Lecture 8 Linear Algebra Fall 2019

A note on the functional calculus: Is it possible to extend the functional calculus for continuous functions $f: \mathbb{C} \to \mathbb{C}$ and for diagonalizable matrices A given by (84.1), to all matrices A?

Consider the 2×2 case and suppose such an extended functional calculus existed. Let $A_h = \begin{pmatrix} 0 & 1 \\ 0 & h \end{pmatrix}$. For $h \in \mathbb{C}, h \neq 0$, A_h is diagonalizable

$$A_h = \begin{pmatrix} 1 & 1 \\ 0 & h \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & h \end{pmatrix} \begin{pmatrix} 1 & -h^{-1} \\ 0 & h^{-1} \end{pmatrix}$$

Now for any diagonalizable matrix B say, we have defined $f(B)$ as in (84.1), and so we should have

$$f(A_h) = \begin{pmatrix} 1 & 1 \\ 0 & h \end{pmatrix} \begin{pmatrix} f(0) & 0 \\ 0 & f(h) \end{pmatrix} \begin{pmatrix} 1 & -h^{-1} \\ 0 & h^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} f(0) & f(h) \\ 0 & h\, f(h) \end{pmatrix} \begin{pmatrix} 1 & -h^{-1} \\ 0 & h^{-1} \end{pmatrix}$$

(102.1) $$= \begin{pmatrix} f(0) & (f(h) - f(0))/h \\ 0 & f(h) \end{pmatrix}$$

Now as in Property (vi), we should have

$f(A_h) \to f(A_0) = f\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ as $h \to 0$. But from (102.1) we see that $\lim_{h\to 0} f(A_h)$ $\exists$ if and only if $f(z)$ is analytic at $z=0$. As A can be chosen arbitrarily, we see that $f(z)$ is necessarily entire!

Thus such a functional calculus $A \mapsto f(A)$ cannot exist for general, continuous $f: \mathbb{C} \to \mathbb{C}$ and <u>all</u> A.

We now return to the Rayleigh-Ritz Theorem 96.1

<u>Exercise 103.1</u> Apply Rayleigh-Ritz to $B = -A$ to conclude that if

$n \times n$

$$\mu_k \equiv \inf_{V_{k-1}} \sup_{\substack{u \perp V_{k-1} \\ \|u\|=1}} (u, Au)$$

then $\mu_1(A) \geq \mu_2(A) \geq \cdots \geq \mu_n(A)$ and, again, the μ_i's are the eigenvalues of A with appropriate multiplicities.

An immediate consequence of the Rayleigh-Ritz min-max theorem is that the k^{th} eigenvalue is a Lipschitz function of A, A Hermitian. In particular, $\lambda_k(A)$ is a continuous function of A, A Hermitian.

<u>Theorem 104.1</u> For Hermitian A, B

$$|\lambda_k(A) - \lambda_k(B)| \leq \|A - B\|$$

<u>Proof</u> Let $C \equiv A - B$. Then

$$(u, Au) = (u, Bu) + (u, Cu)$$

But $|(u, Cu)| \leq \|u\| \, \|Cu\| \leq \|C\| \, \|u\|^2 = \|A-B\| \, \|u\|^2$

Thus

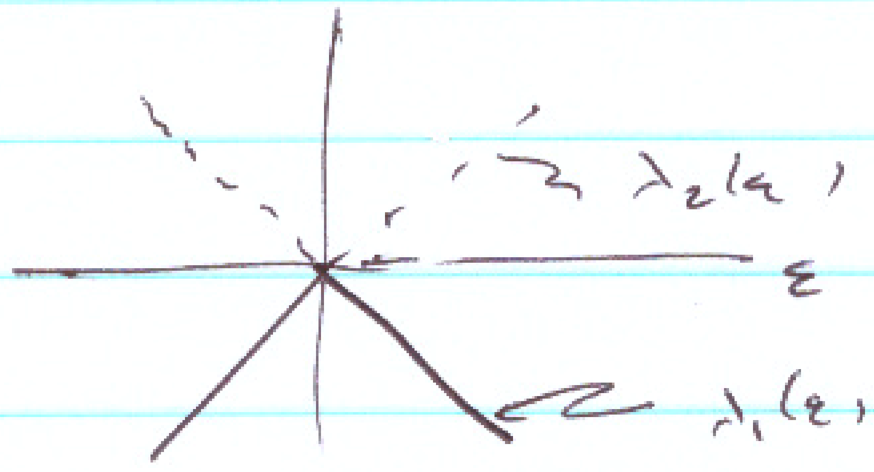
$$(u, Bu) - \|A - B\| \leq (u, Au) \leq (u, Bu) + \|A - B\|$$

for $\|u\| = 1$. Taking inf's over $V_{k-1}^{\perp}$, and then sup's over the V_{k-1}'s, we obtain the result. $\square$.

<u>Remark</u>: This Theorem does not mean that the eigenvalues as given by min-max $\lambda_k(A) = \sup_{V_{k-1}} \inf_{\substack{V_{k-1}^{\perp} \\ \|u\|=1}} (u, A)$ are differentiable functions of (the entries of) A. For example, consider

$$A_\varepsilon = \begin{pmatrix} 0 & \varepsilon \\ \varepsilon & 0 \end{pmatrix}, \quad \varepsilon \text{ real},$$

Have $\det(A_\varepsilon - \lambda I) = \lambda^2 - \varepsilon^2 \Rightarrow \lambda_1(\varepsilon) = -|\varepsilon|, \ \lambda_2(\varepsilon) = +|\varepsilon|$.



Thus $\lambda_1(\varepsilon)$ is not differentiable at $\varepsilon = 0$. But we do have

$$|\lambda_1(\varepsilon) - \lambda_1(0)| = |-|\varepsilon| - 0| = |\varepsilon| = \|A_\varepsilon - A_0\|$$

The following result is very useful.

Theorem 105.1

Let A be a Hermitian matrix. Then

(105.6) $$\|A\| = \max\{|\lambda_i| : \lambda_i \in \sigma(A)\}$$

Proof: By the spectral thm, $A = U \Lambda U^*$, U unitary and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Thus $$\|A\| = \sup_{\|u\|=1} \|U \Lambda U^* u\| = \sup_{\|u\|=1} \|\Lambda U^* u\|$$

$$= \sup_{\|U^* u\|=1} \|\Lambda U^* u\| \quad ; \text{ as } \|w\| = \|U^* u\| = 1,$$

$$= \sup_{\|w\|=1} \|\Lambda w\| = \sup_{\|w\|=1} \left(\sum_{i=1}^{n} \lambda_i^2 |w_i|^2\right)^{\frac{1}{2}}$$

$$= \max\{|\lambda_i| : \lambda_i \in \sigma(A)\}. \quad \square$$

<u>Remark</u> (105.6) fails for a general non-Hermitian matrix. Indeed if $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, then clearly

$\max_i \{|\lambda_i| : \lambda_i \in \sigma(A)\} = 0$. But $\|A\| = \sup_{\|u\|=1} \left\| \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \right\|$

$$= \sup_{\|u\|=1} |u_2| = 1$$

(*) Insert from 106+.

For general A we have

$$\sup_{\substack{\|u\|=1 \\ \|v\|=1}} |(u, Av)| = \sup_{\|v\|=1} \|Av\| = \|A\|$$

ie

(106.1) $$\|A\| = \sup_{\substack{\|u\|=1 \\ \|v\|=1}} |(u, Av)|$$

If A is Hermitian, we have a stronger result.

<u>Corollary to Th^m 105.1</u> For Hermitian A,

$$\|A\| = \sup_{\|u\|=1} |(u, Au)|$$

<u>Proof</u>:

By Th^m (105.1) and Rayleigh-Ritz

$$\|A\| = \max\left(\left|\sup_{\|u\|=1} (u, Au)\right|, \left|\inf_{\|u\|=1} (u, Au)\right| \right)$$

$$= \sup_{\|u\|=1} |(u, Au)| \quad \square.$$

(*) Insert on 106

Although (105.6) fails for general matrices A, we have the following result relating $\operatorname{sp} A$ to $\|A^n\|$ for general A. More precisely,

Thm Let A be an $k \times k$ matrix. Then

(106+.1) $$\sup\{|\lambda| : \lambda \in \sigma(A)\} = \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}}.$$

Furthermore $\|A^n\|^{\frac{1}{n}}$ is monotone decreasing so we in fact have

(106+.2) $$\sup\{|\lambda| : \lambda \in \sigma(A) = \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}} = \inf_n \|A^n\|^{\frac{1}{n}}$$

Thus in our example $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ on p 105,

we have $A^n = 0$ for $n \geq 2$ & so $\lim_{n\to\infty} \|A^n\|^{\frac{1}{n}} = 0$

$= \sup_\lambda \{|\lambda| : \lambda \in \sigma(A)\}$, as it should.

The example $A_\varepsilon = \begin{pmatrix} 0 & \varepsilon \\ \varepsilon & 0 \end{pmatrix}$, $\varepsilon \in \mathbb{R}$, illustrates the following <u>very</u> <u>important</u> fact that we will prove eventually: although $\lambda_1(\varepsilon)$, as defined by min-max, is <u>not</u> a differentiable function of ε, if we set

$$\tilde\lambda_1(\varepsilon) = \lambda_1(\varepsilon) \quad \text{for } \varepsilon \le 0$$
$$= \lambda_2(\varepsilon) \quad \text{for } \varepsilon > 0$$

then $\tilde\lambda_1(\varepsilon)$ is an eigenvalue of A_ε $\forall \varepsilon$, and is differentiable, in fact a real analytic, function of ε: indeed $\tilde\lambda_1(\varepsilon) = \varepsilon$.

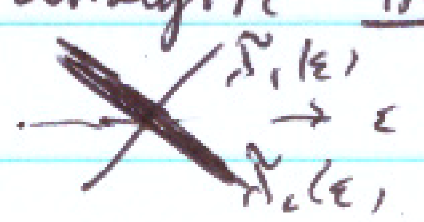
Thus if we want eigenvalues that are analytic functions of the parameter ε, we have to <u>give up</u> the order given by min-max.

We can also set

$$\tilde\lambda_2(\varepsilon) = \lambda_2(\varepsilon), \quad \varepsilon < 0$$
$$= \lambda_1(\varepsilon), \quad \varepsilon > 0$$

Then $\tilde\lambda_2(\varepsilon) = -\varepsilon$ is an analytic eigenvalue.

One says that the 2 analytic <u>branches</u> of the eigenvalues cross at $\varepsilon = 0$



What one does learn from Theorem 104.1, however, is that *if* $\lambda_k(\varepsilon)$, defined by min-max, is known to be differentiable at some ε, then

(108.1) $$|\lambda_k'(\varepsilon)| \le \| A'(\varepsilon)\|$$

We now prove the well-known, and very useful, interlacing theorem

Let A_n be an $n \times n$ Hermitian matrix: write

$$A_n = \begin{pmatrix} A_{n-1} & b \\ b^* & a_{nn} \end{pmatrix}$$

where A_{n-1} is $n-1 \times n-1$ and $b = (a_{1n} \cdots a_{n-1\,n})^T \in \mathbb{C}^{n-1}$. Suppose $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ are the eigenvalues of A and $\gamma_1 \le \gamma_2 \le \cdots \le \gamma_{n-1}$ are the eigenvalues of A_{n-1}.

(108.2) Interlacing Theorem I

The eigenvalues of A_n and A_{n-1} interlace, i.e.,

$$\lambda_1 \le \gamma_1 \le \lambda_2 \le \gamma_2 \le \cdots \le \lambda_{n-1} \le \gamma_{n-1} \le \lambda_n.$$

Proof: Let $V_{k-1} = \langle v_1, \ldots, v_{k-1} \rangle$ be a $(k-1)$-dimensional subspace of $\mathbb{C}^n$. Then

$$\inf_{\substack{v \perp V_{k-1} \\ \|v\|=1 \\ v \in \mathbb{C}^n}} (v, A_n v) \leq \inf_{\substack{\binom{u}{0} \perp V_{k-1} \\ \|\binom{u}{0}\|=1 \\ u \in \mathbb{C}^{n-1}}} \left(\binom{u}{0}, A_n \binom{u}{0} \right)$$

$$= \inf_{\substack{u \perp \tilde{V}_{k-1} \\ \|u\|=1 \\ u \in \mathbb{C}^{n-1}}} \left(\binom{u}{0}, \begin{pmatrix} A_{n-1} & b \\ b^* & a_{nn} \end{pmatrix} \binom{u}{0} \right)$$

where $\tilde{V}_{k-1}$ is the subspace of dimension $\leq k-1$ spanned by $u_1, \dots, u_{k-1}$, where $v_i = \binom{u_i}{v_{in}}$, u_i is of size $n-1$, v_{in} is a scalar

$$= \inf_{\substack{u \perp \tilde{V}_{k-1} \\ \|u\|=1 \\ u \in \mathbb{C}^{n-1}}} (u, A_{n-1} u)$$

$$\leq \gamma_k (A_{n-1}), \quad \text{as } \dim \tilde{V}_{k-1} \leq k-1.$$

Taking sup's over V_k's, we conclude that

(109.1) $$\lambda_k (A_n) \leq \gamma_k (A_{n-1})$$

Conversely, given any $k-1$ dimensional subspace $\tilde{V}_{k-1} = \langle u_1, \dots, u_{k-1} \rangle$ of $\mathbb{C}^{n-1}$, set

$$V_k = \left\langle \binom{u_1}{0}, \dots, \binom{u_n}{0}, e_n \right\rangle \quad \text{where } e_n = (0, \dots, 0, 1)^T.$$

Clearly V_k is a k-dimensional subspace of $\mathbb{C}^n$.
Then

$$\lambda_{k+1}(A_n) \geq \inf_{v \perp V_k, \|v\|=1, v \in \mathbb{C}^n} (v, A_n v)$$

$$= \inf_{\substack{\binom{u}{s} \perp V_k \\ \|\binom{u}{s}\| = 1 \\ u \in \mathbb{C}^{n-1} \\ s \in \mathbb{C}}} \left(\begin{pmatrix} u \\ s \end{pmatrix}, \begin{pmatrix} A_{n-1} & b \\ b^* & a_{nn} \end{pmatrix} \begin{pmatrix} u \\ s \end{pmatrix} \right)$$

But if $\binom{u}{s} \perp V_k$, then s must be 0, and so $u \perp \tilde{V}_{k-1}$. Conversely if $u \perp \tilde{V}_{k-1}$, then $\binom{u}{0} \in V_k^{\perp}$.

Thus $\lambda_{k+1}(A_n) \geq \inf_{\substack{u \perp \tilde{V}_{k-1} \\ \|u\| = 1 \\ u \in \mathbb{C}^{n-1}}} \left(\begin{pmatrix} u \\ 0 \end{pmatrix}, \begin{pmatrix} A_{n-1} & b \\ b^* & a_{nn} \end{pmatrix} \begin{pmatrix} u \\ 0 \end{pmatrix} \right)$

$$= \inf_{\substack{u \perp \tilde{V}_{k-1} \\ \|u\| = 1 \\ u \in \mathbb{C}^{n-1}}} (u, A_{n-1} u)$$

Taking sup's over $\tilde{V}_{k-1}$'s, we obtain $\lambda_{k+1}(A_n) \geq \gamma_k(A_{n-1})$ or

$$\lambda_k(A_n) \geq \gamma_{k-1}(A_{n-1}).$$

This proves the interlacing Theorem. $\square$

There is another very useful and interlacing Theorem

We say that an $n \times n$ matrix is of <u>rank 1</u> if there are two vectors t and s, $t, s \in \mathbb{C}^n$, $t \neq 0$, $s \neq 0$.

such that

$$A = t s^*$$

Thus for any $u \in \mathbb{C}^n$,

$$Au = t\, s^* u = (s, u)\, t$$

Sometimes we write

$$A = (s, \cdot)\, t\ .$$

Clearly $A^* = (t, \cdot)\, s$ Thus A is rank 1 and

Hermitian iff $A = c\,(s, \cdot)\, s$ for some real $c \neq 0$.

(111.1) Interlacing Theorem II

Suppose A and B are $n \times n$ Hermitian matrices

and suppose that $A - B$ is of rank 1. Then

the eigenvalues of A and B interlace i.e.

either $\lambda_1(A) \le \lambda_1(B) \le \lambda_2(A) \le \lambda_2(B) \le \cdots \le \lambda_n(A) \le \lambda_n(B)$

or $\lambda_1(B) \le \lambda_1(A) \le \lambda_2(B) \le \lambda_2(A) \le \cdots \le \lambda_n(B) \le \lambda_n(A)$

The first case corresponds to $A - B \le 0$, the second case to $A - B \ge 0$.

Proof: The proof is a simpler version of the proof of the previous interlacing Theorem, and is left as an <u>Exercise</u>. □

By induction, we immediately have the following

(112.1) <u>Corollary</u>

If A, B are Hermitian and $A-B$ is rank k, i.e. the sum of k independent rank 1 operators, then between any $k+1$ eigenvalues of A there must be an eigenvalue of A, and vice versa.

We now give a <u>geometric</u> proof of Interlacing theorem I. By the spectral theorem for A_{n-1}, we have $A_{n-1} = U \gamma U^*$ where U is unitary in $\mathbb{C}^{n-1}$ and $\gamma = \operatorname{diag}(\gamma_1, \dots, \gamma_{n-1})$ are the eigenvalues of A_{n-1}.

We have

$$A_n - \lambda = \begin{pmatrix} A_{n-1}-\lambda & b \\ b^* & a_{nn}-\lambda \end{pmatrix} = \begin{pmatrix} U & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma-\lambda & c \\ c^* & a_{n}-\lambda \end{pmatrix} \begin{pmatrix} U^* & 0 \\ 0 & 1 \end{pmatrix}$$

where

$$b = Uc \qquad \text{ie } c = U^* b$$

Thus for $\lambda \notin \{\gamma_1, \dots, \gamma_{n-1}, a_{nn}\}$

$$\det(A_n - \lambda) = \det \begin{pmatrix} \gamma-\lambda & c \\ c^* & a_n-\lambda \end{pmatrix} = \det\left[\begin{pmatrix} \gamma-\lambda & 0 \\ 0 & a_n-\lambda \end{pmatrix} + \begin{pmatrix} 0 & c \\ c^* & 0 \end{pmatrix} \right]$$

$$= \det\left[\left(I + \begin{pmatrix} 0 & c \\ c^* & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\gamma-\lambda} & 0 \\ 0 & \frac{1}{a_{nn}-\lambda} \end{pmatrix} \right) \begin{pmatrix} \gamma-\lambda & 0 \\ 0 & a_{nn}-\lambda \end{pmatrix} \right]$$

$$= \left[(a_{nn}-\lambda)\prod_{i=1}^{n-1}(\gamma_i-\lambda)\right] \det\left(I + \begin{pmatrix} 0 & \frac{c}{a_{nn}-\lambda} \\ c^* \frac{1}{\gamma-\lambda} & 0 \end{pmatrix}\right)$$

$$= \left[(a_{nn}-\lambda)\prod_{i=1}^{n-1}(\gamma_i-\lambda)\right] \det\begin{pmatrix} 1 & & & 0 & \frac{c_1}{a_{nn}-\lambda} \\ & 1 & & & \vdots \\ & & \ddots & & \vdots \\ & 0 & & \ddots & \frac{c_{n-1}}{a_{nn}-\lambda} \\ \frac{\bar{c}_1}{\gamma_1-\lambda} & \cdots & & \frac{\bar{c}_{n-1}}{\gamma_{n-1}-\lambda} & 1 \end{pmatrix}$$

$$= (a_{nn}-\lambda)\prod_{i=1}^{n-1}(\gamma_i-\lambda)\left(1 - \frac{1}{a_{nn}-\lambda}\sum_{i=1}^{n-1}\frac{|c_i|^2}{\gamma_i-\lambda}\right)$$

where we have used the fact that for any $(\alpha_1,\dots,\alpha_{n-1})$ and $(\beta_1,\dots,\beta_{n-1})$

$$\det\begin{pmatrix} 1 & 0 & \cdots & \alpha_1 \\ & 1 & 0 & \vdots \\ 0 & & \ddots & \alpha_{n-1} \\ \beta_1 & \cdots & \beta_{n-1} & 1 \end{pmatrix} = 1 - \sum_{i=1}^{n-1}\alpha_i\beta_i$$

which is easily proved by induction (exercise).

Thus

$$\det(A_n-\lambda) = \underbrace{\prod_{i=1}^{n-1}(\gamma_i-\lambda)}_{\det(A_{n-1}-\lambda)}\left(a_{nn}-\lambda - \sum_{i=1}^{n-1}\frac{|c_i|^2}{\gamma_i-\lambda}\right) \tag{113.1}$$

Set

$$f(\lambda) = a_{nn}-\lambda - \sum_{i=1}^{n-1}|c_i|^2/\gamma_i-\lambda$$

Now assume that the δ_i's are distinct and the c_i's are non-zero. Clearly (<u>Exercise</u>) such matrices are dense in the Hermitian $n \times n$ matrices.

Now plot the function $f(\lambda)$: it must look like this

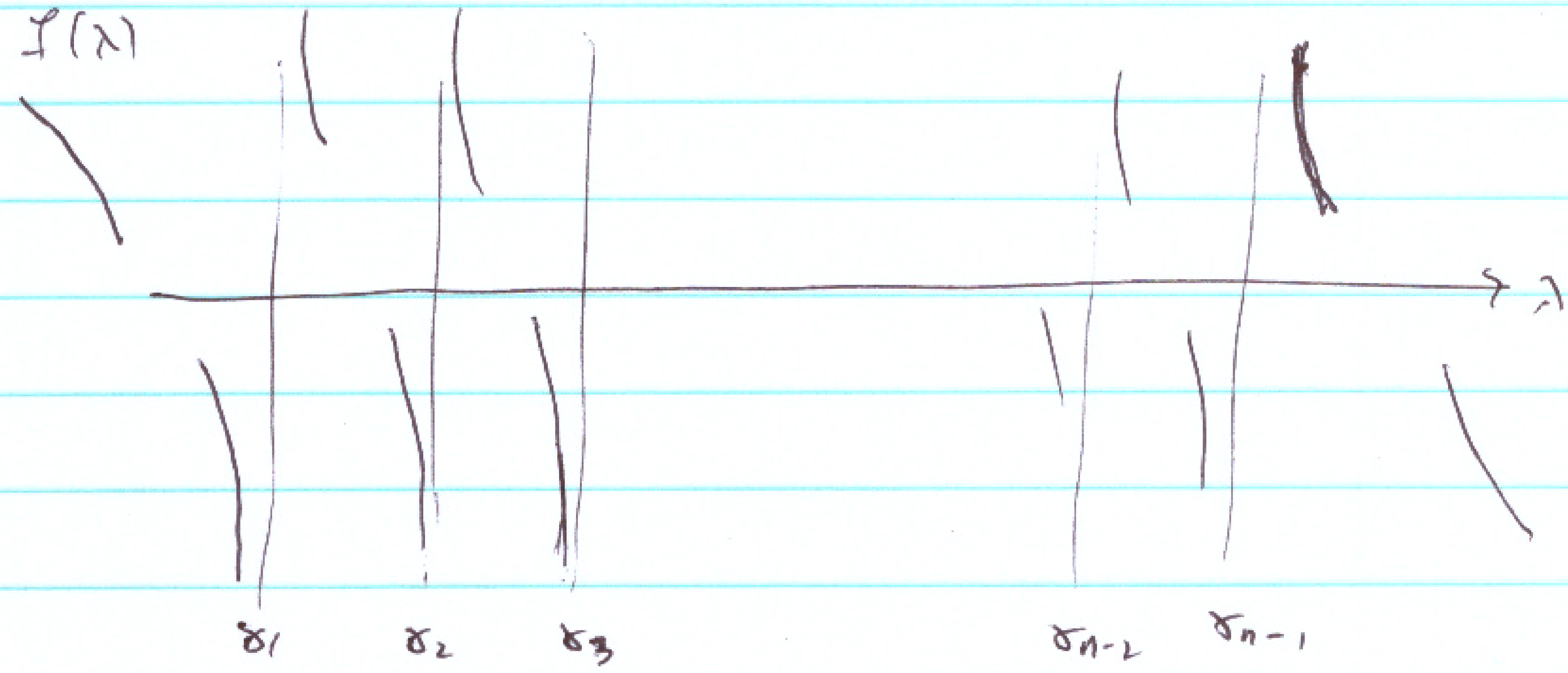


Fig 114.1

near $-\infty, \delta_1, \delta_2, \ldots, \delta_{n-1}$, and $+\infty$.

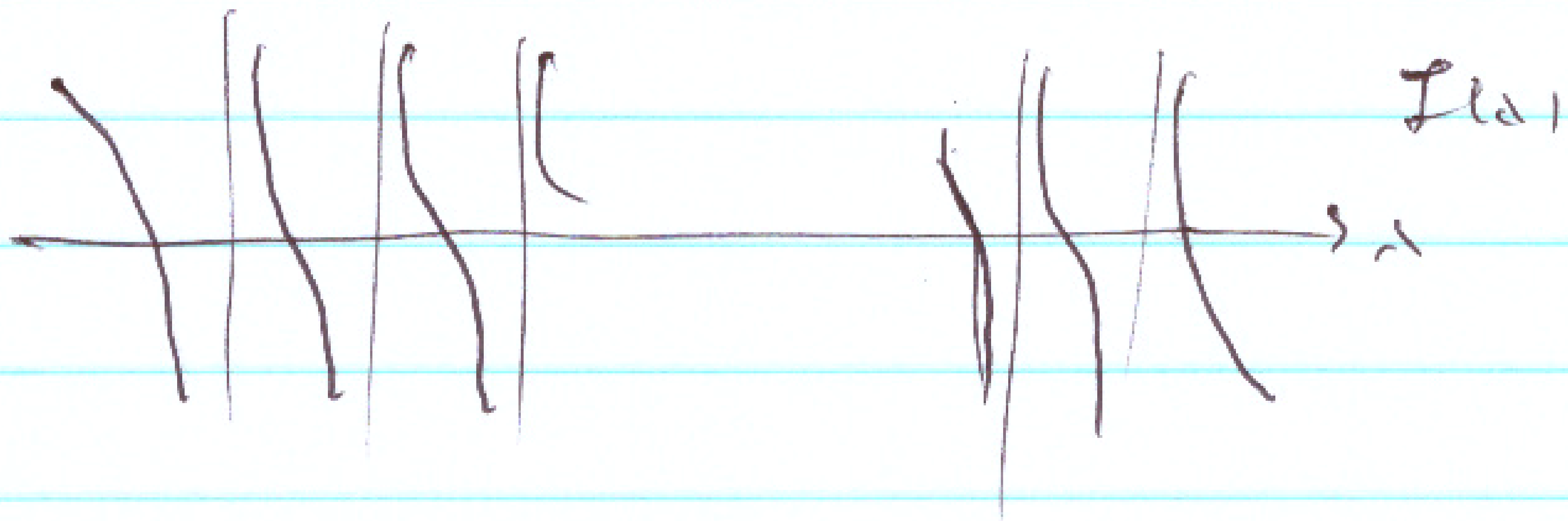
By continuity $f(\lambda)$ must then have zeros in

$$(-\infty, \delta_1), (\delta_1, \delta_2), \ldots, (\delta_{n-2}, \delta_{n-1}), (\delta_{n-1}, \infty).$$

Thus, by (113.1), $\det(A_n - \lambda)$ must have roots in each of these n intervals. But A_n has only n eigenvalues, so these roots are the eigenvalues and the interlacing is clear.

Note that in the intervals, $f'(\lambda) = -1 - \sum_{i=1}^{n-1} |c_i|^2 / |\delta_i - \lambda|^2 < 0$

so that $f(\lambda)$ is monotonic in each interval: In particular it shows explicitly that $f(\lambda)$ has <u>only</u> one root in each interval. Also Fig 114.1 should be completed to



<u>Exercise</u> Give a similar geometric proof of Interlacing Theorem II.

→ $\boxed{\text{Insert 115+}}$

We now prove a famous Theorem of Sylvester.

Given an $n \times n$ Hermitian matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdot & a_{1n} \\ a_{21} & a_{22} & & \\ & & \cdot & \\ a_{n1} & \cdots & & a_{nn} \end{pmatrix}, \quad a_{ij} = \overline{a_{ji}}$$

set $d_1 = a_{11}$, $d_2 = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \ldots, d_n = \det A$

The d_i's are called the principal minors of A. Note that as A is Hermitian, all the d_i's are real.

<u>Theorem 115.1</u>

A is strictly positive definite i.e. $\langle u, Au \rangle > 0 \ \forall u \neq 0 \iff d_i > 0, \ i = 1, \ldots, n$.

Insert on p115

Another consequence of the min-max theorem is the following.

Recall $A = A^*$ is positive definite, written $A \geq 0$, if $(u, Au) \geq 0$, and

strictly positive definite, written $A > 0$, if $(u, Au) > 0$.

for all $u \neq 0$.

Theorem 115+.1

A is positive definite $\Leftrightarrow$ all the eigenvalues λ_i of A are ≥ 0

A is strictly positive definite $\Rightarrow$ all the eigenvalues of A are > 0.

Proof If $A \geq 0$, $\lambda_i \geq 0$ by the min-max theorem.
Conversely, by the spectral theorem $A = \sum_{i=1}^{n} \lambda_i (u_i, \cdot) u_i$
where u_i are the normalized eigenvectors of A.
If $\lambda_i \geq 0$, $(u, Au) = \sum \lambda_i |(u_i, u)|^2 \geq 0 \ \forall u$.
The strictly pos. def. case is similar. $\square$.

<u>Proof:</u> Suppose A is strictly positive definite. Then for any $1 \le k \le n$, and any $u = (u_1, \dots, u_k) \in \mathbb{C}^k$, $u \neq 0$,

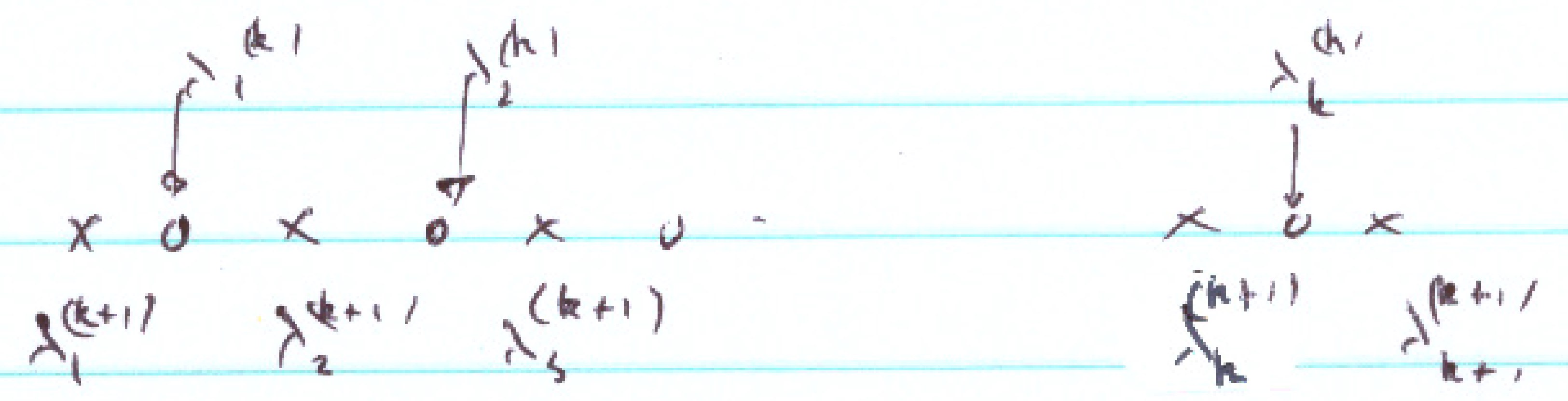
$$0 < \Big(\underset{\in \mathbb{C}^n}{\begin{pmatrix} u \\ 0 \end{pmatrix}}, A \begin{pmatrix} u \\ 0 \end{pmatrix} \Big) = \left(u, \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} u \right)$$

Hence by min-max, all the eigenvalues of $\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$, $\lambda_1^{(k)}, \dots, \lambda_k^{(k)}$, are strictly positive. But then

$$d_k = \lambda_1^{(k)} \cdots \lambda_k^{(k)} > 0, \qquad 1 \le k \le n.$$

Conversely, suppose that $d_k > 0$, $k = 1, \dots, n$.

Assume by induction that all the eigenvalues of $\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$ are strictly positive. This is clearly true for $k = 1$. Now by the interlacing theorem the eigenvalues of $\begin{pmatrix} a_{11} & \cdots & a_{1k+1} \\ a_{k+1,1} & & a_{k+1,k+1} \end{pmatrix}$ interlace with those of $\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$.



It clearly follows that $\lambda_i^{(k+1)} > 0$, $i = 2, \dots, k+1$

But $\lambda_1^{(k+1)} = \dfrac{d_{k+1}}{\lambda_2^{(k+1)} \cdots \lambda_{k+1}^{(k+1)}} > 0$

Hence all the eigenvalues of $\begin{pmatrix} a_{11} & \cdots & a_{1k+1} \\ a_{k+1,1} & \cdots & a_{k+1,k+1} \end{pmatrix}$

are strictly positive. This completes the induction and hence all the eigenvalues of A are strictly positive, which proves that A is strictly positive. $\square$.

Remark:

Note that the statement

$$A \text{ pos. def} \iff d_i \geq 0, \; i=1,\dots,n$$

is not true. All that is true is that A pos. def $\Rightarrow$ $d_i \geq 0$, $i=1,\dots,n$ (same proof as before). However, consider

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Here d_1, d_2 and $d_3 = 0$, but $\det(A-\lambda) = \lambda(1-\lambda^2) = 0$ $\Rightarrow$ $\lambda = 0, 1$ or -1 so that $A \not\geq 0$. (Exercise: what goes wrong in the proof if we try to show $d_i \geq 0 \Rightarrow A \geq 0$?).