Lecture 3

Let $S_n = \{M\}$ denote the set of real $n \times n$ symmetric matrices. Suppose $M \in S_n$ has eigenvalues $\{\lambda_j\}_{j=1}^{n}$ and associated eigenvectors $u(\lambda_j) = (u_{1j}, \dots, u_{nj})^T$.

Let $A_n = \{ M \in S_n :$ for all eigenvectors $u = (u_1, \dots, u_n)^T$ of M, $u_1 \neq 0 \}$

Claim 1:

A_n is a dense, open set in S_n of full measure, i.e., meas$(S_n \setminus A_n) = 0$.

Note that $A_n \subset \{ M \in S_n : \lambda_j(M) \neq \lambda_k(M),\ j \neq k \}$ — (42.1)

Indeed if $M \in A_n$ and $\lambda_j(M) = \lambda_k(M) \equiv \lambda$, then associated with λ there at least two independent eigenvectors

$M u(\lambda_j) = \lambda_j u(\lambda_j) = \lambda u(\lambda_j)$, $M u(\lambda_k) = \lambda_k u(\lambda_k) = \lambda u(\lambda_k)$, from which it follows that there is an eigenector $u \neq 0$, $Mu = \lambda u$, with $u(1) = 0$. This is a contradiction and so (42.1) is true.

Let $C_n = A_n' = S_n \setminus A_n$.

For an $n \times n$ matrix M, let M_1 denote the $(n-1)\times(n-1)$ matrix obtained by deleting the first row and column of M.

Claim 2

$C_n = D_n = \{ M \in S_n : M \text{ and } M_1 \text{ have a common eigenvalue}. \}$

Clearly $C_n \subset D_n$ for if $Mu = \lambda u$, $u = (u_1, u_2, \ldots, u_n)^T$, with $u_1 = 0$, then $v = (u_2, \ldots, u_n)^T \neq 0$ is an eigenvector for M_1, $M_1 v = \lambda v$. Conversely, suppose that $M \in D_n$, but $M \notin C_n$. Let λ be a common eigenvalue of M and M_1. Then $(M - \lambda)u = 0$ for some $u = (u_1, \ldots, u_n)^T$ and as $M \notin C_n$, $u_1 \neq 0$. Without loss, suppose $u_1 = -1$. Write

$$M - \lambda = \begin{pmatrix} a & b^T \\ b & M_1 - \lambda \end{pmatrix}$$

where $a \in \mathbb{R}$ and $b \in \mathbb{R}^{n-1}$. As $(M - \lambda)u = 0$, we have in particular,

$$b = (M_1 - \lambda)\begin{pmatrix} u_2 \\ \vdots \\ u_n \end{pmatrix}$$

and so

$$b^T = (u_2, \dots, u_n)(M_1 - \lambda)^T$$

$$= (u_2, \dots, u_n)(M_1 - \lambda)$$

Now $\exists\, w = (w_2, \dots, w_n)^T \neq 0$ such that

$$(M_1 - \lambda) w = 0$$

and so

$$b^T w = 0$$

But then

$$(M - \lambda)\begin{pmatrix} 0 \\ w \end{pmatrix} = \begin{pmatrix} a & b^T \\ b & M_1 - \lambda \end{pmatrix}\begin{pmatrix} 0 \\ w \end{pmatrix} = 0 .$$

and so M has an eigenvector $u = \begin{pmatrix} 0 \\ w \end{pmatrix}$ with $u_1 = 0$. This is a contradiction and so $D_n \subset C_n$. $\square$

<u>Note:</u> By the proof of Claim 2, we in fact see that if $M \in D_n$, then M and M_1 have a "common" eigenvector i.e. an eigenvector $w = (w_2, \dots, w_n)^T$ for M_1, $M_1 w = \lambda w$, such that $v = (0, w_2, \dots, w_n)$ is an eigenvector for M, $Mv = \lambda v$. Note however that not every eigenvector w of M_1 has the property that $v = \begin{pmatrix} 0 \\ w \end{pmatrix}$ is an eigenvector of M. For example

if $\quad M = \begin{pmatrix} a & b & c \\ b & 0 & 0 \\ c & 0 & 0 \end{pmatrix}, \quad b^2 + c^2 \neq 0, \quad$ then

$M_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} u_2 \\ u_3 \end{pmatrix}$ is an eigenvector of M_1. However

$\begin{pmatrix} 0 \\ u_2 \\ u_3 \end{pmatrix}$ is an eigenvector of M if and only if $\begin{pmatrix} u_2 \\ u_3 \end{pmatrix}$ is a multiple

of $\begin{pmatrix} -c \\ b \end{pmatrix}$.

We now show that D_n has measure 0. Let

$$p(\lambda) = \det(\lambda - M)$$
$$q(\lambda) = \det(\lambda - M_1)$$

Write

$$p(\lambda) = \lambda^n + p_{n-1}\lambda^{n-1} + \dots + p_0$$
$$q(\lambda) = \lambda^{n-1} + q_{n-2}\lambda^{n-2} + \dots + q_0$$

and consider the <u>resultant</u> R of p and q

$$R = \det \begin{pmatrix} p_0 & p_1 & \cdots & p_{n-1} & 1 & & O \\ 0 & p_0 & \cdots & \cdots & p_{n-1} & 1 & \\ & & \ddots & & & & \\ & O & & \ddots & & & \\ & & & p_0 & \cdots & \cdots & 1 \\ q_0 & q_1 & \cdots & q_{n-2} & 1 & & O \\ 0 & q_0 & \cdots & \cdots & & 1 & \\ & & \ddots & & & & \\ & O & & \ddots & & & \\ & & & q_0 & \cdots & & 1 \end{pmatrix}$$

where $(p_0 \cdots p_{n-1} \; 1)$ is repeated $n-1$ times

and $(q_0 \cdots q_{n-2} \; 1)$ is repeated n times. Clearly R is a determinant of size $n+(n-1) = 2n-1$. By standard theory (exercise) for any 2 polynomials p and q

$$R = 0 \iff p(\lambda) \text{ and } q(\lambda) \text{ have a common root}$$

Now $R = R(M)$ is a real analytic function (in fact a polynomial) in the entries of M and hence if it vanishes on a set of positive measure in $\mathbb{R}^{\frac{n(n+1)}{2}} \cong S_n$, it is identically zero (exercise). But

$$M = \begin{pmatrix} 1 & 1 & \cdots & \cdots & 1 \\ 1 & 2 & & 0 & \\ \vdots & & 3 & & \\ & & & \ddots & \\ 1 & & 0 & & n \end{pmatrix}$$

does not have a common eigenvalue with $M_1 = \text{diag}(2,3,\ldots,n)$, spec $M_1 = \{2,3,\ldots,n\}$. Indeed, if $Mu = ju$, $u = (u_1,\ldots,u_n)^T \neq 0$, for $2 \le j \le n$, then $0 = (e_j, (M-j)u) = u_1$. But the $(k-j)u_k = 0$, $2 \le k \le n$, $k \neq j$; so $u = (0 \cdots 0 \, u_j \, 0 \cdots 0)^T$. But then from the first row of $(M-j)$ we see that $u_j = 0$. Thus $R \not\equiv 0$. Hence we must have $\text{meas}(D_n) = 0$.

<u>example</u> Suppose

$$p(z) = p_0 + p_1 z + z^2$$

$$q(z) = q_0 + z$$

$$2n-1 = 4-1 = 3$$

Then

$$R = \det\begin{pmatrix} p_0 & p_1 & 1 \\ q_0 & 1 & 0 \\ 0 & q_0 & 1 \end{pmatrix} = p_0 - p_1 q_0 + q_0^2$$

So if $z = -q_0$ is a root of q and also of

$p(z) = p_0 + p_1(-q_0) + (-q_0)^2$, we see that $R = 0$.

<u>Exercise</u> Generalize $R(p,q)$ for any 2 polynomials of arbitrary order.

Exercise Instead of R, we can consider associated tensors.

For a polynomial $p(z) = p_0 + p_1 z + \cdots + p_{n-1} z^{n-1} + z^n$ the companion matrix for p is defined as follows:

$$C_p = \begin{pmatrix} 0 & - & - & - & 0 & -p_0 \\ 1 & 0 & - & - & 0 & -p_1 \\ 0 & 1 & & & & \vdots \\ & & \ddots & & & \vdots \\ & 0 & & & 1 & -p_{n-1} \end{pmatrix} \qquad C_p \text{ is } n \times n$$

Then (exercise) $p(z) = 0$ if and only if z is an eigenvalue of C_p ie $\det(z - C_p) = 0$.

If $q(z) = q_0 + \cdots + q_{m-1} z^{m-1} + z^m$ is a second polynomial, with companion matrix C_q, C_q is $m \times m$, set

$$A = C_p \otimes I_m - I_n \otimes C_q$$

Show det

$$\det A = \prod_{\substack{1 \le i \le n \\ 1 \le j \le m}} (\lambda_i - \mu_j)$$

and so $\det A = 0 \iff p$ and q have a common root

If $p(z) = \det(z - M_1)$, $q(z) = \det(z - M_2)$, we see that $\det A$ is a polynomial in the entries of M, and the argument proceeds as before for R.

Finally we show that A_n is dense and open, which completes the proof of Claim 1.

If $M_0 \in A_n$, then M_0 has distinct spectrum $\lambda_1 < \lambda_2 < \dots < \lambda_n$, (see p42) and $u_1 \neq 0$ for all eigenvectors $u = (u_1, \dots, u_n)^T$ of M_0.

But as the spectrum of M_0 is simple it follows by standard spectral theory that the eigenvalues and eigenvectors are continuous for M in a neighborhood of M_0. This shows that A_n is open. On the other hand, as meas $A_n' = 0$, A_n is certainly dense. We are done.

It follows from the above calculations that the Spectral Theorem

$$M = O \Lambda O^T, \quad \Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_N)^T$$

induces a well-defined and smooth map

$$\varphi : M \mapsto (\lambda_1 < \dots < \lambda_N, O)$$

from A_N into $\Lambda_N^{\uparrow} \times O_1^+$

where O_1^+ is the set of orthogonal matrices whose columns

$o_j = Oe_j$ have positive first entries, $o_j(1) > 0$,

$j = 1, \dots, N$, an

and where $\Lambda_N^{\uparrow}$ is the vector in $\mathbb{R}^N$

$$\{\lambda_1 < \dots < \lambda_N\}.$$

In fact φ is a bijection with a smooth

inverse φ^{-1}: Indeed if $\varphi(M) = \varphi(\tilde{M})$

then $\tilde{\lambda}_i = \lambda_i$ $i = 1, \dots, N$ and $\tilde{O} = O$ and

so

$$M = O \Lambda O^T = \tilde{O} \tilde{\Lambda} \tilde{O}^T = \tilde{M}$$

so φ is $1 \to 1$. Also if $\Lambda = \{\lambda_1 < \dots < \lambda_N\} \in \Lambda_N^{\uparrow}$

and $O \in O_1^+$ then $M = O \Lambda O^T$

is a real symmetric matrix with simple spectrum

Moreover the columns of O are clearly the eigenvectors of M and $Oe_j(1) \neq 0$. Hence $M \in A_n$ and clearly $\varphi(M) = (\Lambda, O)$. Thus φ is a bijection. Clearly the inverse of φ is given by

$$\varphi^{-1} : (\Lambda, O) \to O \Lambda O^T$$

and φ^{-1} is smooth.

Now as A_n is an open set of full measure, we can use φ as a change of variables to compute probabilities. The first order of business is to compute the Jacobian

$$\det \left(\frac{\partial (M)}{\partial (\Lambda, O)} \right)$$

on A_n.

Fix $M_0 \in A_n$ and let $(\Lambda_0, O_0) = \varphi(M_0)$

Let $p = (p_1, \dots, p_\ell)$, $\ell = \frac{N(N-1)}{2}$, be local

co-ordinates on O_1^+ in a small neighborhood of

O_0, $\sum_{i=1}^{\ell} p_i^2 < \varepsilon^2$.

$$(p_1, \dots, p_\ell) \longmapsto O(p_1, \dots, p_\ell)$$

where $(0, \dots, 0) \mapsto O_0$

For ~~just~~ sufficiently small $\varepsilon > 0$, $\lambda_1, \dots, \lambda_N$, $p_1, \dots, p_\ell$ with

$$\lambda_1 < \dots < \lambda_N, \quad \sum_{i=1}^{\ell} p_i^2 < \varepsilon^2$$

are co-coordinates ~~for~~ an open nbhood of M_0

$$O_{M_0} = \varphi^{-1}\left((\lambda_1 < \dots < \lambda_N), \sum_{i=1}^{\ell} p_i^2 < \varepsilon^2 \right)$$

$$= \left\{ M = M(\Lambda, p) = O(p_1, \dots, p_\ell) \Lambda O^T(p_1, \dots, p_\ell), \; \lambda_1 < \dots < \lambda_N, \sum_{i=1}^{\ell} p_i^2 < \varepsilon^2 \right\},$$

Differentiating $M = M(\Lambda, p) = O(p) \Lambda O(p)^T$ wrt p_j,

we find

$$M_{p_j} = O_{p_j} \Lambda O^T + O \Lambda O_{p_j}^T, \quad 1 \le j \le \ell.$$

But $O^T O = I$. Hence $O_{p_j}^T O + O^T O_{p_j} = 0$ and so

(53.1) $$S_j = O^T O_{p_j}$$

is skew-symmetric. It follows that for $1 \le j \le \ell$

(53.2) $$O^T M_{p_j} O = S_j \Lambda - \Lambda S_j = [S_j, \Lambda]$$

Similarly

(53.3) $$O^T M_{\lambda_j} O = \Lambda_{\lambda_j} \qquad j = 1, \dots, \ell$$

Consider the map

(53.4) $$V_O : C \mapsto V_O(C) = O^T C O$$

mapping real symmetric matrices to real symmetric matrices. Let $\mathcal{J}$ denote the bijection of real symmetric $N \times N$ matrices onto $\mathbb{R}^{N(N+1)/2}$ given by (cf $M \mapsto \tilde{M}$, p.29)

$$\mathcal{J} : M \mapsto (M_{11}, \dots, M_{N,N}, M_{12}, M_{13}, \quad , M_{N-1,N})^T$$

$$\equiv \vec{M} = \mathcal{J}(M)$$

Define the inner product on $\mathbb{R}^{N(N+1)/2}$ as before by

(54)

(54.1) $$(\vec{M}, \vec{M})_{tr} \equiv \operatorname{tr} M^2$$

where $\vec{M} = \mathcal{F}(M)$. We find for $M = M^T = \overline{M}$,

$$(\overrightarrow{V_0 M}, \overrightarrow{V_0 M})_{tr} = \operatorname{tr}(V_0 M)^2$$

$$= \operatorname{tr}(O^T M^2 O)$$

$$= \operatorname{tr} M^2$$

$$= (\vec{M}, \vec{M})_{tr}$$

It follows that

(54.2) $$\vec{V}_0(\vec{M}) \equiv \overrightarrow{V_0(M)}$$

where $\vec{M} = \mathcal{F}(M)$ is orthogonal and as before

(54.3) $$\det \vec{V}_0 = \pm 1$$

More precisely, if the matrix T_0 represents $\vec{V}_0$ in an orthonormal basis $\{e_i\}_{i=1}^{\frac{N(N+1)}{2}}$ for $(\mathbb{R}^{\frac{N(N+1)}{2}}, (\cdot,\cdot)_{tr})$, ie $\hat{V}_0 e_j = \sum_{i=1}^{N(N+1)/2} (T_0)_{ij} e_i$, then T_0 is orthogonal and so $\det T_0 = \pm 1$. Now (53.2) (53.3) can be written in the form

(54.4) $$V_0(M_{\lambda_1}, \ldots, M_{\lambda_N}, M_{p_1}, \ldots, M_{p_\ell})$$

$$= (\Lambda_{\lambda_1}, \ldots, \Lambda_{\lambda_N}, [S_1, \Lambda], \ldots, [S_\ell, \Lambda])$$

where the elements on the RHS are $N \times N$ matrices,

or

(55.1) $$\vec{V}_0\left(\vec{M}_{\lambda_1}, \vec{M}_{\lambda_2}, \ldots, \vec{M}_{p_1}, \ldots, \vec{M}_{p_\ell}\right)$$

$$= \left(\vec{\Lambda}_{\lambda_1}, \ldots, \vec{\Lambda}_{\lambda_N}, \overrightarrow{[S_1, \Lambda]}, \ldots, \overrightarrow{[S_\ell, \Lambda]}\right).$$

where the elements on the RHS are column vectors

of size $N(N+1)/2$. Hence by (54.3) and the representation $\vec{V}_0(\vec{M}) = T_0 \vec{M}$,

(55.2) $$\left|\det\left(\vec{M}_{\lambda_1}, \ldots, \vec{M}_{\lambda_N}, \vec{M}_{p_1}, \ldots, \vec{M}_{p_\ell}\right)\right|$$

$$= \left|\det\left(\vec{\Lambda}_{\lambda_1}, \ldots, \vec{\Lambda}_{\lambda_N}, \overrightarrow{[S_1, \Lambda]}, \ldots, \overrightarrow{[S_\ell, \Lambda]}\right)\right|$$

Now

(55.3) $$\vec{\Lambda}_{\lambda_j} = (0, \ldots, 0, 1, 0, \ldots, 0)^T \in \mathbb{R}^{N(N+1)/2}.$$

where 1 is at the j^{th} place. Also $[S_q, \Lambda]_{ij} = (\lambda_j - \lambda_i)(S_q)_{ij}$

for $q = 1, \ldots, \ell$, and so for $1 \le q \le \ell$

(55.4) $$\overrightarrow{[S_q, \Lambda]} = \left(0, \ldots, 0, (\lambda_2 - \lambda_1)(S_q)_{12}, (\lambda_3 - \lambda_1)(S_q)_{13}, \ldots, (\lambda_N - \lambda_{N-1})(S_q)_{N-1,N}\right)$$

with 0's in the first N entries. Thus graphically we have

(56.0)
$$\left(\vec{\Lambda}_{\lambda_1}, \ldots, \vec{\Lambda}_{\lambda_N}, \overrightarrow{[S_1, \Lambda]}, \ldots, \overrightarrow{[S_q, \Lambda]}\right) = \left(\begin{array}{c|c} I_N & 0 \\ \hline 0 & X \end{array}\right)$$

where X denotes the $\frac{N(N-1)}{2} \times \frac{N(N-1)}{2}$ matrix

(56.1)
$$\begin{pmatrix} (\lambda_2-\lambda_1)(S_1)_{12} & (\lambda_2-\lambda_1)(S_2)_{12} & \cdots & (\lambda_2-\lambda_1)(S_\ell)_{12} \\ (\lambda_3-\lambda_1)(S_1)_{13} & (\lambda_3-\lambda_1)(S_2)_{13} & \cdots & (\lambda_3-\lambda_1)(S_\ell)_{13} \\ \vdots & \vdots & & \vdots \\ (\lambda_N-\lambda_{N-1})(S_1)_{N-1,N} & (\lambda_N-\lambda_{N-1})(S_2)_{N-1,N} & \cdots & (\lambda_N-\lambda_{N-1})(S_\ell)_{N-1,N} \end{pmatrix}$$

and so

$$|\det X| = \prod_{1 \le i < j \le N} |\lambda_i - \lambda_j| \quad f_1(0)$$

which implies by (55.2)(56.0)

(56.2)
$$\left|\det\left(\frac{\partial \vec{M}}{\partial(\Lambda, 0)}\right)\right| = \prod_{1 \le i < j \le N} |\lambda_i - \lambda_j| \quad f_1(0)$$

As φ and φ^{-1} are smooth, $f_1(0) > 0$

(why?)