Lecture 37

Cyclical Behavior and Periodicity: Part 1

Arnab Hazra

Johnson & Johnson Quarterly Earnings (recap)

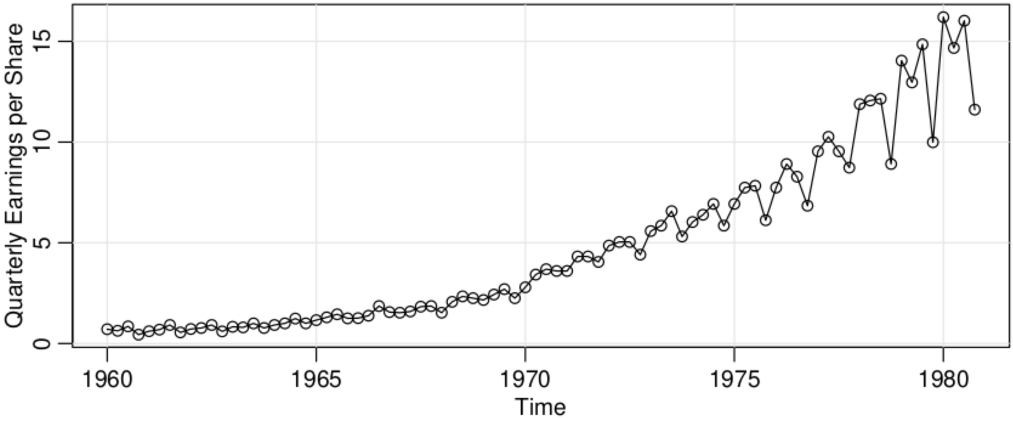

Fig. 1.1. Johnson & Johnson quarterly earnings per share, 84 quarters, 1960-I to 1980-IV.

Johnson & Johnson Quarterly Earnings (recap)

- There are 84 quarters (21 years) measured from the first quarter of 1960 to the last quarter of 1980.
- Note the gradually increasing underlying trend and the rather regular variation superimposed on the trend that seems to repeat over quarters.

Speech Data (recap)

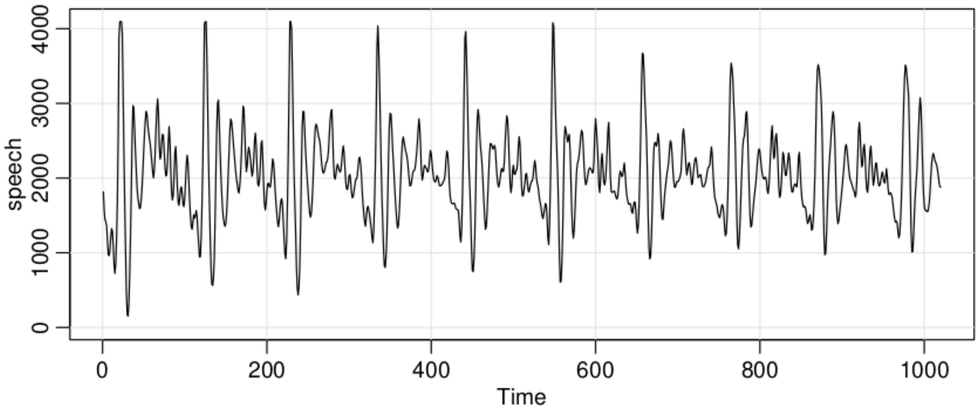

Fig. 1.3. *Speech recording of the syllable aaa $\cdots$ hhh sampled at 10,000 points per second with $n = 1020$ points.*

Speech Data (recap)

- We note the repetitive nature of the signal and the rather regular periodicities.
- Computer recognition of speech require converting this particular signal into the recorded phrase aaa ... hhh.
- Spectral analysis can be used to produce a signature of this phrase that can be compared with signatures of various library syllables to look for a match.
- One can immediately notice the rather regular repetition of small wavelets. The separation between the packets is known as the pitch period.
- Pitch period represents the response of the vocal tract filter to a periodic sequence of pulses stimulated by the opening and closing of the glottis.

fMRI Imaging (recap)

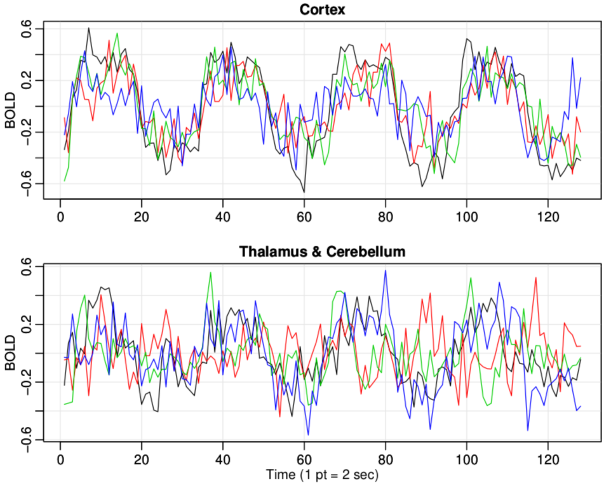

Fig. 1.6. *fMRI data from various locations in the cortex, thalamus, and cerebellum;* $n = 128$ *points, one observation taken every 2 seconds.*

fMRI Imaging (recap)

- Five subjects were given periodic brushing on the hand. The stimulus was applied for 32 seconds and then stopped for 32 seconds; thus, the signal period is 64 seconds.
- The sampling rate was one observation every 2 seconds for 256 seconds ($n = 128$). For this example, results are averaged over subjects.
- The series are consecutive measures of blood oxygenation-level dependent (bold) signal intensity, which measures areas of activation in the brain.
- Notice that the periodicities appear strongly in the motor cortex series and less strongly in the thalamus and cerebellum.
- The fact that one has series from different areas of the brain suggests testing whether the areas are responding differently to the brush stimulus.

Earthquakes and Explosions (recap)

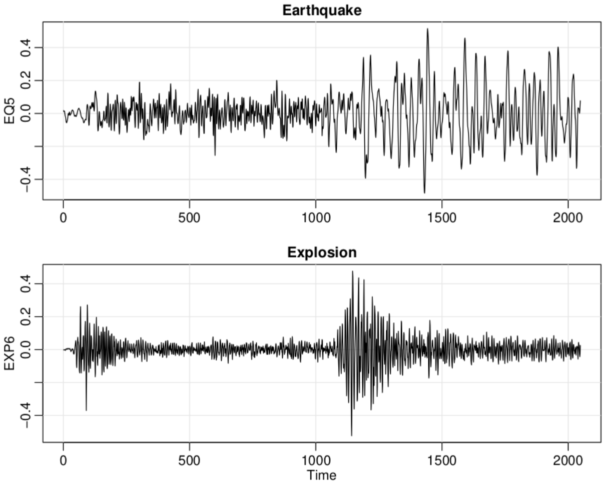

Fig. 1.7. *Arrival phases from an earthquake (top) and explosion (bottom) at 40 points per second.*

Earthquakes and Explosions (recap)

- The series represent two phases along the surface, denoted by P $(t = 1, \ldots, 1024)$ and S $(t = 1025, \ldots, 2048)$, at a seismic recorder.
- The general problem of interest is in distinguishing or discriminating between waveforms generated by earthquakes and those generated by explosions.
- Features that may be important are the rough amplitude ratios of the first phase P to the second phase S.
- The ratio of maximum amplitudes appears to be somewhat less than 0.5 for the earthquake and about 1 for the explosion.
- A subtle difference exists in the periodic nature of the S phase for the earthquake.
- We can again think about spectral analysis of variance for testing the equality of the periodic components of earthquakes and explosions.

SOI and Fish Population (recap)

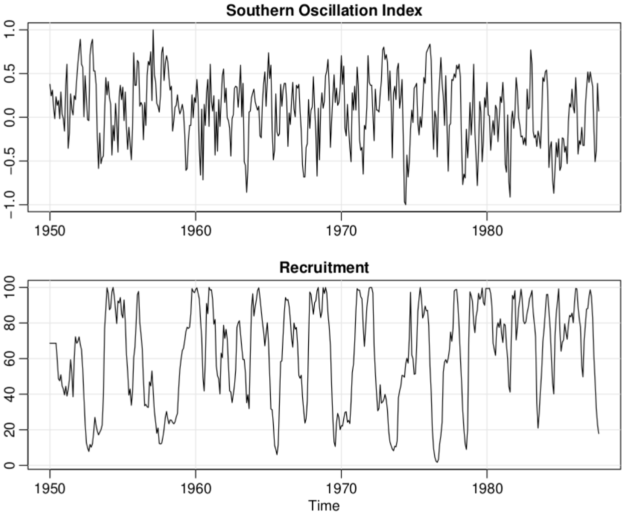

Fig. 1.5. *Monthly SOI and Recruitment (estimated new fish), 1950-1987.*

SOI and Fish Population (recap)

- Both series are for a period of 453 months ranging over the years 1950–1987.
- The SOI measures El Nino, i.e., changes in air pressure, related to sea surface temperatures in the central Pacific Ocean.
- Both series exhibit repetitive behavior, with regularly repeating cycles that are easily visible.
- This periodic behavior is of interest because underlying processes of interest may be regular and the rate or frequency of oscillation characterizing the behavior of the underlying series would help to identify them.
- The series show two basic oscillations types, an obvious annual cycle, and a slower frequency that seems to repeat about every 4 years.
- The two series are also related.

Frequency and period of a time series

- We measure frequency ω by cycles per time point.
- In the Johnson & Johnson data set, the predominant frequency of oscillation is one cycle per year (4 quarters), or $\omega = 0.25$ cycles per observation.
- The predominant frequency in the SOI and fish populations series is 1 cycle every 12 months, or $\omega = 0.083$ cycles per observation.
- For a discrete time series, we will need at least two points to determine a cycle, so the highest frequency, called the folding frequency, is 0.5.
- The period of a time series is the number of points in a cycle, i.e., $1/\omega$.
- The predominant period of the Johnson & Johnson series is $1/0.25$ or 4 quarters per cycle.
- The predominant period of the SOI series is 12 months per cycle.

Periodic process

- We first define a cycle as one complete period of a sine or cosine function defined over a unit time interval.
- We consider the periodic process

$$X_t = A\cos(2\pi\omega t + \phi)$$

  for $t = 0, \pm 1, \pm 2, \ldots$, where ω is a frequency index.
- Here A determines the height or amplitude of the function and ϕ, called the phase, determining the start point of the cosine function.
- We can introduce random variation in this time series by allowing the amplitude and phase to vary randomly.

Periodic process (contd.)

- For purposes of data analysis, it is easier to write X_t as

$$X_t = U_1 \cos(2\pi\omega t) + U_2 \sin(2\pi\omega t),$$

  where $U_1 = A\cos(\phi)$ and $U_2 = -A\sin(\phi)$.
- We then often take U_1 and U_2 to be normally distributed.
- The amplitude is $A = \sqrt{U_1^2 + U_2^2}$ and the phase is $\phi = \tan^{-1}(-U_2/U_1)$.
- Here, A and ϕ are independent random variables if U_1 and U_2 are independent standard normal random variables.
- Then $A^2 \sim \chi_2^2$ and $\phi \sim \mathrm{Unif}(-\pi, \pi)$.
- Straightforward Jacobian calculations show that the reverse is also true.

Moments of X_t

If we assume that U_1 and U_2 are uncorrelated random variables with mean 0 and variance σ^2, then

- Mean $E(X_t) = 0$.
- Covariance $\mathrm{Cov}(X_{t+h}, X_t) = \sigma^2 \cos(2\pi\omega h)$.
- Variance $\mathrm{Var}(X_t) = \sigma^2$
- If we observe U_1 and U_2, an estimate of σ^2 is the sample variance

$$S^2 = \frac{U_1^2 + U_2^2}{2 - 1} = U_1^2 + U_2^2.$$

Further generalization of periodic processes

- We can allow mixtures of periodic series with multiple frequencies and amplitudes,

$$X_t = \sum_{k=1}^{K} [U_{k1} \cos(2\pi\omega_k t) + U_{k2} \sin(2\pi\omega_k t)]$$

  where U_{k1}, U_{k2}, for $k = 1, 2, \ldots, K$, are uncorrelated zero-mean random variables with variances σ_k^2, and the ω_k are distinct frequencies.
- Notice that the process is a sum of uncorrelated components, with variance σ_k^2 for frequency ω_k.
- The autocovariance function of X_t is

$$\mathrm{Cov}(X_{t+h}, X_t) = \sum_{k=1}^{K} \sigma_k^2 \cos(2\pi\omega_k h), \quad \mathrm{Var}(X_t) = \sum_{k=1}^{K} \sigma_k^2.$$

- Here $\widehat{\mathrm{Var}}(X_t) = \sum_{k=1}^{K} \hat{\sigma}_k^2 = \sum_{k=1}^{K} [U_{k1}^2 + U_{k2}^2]$.

Sum of periodic functions

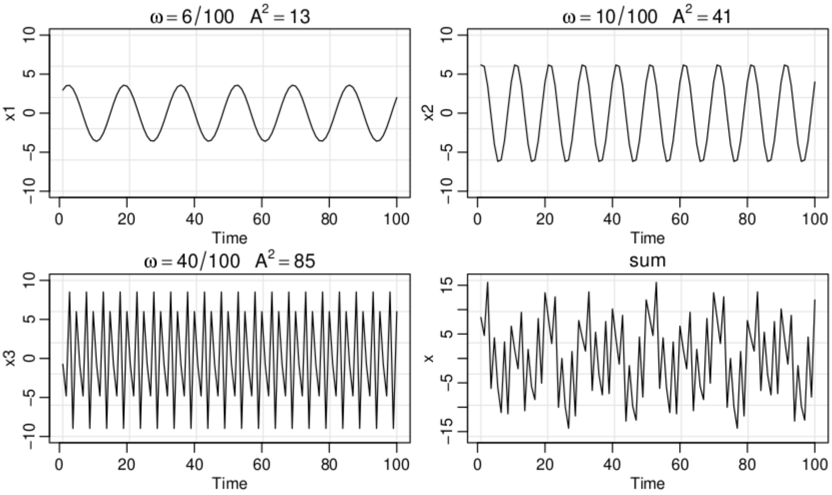

Fig. 4.1. *Periodic components and their sum as described in Example 4.1.*

Our goal

- The systematic sorting out of the essential frequency components, including their relative contributions, is one of the main objectives of *spectral analysis*.
- The moments of X_t only discuss the population properties but no statistical inference.
- If we can observe U_{k1} and U_{k2}, then $\hat{\sigma}_k^2 = U_{k1}^2 + U_{k2}^2$.
- In practice, we only observe $X_1, \ldots, X_T$ but not U_{k1}'s and U_{k2}'s.
- Hence, we next discuss the practical aspects of how, given data $X_1, \ldots, X_T$, to actually estimate the variance components $\sigma_k^2, k = 1, \ldots, K$.

Thank you!