrix.

Equivalent Definitions:

Positive Definite Matrices:

- $\mathbb{A} = \mathbb{A}^{\mathrm{T}}$ is a PDM if

$$\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x} > 0 \;\; \forall \; \vec{x} \neq \vec{0}$$

Semi-Positive Definite Matrices:

- $\mathbb{A} = \mathbb{A}^{\mathrm{T}}$ is a SPDM if

$$\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x} \geq 0$$

Connecting these definitions:

Spectral Theorem for Symmetric Matrices:

$$\boxed{\mathbb{A} = \mathbb{Q}\,\Lambda\,\mathbb{Q}^{\mathrm{T}}}$$

where $\mathbb{Q}^{-1} = \mathbb{Q}^{\mathrm{T}}$,

$$\mathbb{Q} = \begin{bmatrix} | & | & \cdots & | \\ \hat{v}_1 & \hat{v}_2 & \cdots & \hat{v}_n \\ | & | & \cdots & | \end{bmatrix}, \text{ and } \Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

- Special form of $\mathbb{A} = \mathbb{S}\Lambda\mathbb{S}^{-1}$ that arises when $\mathbb{A} = \mathbb{A}^{\mathrm{T}}$.

Understanding $\vec{x}^{\mathrm{T}} \mathbb{A}\, \vec{x}$:

- Substitute $\mathbb{A} = \mathbb{Q}\, \Lambda\, \mathbb{Q}^{\mathrm{T}}$ into $\vec{x}^{\mathrm{T}} \mathbb{A}\, \vec{x}$:

$$= \vec{x}^{\mathrm{T}} \left(\mathbb{Q}\, \Lambda\, \mathbb{Q}^{\mathrm{T}}\right) \vec{x} = \left(\vec{x}^{\mathrm{T}} \mathbb{Q}\right)\, \Lambda\, \left(\mathbb{Q}^{\mathrm{T}} \vec{x}\right) = (\mathbb{Q}^{\mathrm{T}} \vec{x})^{\mathrm{T}}\, \Lambda\, (\mathbb{Q}^{\mathrm{T}} \vec{x}).$$

- We now see $\vec{x}$ transforming from the natural basis to $\mathbb{A}$'s eigenvector basis: $\vec{y} = \mathbb{Q}^{\mathrm{T}} \vec{x}$.

$$:\vec{x}^{\mathrm{T}} \mathbb{A}\, \vec{x} = \vec{y}^{\mathrm{T}} \Lambda\, \vec{y}$$

$$= \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2.$$

The UNIVERSITY of VERMONT

Understanding $\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x}$:

So now we have...

$$\boxed{\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2}$$

- Can see whether or not $\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x} > 0$ depends on the λ_i since each $y_i^2 > 0$.
- So a PDM must have each $\lambda_i > 0$.
- And a SPDM must have $\lambda_i \geq 0$.

The UNIVERSITY of VERMONT

More understanding of $\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x}$:

- Substitute $\mathbb{A} = \mathbb{L}\,\mathbb{D}\,\mathbb{L}^{\mathrm{T}}$ into $\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x}$:

$$= \vec{x}^{\mathrm{T}} \left(\mathbb{L}\,\mathbb{D}\,\mathbb{L}^{\mathrm{T}}\right) \vec{x} = \left(\vec{x}^{\mathrm{T}}\mathbb{L}\right) \mathbb{D} \left(\mathbb{L}^{\mathrm{T}}\vec{x}\right) = \left(\mathbb{L}^{\mathrm{T}}\vec{x}\right)^{\mathrm{T}} \mathbb{D} \left(\mathbb{L}^{\mathrm{T}}\vec{x}\right).$$

- Change from eigenvalue story: $\vec{x}$ is transformed into $\vec{z} = \mathbb{L}^{\mathrm{T}}\vec{x}$ but this is not a change of basis.

$$:\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x} = \vec{z}^{\mathrm{T}} \mathbb{D} \vec{z}$$

$$= \begin{bmatrix} z_1 & z_2 & \cdots & z_n \end{bmatrix} \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}$$

$$= d_1 z_1^2 + d_2 z_2^2 + \cdots + d_n z_n^2.$$

More understanding of $\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x}$:

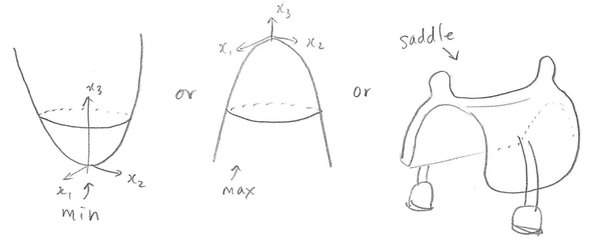
So now we have...

$$\boxed{\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x} = d_1 z_1^2 + d_2 z_2^2 + \cdots + d_n z_n^2}$$

- Can see whether or not $\vec{x}^{\mathrm{T}} \mathbb{A} \vec{x} > 0$ depends on the d_i since each $z_i^2 > 0$.
- So a PDM must have each $d_i > 0$.
- And a SPDM must have $d_i \geq 0$.

Back to general 2×2 example:

$$f(x, y) = \vec{x}^{\mathrm{T}} \mathbb{A} \, \vec{x} = a x_1^2 + 2 b x_1 x_2 + c x_2^2$$



Focus on eigenvalues—We can now see:

- $f(x, y)$ has a minimum at $x = y = 0$ iff $\mathbb{A}$ is a PDM, i.e., if $\lambda_1 > 0$ and $\lambda_2 > 0$.
- Maximum: if $\lambda_1 < 0$ and $\lambda_2 < 0$.
- Saddle: if $\lambda_1 > 0$ and $\lambda_2 < 0$.

Back to simple example problem 1 of 2:

$$f(x_1, x_2) = 2x_1^2 - 2x_1x_2 + 2x_2^2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Compute eigenvalues...

- Find $\lambda_1 = +3$ and $\lambda_2 = +1$: f is a minimum.

General problem:

- How do we easily find the signs of λs...?

Excitement about symmetric matrices:

- We recall with alacrity the totally amazing fact that real symmetric matrices always have (1) real eigenvalues, and (2) orthogonal eigenvectors forming a basis for R^n.
- We now see that knowing the signs of the λs is also important...

Test cases:

- $\mathbb{A}_1 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$, $\mathbb{A}_2 = \begin{bmatrix} 2 & -1 \\ -1 & -2 \end{bmatrix}$, $\mathbb{A}_3 = \begin{bmatrix} -2 & -1 \\ -1 & -2 \end{bmatrix}$

Some minor struggling leads to::

- $\mathbb{A}_1 : \lambda_1 = +3, \lambda_2 = +1,$ (PDM, happy),
- $\mathbb{A}_2 : \lambda_1 = +\sqrt{5}, \lambda_2 = -\sqrt{5},$ (sad),
- $\mathbb{A}_3 : \lambda_1 = -1, \lambda_2 = -3,$ (sad)

Pure madness:

Extremely Sneaky Result #632:

If $\mathbb{A} = \mathbb{A}^{\mathrm{T}}$ and $\mathbb{A}$ is real, then

- # +ve eigenvalues = # +ve pivots
- # -ve eigenvalues = # -ve pivots
- # 0 eigenvalues = # 0 pivots

Notes:

- Previously, we had for general $\mathbb{A}$ that $|\mathbb{A}| = \prod \lambda_i = \pm \prod d_i$.
- The bonus here is for real symmetric $\mathbb{A}$.
- Eigenvalues are pivots come from very different parts of linear algebra.
- Crazy connection between eigenvalues and pivots!

Pivots and Eigenvalues:

More notes:

- All very exciting: Pivots are much, much easier to compute.
- (cue balloons, streamers)

Check for our three examples:

- $\mathbb{A}_1 : d_1 = +2,\ d_2 = +\frac{3}{2}$
  ✓ signs match with $\lambda_1 = +3, \lambda_2 = +1$.
- $\mathbb{A}_2 : d_1 = +2,\ d_2 = -\frac{5}{2}$
  ✓ signs match with $\lambda_1 = +\sqrt{5}, \lambda_2 = -\sqrt{5}$.
- $\mathbb{A}_3 : d_1 = -2,\ d_2 = -\frac{3}{2}$
  ✓ signs match with $\lambda_1 = -1, \lambda_2 = -3$.

Beautiful reason:

- Let's show how the signs of eigenvalues match signs of pivots for

$$\mathbb{A}_2 = \begin{bmatrix} 2 & -1 \\ -1 & -2 \end{bmatrix}, \quad \lambda_{1,2} = \pm\sqrt{5}$$

- Compute $\mathbb{LU}$ decomposition:

$$\mathbb{A}_2 = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 0 & -\frac{5}{2} \end{bmatrix} = \mathbb{LU}$$

- $\mathbb{A}_2$ is symmetric, so we can go further:

$$\mathbb{A}_2 = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -\frac{5}{2} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{bmatrix} = \mathbb{LDL}^{\mathrm{T}}$$

Beautiful reason:

- We're here:

$$\mathbb{A}_2 = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -\frac{5}{2} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{bmatrix} = \mathbb{LDL}^{\mathbb{T}}$$

- Now think about this matrix:

$$\mathbb{B}(\ell_{21}) = \begin{bmatrix} 1 & 0 \\ \ell_{21} & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -\frac{5}{2} \end{bmatrix} \begin{bmatrix} 1 & \ell_{21} \\ 0 & 1 \end{bmatrix}$$

- When $\ell_{21} = -\frac{1}{2}$, we have $B(-\frac{1}{2}) = \mathbb{A}_2$.
- Think about what happens as ℓ_{21} changes smoothly from $-\frac{1}{2}$ to 0.
-

$$\mathbb{B}(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -\frac{5}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbb{IDI} = \begin{bmatrix} 2 & 0 \\ 0 & -\frac{5}{2} \end{bmatrix}$$

1. $\mathbb{B}(0) = \mathbb{D}$'s eigenvalues and pivots are both $2, -\frac{5}{2}$.
2. Stronger: As we alter $\mathbb{B}(\ell_{21})$, the pivots do not change!
3. But eigenvalues do change from $+\sqrt{5}$ and $-\sqrt{5}$ to $2, -\frac{5}{2}$.
4. Big deal: because the pivots don't change, the determinant of $\mathbb{B}(\ell_{21})$ never changes:

$$\det \mathbb{B}(\ell_{21}) = d_1 \cdot d_2 = 2 \cdot \left(-\frac{5}{2}\right) = -5 \neq 0$$

5. But we also know $\det \mathbb{B}(\ell_{21}) = \lambda_1 \cdot \lambda_2$.
6. $\therefore$ as ℓ_{21} changes, eigenvalues cannot pass through 0 as determinant would be 0, not -5.
7. $\therefore$ eigenvalues cannot change sign as ℓ_{21} changes...
8. Signs of eigenvalues of $\mathbb{A}_2 = \mathbb{B}(-\frac{1}{2})$ must match signs of eigenvalues of $\mathbb{B}(0)$ which match signs of pivots of $\mathbb{B}(0)$.

- n.b.: Above assumes pivots $\neq 0$; proof is tweakable.

General argument:

- Can see argument extends to *n* by *n*'s.
- Take $\mathbb{A} = \mathbb{A}^{\mathrm{T}} = \mathbb{L}\mathbb{D}\mathbb{L}^{\mathrm{T}}$ and smoothly change $\mathbb{L}$ to $\mathbb{I}$.
- Write $\hat{\mathbb{L}}(t) = \mathbb{I} + t(\mathbb{L} - \mathbb{I})$ and

$$\mathbb{B}(t) = \hat{\mathbb{L}}(t)\,\mathbb{D}\,\hat{\mathbb{L}}(t)^{\mathrm{T}}$$

- When $t = 1$, we have $\hat{\mathbb{L}}(1) = \mathbb{L}$ and $\mathbb{B}(1) = \mathbb{A}$.
- When $t = 0$, $\hat{\mathbb{L}}(0) = \mathbb{I}$, and $\mathbb{B}(0) = \mathbb{D}$.
- Again, pivots don't change as we move t from 1 to 0, and determinant must stay the same.
- Same story: eigenvalues cannot cross zero and must have the same signs for all t, including $t = 0$ when eigenvalues and pivots are equal $\mathbb{A} = \mathbb{D}$.

Further down the rabbit hole:

'Complete the square' for our first example:

$$f(x_1, x_2) = 2x_1^2 - 2x_1 x_2 + 2x_2^2$$

$$= 2(x_1^2 - x_1 x_2) + 2x_2^2 = 2(x_1^2 - x_1 x_2 + \frac{1}{4}x_2^2 - \frac{1}{4}x_2^2) + 2x_2^2$$

$$= 2(x_1^2 - x_1 x_2 + \frac{1}{4}x_2^2) - \frac{1}{2}x_2^2 + 2x_2^2 = 2(x_1 - \frac{1}{2}x_2)^2 + \frac{3}{2}x_2^2$$

- We see the pivots $d_1 = 2$ and $d_2 = \frac{3}{2}$ and the multiplier $\ell_{21} = -\frac{1}{2}$ appear:

$$f(x_1, x_2) = d_1(x_1 + \ell_{21}x_2)^2 + d_2 x_2^2.$$

- Super cool—this is exactly $\vec{x}^{\mathrm{T}}\mathbb{A}\vec{x} = (\mathbb{L}^{\mathrm{T}}\vec{x})\,\mathbb{D}\,(\mathbb{L}^{\mathrm{T}}\vec{x})^{\mathrm{T}} = d_1 z_1^2 + d_2 z_2^2$.
- The minimum is now obvious (sum of squares).

Another example:

- Take the matrix $\mathbb{A}_2$:

$$f(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

- Complete the square:

$$f(x_1, x_2) = 2x_1^2 - 2x_1x_2 - 2x_2^2 = 2(x_1 - \frac{1}{2}x_2)^2 - \frac{5}{2}x_2^2.$$

- Matches: Pivots $d_1 = 2$, $d_2 = -\frac{5}{2}$, so $x_1 = x_2 = 0$ is a saddle.
- Completing the square matches up with elimination...

Principle Axis Theorem:

Back to our second simple problem:

- Graph $2x_1^2 + 2x_1x_2 + 2x_2^2 = 1$.
- We'll simplify with linear algebra to find an equation of an ellipse...
- From before, our equation can be rewritten as

$$\vec{x}^{\mathrm{T}} \mathbb{A} \, \vec{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 1$$

- Again use spectral decomposition, $\mathbb{A} = \mathbb{Q} \, \Lambda \, \mathbb{Q}^{\mathrm{T}}$, to diagonalize giving $(\mathbb{Q}^{\mathrm{T}} \vec{x})^{\mathrm{T}} \Lambda \, (\mathbb{Q}^{\mathrm{T}} \vec{x}) = 1$ where

$$\mathbb{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \underbrace{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{\mathbb{Q}} \underbrace{\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}}_{\Lambda} \underbrace{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}}_{\mathbb{Q}^{\mathrm{T}}}$$

Principle Axis Theorem:

So $2x_1^2 + 2x_1x_2 + 2x_2^2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 1$

crazily becomes

$$\left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right)^{\mathrm{T}} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right)^{\mathrm{T}} = 1$$

$$: \begin{bmatrix} \frac{x_1+x_2}{\sqrt{2}} & \frac{x_1-x_2}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{x_1+x_2}{\sqrt{2}} \\ \frac{x_1-x_2}{\sqrt{2}} \end{bmatrix} = 1$$

$$:3 \left(\frac{x_1 + x_2}{\sqrt{2}} \right)^2 + \left(\frac{x_1 - x_2}{\sqrt{2}} \right)^2 = 1$$

Principle Axis Theorem:

If we change to eigenvector coordinate system,

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \mathbb{Q}^{\mathrm{T}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{x_1+x_2}{\sqrt{2}} \\ \frac{x_1-x_2}{\sqrt{2}} \end{bmatrix},$$

then our equation simplifies greatly:

$$\begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 1,$$

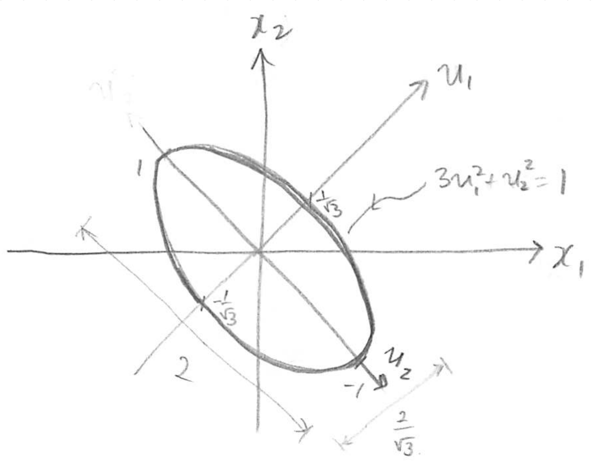
which is just

$$3 \cdot u_1^2 + 1 \cdot u_2^2 = 1.$$

Very nice! PDM : ellipse.

Principle Axis Theorem:

Finally, we can draw a picture of $2x_1^2 + 2x_1x_2 + 2x_2^2$:



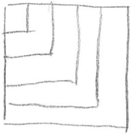
$\boxed{3 \cdot u_1^2 + 1 \cdot u_2^2 = 1}$ where $u_1 = \frac{x_1 + x_2}{\sqrt{2}}$ and $u_2 = \frac{x_1 - x_2}{\sqrt{2}}$.

Nutshell:

- $\vec{x}^{\mathrm{T}}\mathbb{A}\vec{x}$ is a commonly occurring construction.
- Big deals: Positive Definiteness and Semi-Positive Definiteness of $\mathbb{A}$.
- Positive eigenvalues : PDM.
- Non-negative eigenvalues : SPDM.
- Signs of pivots (easy test) match signs of eigenvalues.
- Gaussian elimination $\equiv$ completing the square.
- Standard questions: determine if a matrix is a PDM, convert a quadratic function into matrix $\vec{x}^{\mathrm{T}}\mathbb{A}\vec{x}$, sketch a quadratic curve (e.g., an ellipse).

Another connection:

ST #731:



For a real symmetric $\mathbb{A}$, if all upper left determinants of $\mathbb{A}$ are +ve, so are $\mathbb{A}$'s eigenvalues, and vice versa.

Check:

- $\mathbb{A}_1 : |2| > 0, \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3 > 0$: yes.
- $\mathbb{A}_2 : |2| > 0, \begin{vmatrix} 2 & -1 \\ -1 & -2 \end{vmatrix} = -5 < 0$: no.
- $\mathbb{A}_3 : |-2| < 0, \begin{vmatrix} -2 & -1 \\ -1 & -2 \end{vmatrix} = 3 > 0$: no.

Reasoning for 2×2 case:

- Take general symmetric matrix 2×2: $\mathbb{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$
- Upper left determinants: a and $ac - b^2$.
- Eigenvalues (from Assignment 9):

$$\lambda_1 = \frac{(a+c) + \sqrt{(a-c)^2 + 4b^2}}{2}$$

$$\lambda_2 = \frac{(a+c) - \sqrt{(a-c)^2 + 4b^2}}{2}$$

- Objective:
  show $a > 0$ and $ac - b^2 > 0 \Leftrightarrow \lambda_1, \lambda_2 > 0$.

Reasoning for 2×2 case:

Reuse previous sneakiness:

$$|\mathbb{A} - \lambda \mathbb{I}| = \begin{vmatrix} a-\lambda & b \\ b & c-\lambda \end{vmatrix} = (a-\lambda)(c-\lambda) - b^2$$

$$= \lambda^2 - (a+c)\lambda + ac - b^2$$

$$= \lambda^2 - \mathrm{Tr}(\mathbb{A}) + \det(\mathbb{A})$$

$$= \lambda^2 - (\lambda_1 + \lambda_2) + (\lambda_1 \cdot \lambda_2)$$

$$:\lambda_1 + \lambda_2 = a + c, \quad \lambda_1 \cdot \lambda_2 = ac - b^2$$

Show $a > 0,\ ac - b^2 > 0 \Leftrightarrow \lambda_1,\ \lambda_2 > 0$:

Show ":":

- Given $ac - b^2 > 0$ then $\lambda_1 \cdot \lambda_2 > 0$, so both eigenvalues are positive or both are negative.
- Given $a > 0$ then $c > 0$ b/c otherwise $ac - b^2 < 0$.
- This means $a + c = \lambda_1 + \lambda_2 > 0 \rightarrow$ both eigenvalues are positive.

Show "$\Leftarrow$":

- Given $\lambda_1,\ \lambda_2 > 0$, then $ac - b^2 = \lambda_1 \cdot \lambda_2 > 0$
- Know $a + c = \lambda_1 + \lambda_2 > 0$, so either $a,\ c > 0$, or one is negative.
- But again, $ac - b^2 > 0$ implies $a,\ c$ must have same sign, $\rightarrow a > 0$.

Finding PDMs...

- Upshot: We can compute determinants instead of eigenvalues to find signs.
- But: Computing determinants still isn't a picnic either...
- A much better way is to use the connection between pivots and eigenvalues.
- Another weird connection.

References I