University of California, Los Angeles
Department of Statistics

Statistics 100C **Instructor: Nicolas Christou**

Useful results

Eigenvalues (characteristic values) and eigenvectors (characteristic vectors):
Let $\mathbf{A}$ be a $k \times k$ square matrix and $\mathbf{I}$ be the $k \times k$ identity matrix. Then the scalars $\lambda_1, \lambda_2, \ldots, \lambda_k$ of the solution of $|\mathbf{A} - \lambda\mathbf{I}| = 0$ are called the eigenvalues or characteristic values. The equation $|\mathbf{A} - \lambda\mathbf{I}| = 0$ is a polynomial (function of λ). Here is an example: Let

$$\mathbf{A} = \begin{pmatrix} 1 & 5 \\ 3 & 4 \end{pmatrix}$$

Then

$$|\mathbf{A} - \lambda\mathbf{I}| = \left| \begin{pmatrix} 1 & 5 \\ 3 & 4 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = \left| \begin{pmatrix} 1-\lambda & 5 \\ 3 & 4-\lambda \end{pmatrix} \right| = 0 \Rightarrow (1-\lambda)(4-\lambda) - 15 = 0.$$

And the roots of the function $\lambda^2 - 5\lambda - 11 = 0$ are $\lambda_1 = -1.653312$ and $\lambda_2 = 6.653312$. These are the eigenvalues of the matrix $\mathbf{A}$.

Now let's find the eigenvectors or characteristic vectors of the matrix $\mathbf{A}$:
Let $\mathbf{A}$ be a $k \times k$ matrix and let λ be an eigenvalue of $\mathbf{A}$. If $\mathbf{x}$ is a nonzero $k \times 1$ vector such that $\mathbf{Ax} = \lambda\mathbf{x}$ we say that $\mathbf{x}$ is an eigenvector associated with the eigenvalue λ.

For the example above we have:

$$\begin{pmatrix} 1 & 5 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -1.653312 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 5 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 6.653312 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Here are the results using `R`:

```
> a <- c(1,5,3,4)
> A <- matrix(a, 2, 2, byrow=TRUE)
> A
     [,1] [,2]
[1,]    1    5
[2,]    3    4
> eigen(A)
$values
[1]  6.653312 -1.653312

$vectors
           [,1]       [,2]
[1,] -0.6624990 -0.8833307
[2,] -0.7490628  0.4687504
```

Orthogonal matrix:
A matrix $\mathbf{Q}$ is said to be orthogonal if $\mathbf{QQ}' = \mathbf{Q}'\mathbf{Q} = \mathbf{I}$, or $\mathbf{Q}' = \mathbf{Q}^{-1}$.

Let $\mathbf{A}$ be a symmetric $k \times k$ matrix. The spectral decomposition of $\mathbf{A}$ is expressed as:

$$\mathbf{A} = \mathbf{P\Lambda P}' = \mathbf{P\Lambda P}^{-1}, \text{ (since } \mathbf{P} \text{ is orthogonal).}$$

Note: The columns of the $k \times k$ matrix $\mathbf{P}$ are the normalized eigenvectors which are orthogonal because $\mathbf{A}$ is symmetric. The $k \times k$ matrix $\mathbf{\Lambda}$ is diagonal with diagonal elements the eigenvalues.

Theorem:
If $\mathbf{A}$ is idempotent $k \times k$ matrix then the eigenvalues are 0 or 1.

Proof:
$\mathbf{A}$ is idempotent, therefore, $\mathbf{AA} = \mathbf{A}$. But also from the previous theorem we have $\mathbf{A} = \mathbf{P\Lambda P}'$, therefore,

$$\mathbf{AA} = \mathbf{P\Lambda P}'\mathbf{P\Lambda P}' = \mathbf{P\Lambda\Lambda P}' = \mathbf{P\Lambda}^2\mathbf{P}'$$

It follows that

$$\begin{aligned} \mathbf{A} &= \mathbf{AA} \\ \mathbf{P\Lambda P}' &= \mathbf{P\Lambda}^2\mathbf{P}' \\ \mathbf{\Lambda} &= \mathbf{\Lambda}^2 \end{aligned}$$

But $\mathbf{\Lambda} = diag(\lambda_1, \lambda_1, \ldots, \lambda_k)$, therefore we must have $\lambda_i = \lambda_i^2$, which means $\lambda_i = 0$ or 1.

Theorem:
Let $\mathbf{A}$ be a $k \times k$ idempotent matrix with $tr(\mathbf{A}) = m$. Then there exists $m \times k$ matrix $\mathbf{E}$ such that $\mathbf{E}'\mathbf{E} = \mathbf{A}$ and $\mathbf{EE}' = \mathbf{I}_m$.

Proof:
Since $\mathbf{A}$ is idempotent, $\mathbf{A} = \mathbf{CDC}'$. From the previous theorem the diagonal elements are either equal to 0 or 1. Therefore, we can write the $\mathbf{D}$ matrix as:

$$\begin{pmatrix} \mathbf{I}_q & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \text{ where } q \text{ is the number of ones.}$$

Now, partition $\mathbf{C}$ as follows: $\mathbf{C} = (\mathbf{E}'|\mathbf{F}')$, where $\mathbf{E}$ is a $q \times k$ matrix and $\mathbf{F}$ is $(k-q) \times k$ matrix. We can write $\mathbf{A}$ as:

$$\mathbf{A} = \mathbf{CDC}' = (\mathbf{E}'|\mathbf{F}') \begin{pmatrix} \mathbf{I}_q & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{F} \end{pmatrix} \Rightarrow \mathbf{A} = \mathbf{E}'\mathbf{E}.$$

In addition, since $\mathbf{C}$ is orthogonal we have that $\mathbf{C}'\mathbf{C} = \mathbf{I}_k$ or

$$\begin{aligned} \mathbf{C}'\mathbf{C} = \mathbf{I}_k &= \begin{pmatrix} \mathbf{I}_q & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{k-q} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{E} \\ \mathbf{F} \end{pmatrix} (\mathbf{E}'|\mathbf{F}') = \begin{pmatrix} \mathbf{EE}' & \mathbf{EF}' \\ \mathbf{FE}' & \mathbf{FF}' \end{pmatrix} \Rightarrow \mathbf{EE}' = \mathbf{I_q}. \end{aligned}$$

But $m = tr(\mathbf{A}) = tr(\mathbf{CDC}') = tr(\mathbf{DC}'\mathbf{C})$.

$$tr(\mathbf{DC}'\mathbf{C}) = tr\left[\begin{pmatrix} \mathbf{I}_q & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{EE}' & \mathbf{EF}' \\ \mathbf{FE}' & \mathbf{FF}' \end{pmatrix}\right] = tr\begin{pmatrix} \mathbf{EE}' & \mathbf{EF}' \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = tr(\mathbf{EE}') = m.$$

Therefore, $\mathbf{EE}' = \mathbf{I}_m$.

See numerical example using `R` on the next page to verify this important theorem.

Example:
Consider the following data:

```
> X
     ones
[1,]    1 11.7 85 1022
[2,]    1  8.6 81 1141
[3,]    1  6.5 68  640
[4,]    1  2.6 81  257
[5,]    1  2.8 48  269
[6,]    1  3.0 61  281
```

The idempotent hat matrix $\mathbf{H}$ is given by: $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$:

```
> H <- X %*% solve(t(X) %*% X) %*% t(X)
> H
            [,1]         [,2]       [,3]         [,4]        [,5]        [,6]
[1,]  0.92191692 -0.037684304 0.24729375 -0.015589431 -0.09295479 -0.02298215
[2,] -0.03768430  0.980843706 0.12275678  0.001566411 -0.03053258 -0.03695002
[3,]  0.24729375  0.122756784 0.20481955  0.017404606  0.24400113  0.16372417
[4,] -0.01558943  0.001566411 0.01740461  0.911627909 -0.15295598  0.23794648
[5,] -0.09295479 -0.030532576 0.24400113 -0.152955978  0.67748630  0.35495591
[6,] -0.02298215 -0.036950020 0.16372417  0.237946483  0.35495591  0.30330561
```

Let's find the eigenvalues and eigenvectors of $\mathbf{H}$:

```
r <- eigen(H)
val <- r$values
vec <- r$vectors

> val
[1] 1.000000e+00 1.000000e+00 1.000000e+00 1.000000e+00 1.422563e-15 3.400136e-16

> vec
            [,1]         [,2]        [,3]         [,4]         [,5]        [,6]
[1,]  0.44837342  0.518084668 -0.08826909  0.666839597  0.001865000 -0.27942727
[2,] -0.62890209 -0.369968471  0.08191754  0.664634275  0.032030844 -0.13464887
[3,]  0.16826715 -0.008915007  0.29076656  0.303118870 -0.115386559  0.88423209
[4,] -0.46981644  0.647071341  0.50151216 -0.143821616 -0.291613666 -0.05773699
[5,]  0.38195825 -0.419369053  0.59636471  0.008539588 -0.458046335 -0.33571901
[6,]  0.09224981  0.007772225  0.54202174 -0.030783142  0.831150716 -0.07669994
```

The previous theorem states that:

```
> vec[1:6,1:4] %*% t(vec[1:6,1:4])
            [,1]         [,2]       [,3]         [,4]        [,5]        [,6]
[1,]  0.92191692 -0.037684304 0.24729375 -0.015589431 -0.09295479 -0.02298215
[2,] -0.03768430  0.980843706 0.12275678  0.001566411 -0.03053258 -0.03695002
[3,]  0.24729375  0.122756784 0.20481955  0.017404606  0.24400113  0.16372417
[4,] -0.01558943  0.001566411 0.01740461  0.911627909 -0.15295598  0.23794648
[5,] -0.09295479 -0.030532576 0.24400113 -0.152955978  0.67748630  0.35495591
[6,] -0.02298215 -0.036950020 0.16372417  0.237946483  0.35495591  0.30330561

> t(vec[1:6,1:4]) %*% vec[1:6,1:4]
              [,1]          [,2]          [,3]          [,4]
[1,]  1.000000e+00 -3.122502e-17 -5.551115e-17  2.346214e-16
[2,] -3.122502e-17  1.000000e+00  1.092876e-16 -2.375758e-16
[3,] -5.551115e-17  1.092876e-16  1.000000e+00 -7.632783e-17
[4,]  2.346214e-16 -2.375758e-16 -7.632783e-17  1.000000e+00
```

Confirm that

$$\mathbf{H} = \mathbf{P\Lambda P'} = \mathbf{P\Lambda P^{-1}}$$

```
> diag(val)
     [,1] [,2] [,3] [,4]         [,5]         [,6]
[1,]    1    0    0    0 0.000000e+00 0.000000e+00
[2,]    0    1    0    0 0.000000e+00 0.000000e+00
[3,]    0    0    1    0 0.000000e+00 0.000000e+00
[4,]    0    0    0    1 0.000000e+00 0.000000e+00
[5,]    0    0    0    0 1.422563e-15 0.000000e+00
[6,]    0    0    0    0 0.000000e+00 3.400136e-16

> vec %*% diag(val) %*% t(vec)
            [,1]         [,2]       [,3]         [,4]        [,5]        [,6]
[1,]  0.92191692 -0.037684304 0.24729375 -0.015589431 -0.09295479 -0.02298215
[2,] -0.03768430  0.980843706 0.12275678  0.001566411 -0.03053258 -0.03695002
[3,]  0.24729375  0.122756784 0.20481955  0.017404606  0.24400113  0.16372417
[4,] -0.01558943  0.001566411 0.01740461  0.911627909 -0.15295598  0.23794648
[5,] -0.09295479 -0.030532576 0.24400113 -0.152955978  0.67748630  0.35495591
[6,] -0.02298215 -0.036950020 0.16372417  0.237946483  0.35495591  0.30330561

> vec %*% diag(val) %*% solve(vec)
            [,1]         [,2]       [,3]         [,4]        [,5]        [,6]
[1,]  0.92191692 -0.037684304 0.24729375 -0.015589431 -0.09295479 -0.02298215
[2,] -0.03768430  0.980843706 0.12275678  0.001566411 -0.03053258 -0.03695002
[3,]  0.24729375  0.122756784 0.20481955  0.017404606  0.24400113  0.16372417
[4,] -0.01558943  0.001566411 0.01740461  0.911627909 -0.15295598  0.23794648
[5,] -0.09295479 -0.030532576 0.24400113 -0.152955978  0.67748630  0.35495591
[6,] -0.02298215 -0.036950020 0.16372417  0.237946483  0.35495591  0.30330561
```