Recitation 10. May 18

Focus: statistics, Fourier series

Consider running a measurement n times, and getting the samples $x_1, x_2, \ldots, x_n$. The collection of these n numbers is known as a data seat. The **mean** of the data set is:

$$\mu = \frac{1}{n}\Big(x_1 + \cdots + x_n\Big)$$

Given two data sets $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ with means μ and ν, their **covariance** is:

$$\Sigma_{xy} = \frac{1}{n-1}\Big((x_1-\mu)(y_1-\nu) + \cdots + (x_n-\mu)(y_n-\nu)\Big)$$

(you get $n-1$ instead of n in the denominator due to Bessel's correction).

The covariance of the data set $x_1, \ldots, x_n$ with itself is called its **variance** $\Sigma = \frac{1}{n-1}((x_1-\mu)^2 + \cdots + (x_n-\mu)^2)$.

In terms of the vectors $\boldsymbol{o} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$, $\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $\boldsymbol{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$, the (co)variance is given by:

$$\boxed{\Sigma_{xy} = \frac{\boldsymbol{x}^T P \boldsymbol{y}}{n-1}} \qquad \text{where } P = I - \frac{\boldsymbol{o}\boldsymbol{o}^T}{\boldsymbol{o}^T\boldsymbol{o}} \text{ is the projection matrix onto the orthogonal complement of } \boldsymbol{o}$$

In general, let $\boldsymbol{A} = \begin{bmatrix} x_1 & y_1 & z_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & z_n & \cdots \end{bmatrix}$ a matrix of different data sets. Their **covariance matrix** is computed by:

$$\boxed{K = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} & \Sigma_{xz} & \cdots \\ \Sigma_{yx} & \Sigma_{yy} & \Sigma_{yz} & \cdots \\ \Sigma_{zx} & \Sigma_{zy} & \Sigma_{zz} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \frac{\boldsymbol{A}^T P \boldsymbol{A}}{n-1}}$$

Any 2π-periodic function $f(x)$ can be written as a **Fourier series**:

$$f(x) = a_0 + a_1 \cos x + a_2 \cos 2x + a_3 \cos 3x + \cdots + b_1 \sin x + b_2 \sin 2x + b_3 \sin 3x + \cdots$$

where:

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)dx$$
$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)dx$$
$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)dx$$

for all $n > 0$. Alternatively, one can define complex-valued Fourier series, and write any 2π-periodic function as:

$$f(x) = \sum_{k\in\mathbb{Z}} c_k e^{ikx}$$

where:

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}dx$$

1. Consider the matrix:

$$\begin{bmatrix} 1 & 2 \\ a & b \end{bmatrix}$$

For some constants a and b. Suppose it is a covariance matrix of two random variables X and Y.

- What can you say about a and b based on the information above?

- What can you say about a and b if, on top of the information above, you know that there is some linear combination of X and Y that are constant?

Solution: Any covariance matrix is a symmetric positive semidefinite matrix. By symmetry, we conclude that $a = 2$. By positive definiteness, we conclude that $\text{Tr } = 1 + b$ and $\det = b - 4$ should both be non-negative, so this implies $b \geq 4$. If we also know that a certain linear combination of X and Y is constant, then the energy of the corresponding vector (of coefficients in that linear combination) is 0. This only happens if the matrix is not positive definite, hence $\det = 0$, hence $b = 4$ (FYI: the linear combination which is constant would be $2X - Y$, since $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ spans the null-space of the covariance matrix).

2. Consider the following measurements for temperature and pressure (don't worry about units):

$$T = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} \quad \text{and} \quad P = \begin{bmatrix} 6 \\ 1 \\ 2 \end{bmatrix}$$

- Compute the covariance matrix of T and P.
- Find linear combinations of temperature and pressure that are uncorrelated.

Solution: First we put the above samples into a matrix:

$$A = \begin{bmatrix} 1 & 6 \\ 2 & 1 \\ -3 & 2 \end{bmatrix}$$

Also consider the matrix:

$$P = \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

Then the covariance matrix is given by:

$$K = \frac{A^T P A}{3 - 1} = \begin{bmatrix} 7 & 1 \\ 1 & 7 \end{bmatrix}$$

To find independent random variables we need to diagonalize K. Note that the eigenvalues are given by $\lambda_1 = 6$ and $\lambda_2 = 8$ with eigenvectors:

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Thus we conclude that $T - P$ and $T + P$ are independent random variables.

3. Consider the 2π-periodic square wave, which on the interval $[-\pi, \pi]$ is described by the function:

$$f(x) = \begin{cases} 0, & \text{if } -\pi \leq x \leq 0 \\ 1, & \text{if } 0 < x \leq \pi \end{cases}$$

Compute the Fourier series expansion of $f(x)$, in terms of either sines/cosines or complex exponentials.

Solution: The various Fourier coefficients are calculated as in the formulas on the first page:

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)dx = \frac{1}{2\pi}\int_0^{\pi} 1dx = 1/2$$
$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)dx = \frac{1}{\pi}\int_0^{\pi} \cos(nx)dx = 0$$
$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)dx = \frac{1}{\pi}\int_0^{\pi} sin(nx)dx = \frac{-cos(n\pi)+\cos(0)}{\pi n} = \frac{1-(-1)^n}{\pi n} = \begin{cases} \frac{2}{\pi n} & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even} \end{cases}$$

for all $n > 0$. We conclude that:

$$f(x) = \frac{1}{2} + \sum_{n=1}^{\infty} \frac{1-(-1)^n}{\pi n}\sin(nx) = \frac{1}{2} + \frac{2\sin(x)}{\pi} + \frac{2\sin(3x)}{3\pi} + \frac{2\sin(5x)}{5\pi} + \cdots$$

(by the way, we could have predicted that the a_n for $n > 0$ are 0, since $f(x) - \frac{1}{2}$ is an odd function like $\sin(nx)$ and unlike $\cos(nx)$).

For the complex Fourier series, you could either convert in the formula above all sines and cosines to complex exponentials via:

$$\cos kx = \frac{e^{ikx}+e^{-ikx}}{2} \qquad \text{and} \qquad \sin kx = \frac{e^{ikx}-e^{-ikx}}{2i}$$

or you could compute the complex Fourier coefficients using the formula on the first page. Let us go for the latter route:

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}dx = \frac{1}{2\pi}\int_0^{\pi} e^{-ikx}dx = \begin{cases} \frac{1}{2} & \text{if } k = 0 \\ \frac{-1}{2\pi ik}e^{-ikx}\Big|_0^{\pi} = \frac{1-(-1)^k}{2\pi ik} & \text{if } k \neq 0 \end{cases}$$

Therefore, we conclude that:

$$f(x) = \frac{1}{2} + \sum_{k \text{ odd integer}} \frac{e^{ikx}}{\pi ik} = \frac{1}{2} + \sum_{k \text{ odd positive integer}} \frac{e^{ikx}-e^{-ikx}}{\pi ik}$$