Lecture 4: Time series. Autocovariance

Math 586

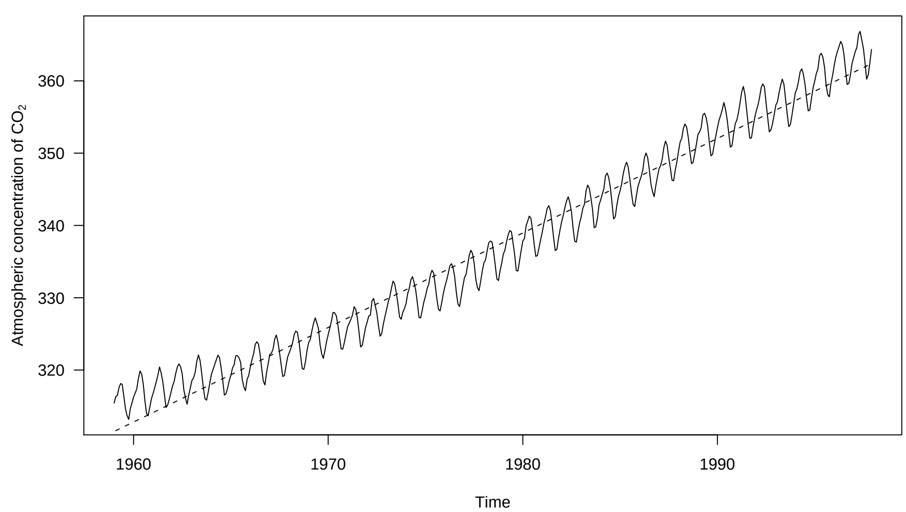
February 2, 2009

Time series: one-dimensional array, regular spacing: observed at times $t_0, t_0 + \Delta t, t_0 + 2\Delta t, ...$
May also apply to geo. data, e.g. measurements along a given transect, depth etc.

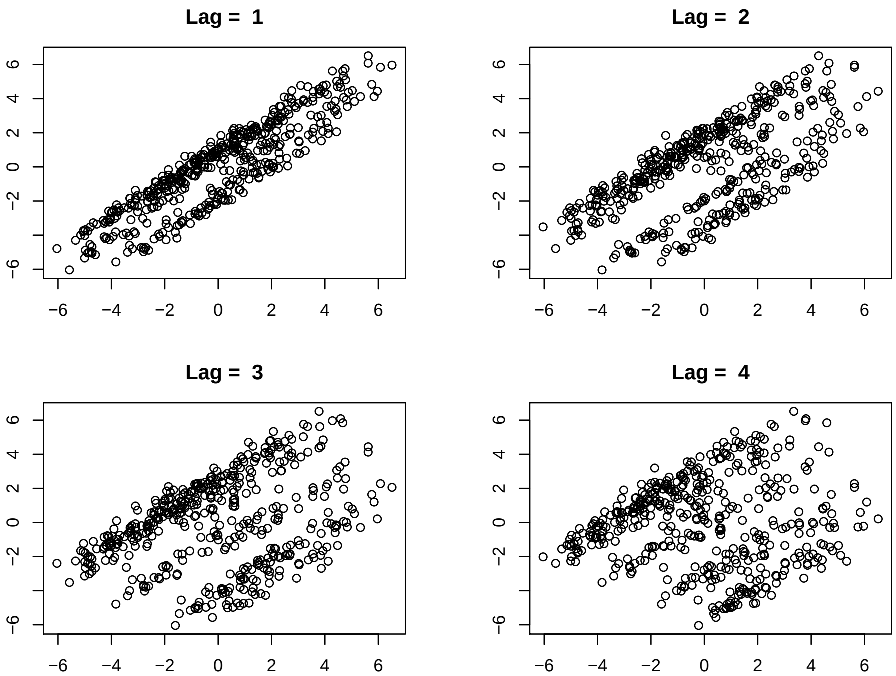
Need a way to describe dependence between measurements. For simplicity, let $\Delta = 1$ and consider relation between X_t and X_{t+k}, $k =$ **lag**.

If we plot the vector $X_1, ..., X_{n-k}$ against the vector $X_{k+1}, ..., X_n$, we might notice some structure.

Example: Mauna Loa Atmospheric CO_2 Concentration (ppm), a time series of 468 observations; monthly from 1959 to 1997.
Source: R built-in data set, primarily at
`ftp://cdiac.esd.ornl.gov/pub/maunaloa-co2/maunaloa.co2`



Features: increasing trend, yearly oscillations, some long-term “mini-trends”.
Lagged plots (after subtracting the trend):



> **Definition**
> Time series X_t is **strictly stationary** if the joint distribution of $X_{k+1}, ..., X_{k+n}$, for each n, does not depend on lag k.
>
> - is **second-order stationary** if its mean, variance and autocovariances $C(k) = Cov(X_i, X_{i+k})$ do not depend on i.

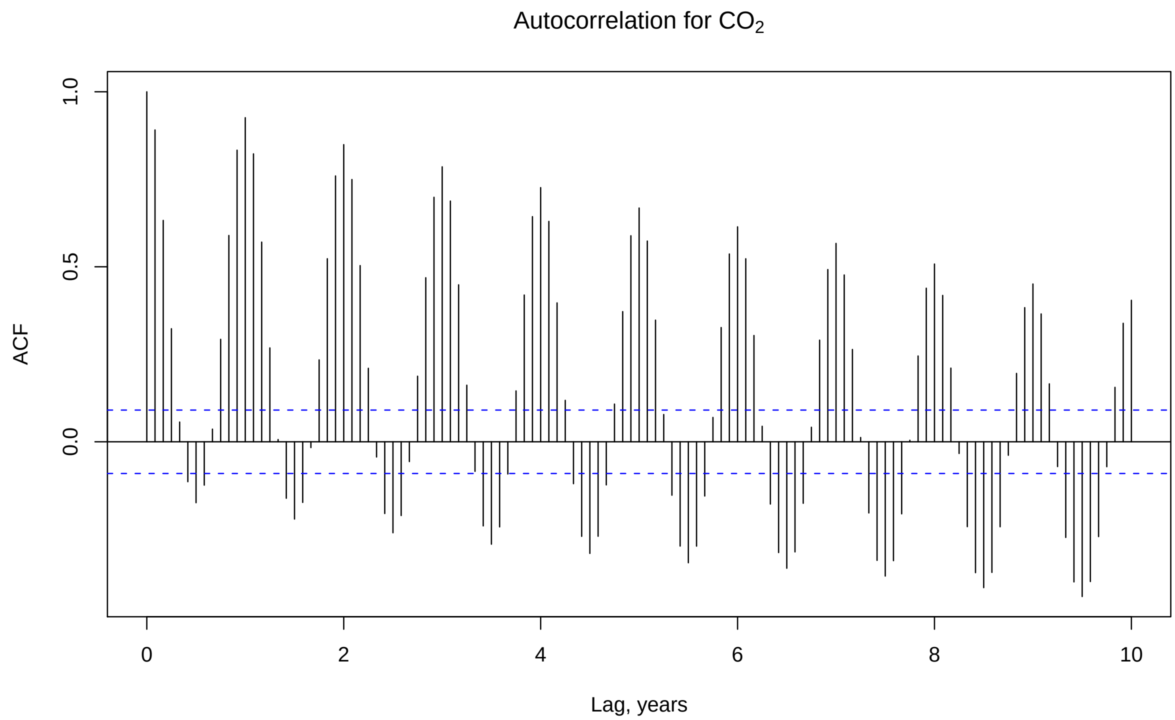
(implies, in particular, that there is no trend).

> **Definition**
> **Sample autocovariance function** $\hat{C}(k)$ for lag k is computed by
>
> $$\hat{C}(k) = \frac{1}{n}\sum_{i=1}^{n-k}(X_i - \overline{x})(X_{i+k} - \overline{x})$$

Note:

1. divided by n and not by $n - k$ for admissibility.
2. Sample autocorrelation function (ACF) is obtained as $\hat{C}(k)/\hat{C}(0)$, analogous to sample correlation.
3. $\hat{C}(0)$ is just an estimate of variance of X.

ACF of $residuals = observed - trend$ is shown:



Properties of Autocovariance function:

(i.) $C(k) = C(-k)$ symmetry (also for $\hat{C}$).

(ii.) $C(k) \to 0$ as $k \to \infty$ in most cases (if there is no trend)

(iii.) $|C(k)| \leq C(0)$ (because correlation can't exceed 1)

(iv.) Function C is **positive-semidefinite**, i.e. for every N and every $b_1, ..., b_N$, $t_1, ..., t_N$

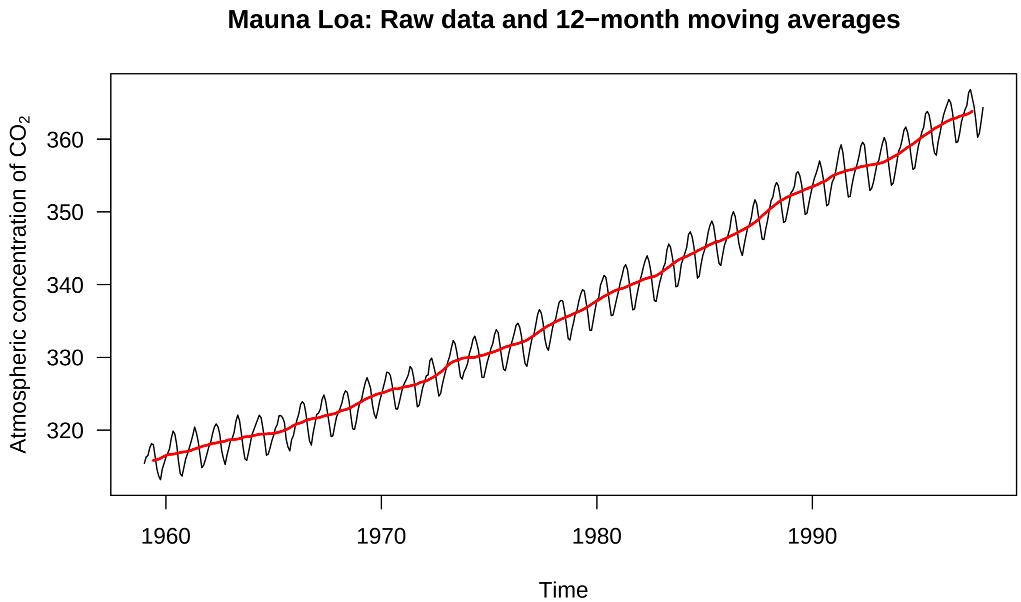
$$Var\left(\sum_{i=1}^{N} b_i X(t_i)\right) = \sum_{i=1}^{N}\sum_{j=1}^{N} b_i b_j C(t_i - t_j) \geq 0$$

Moving Averages

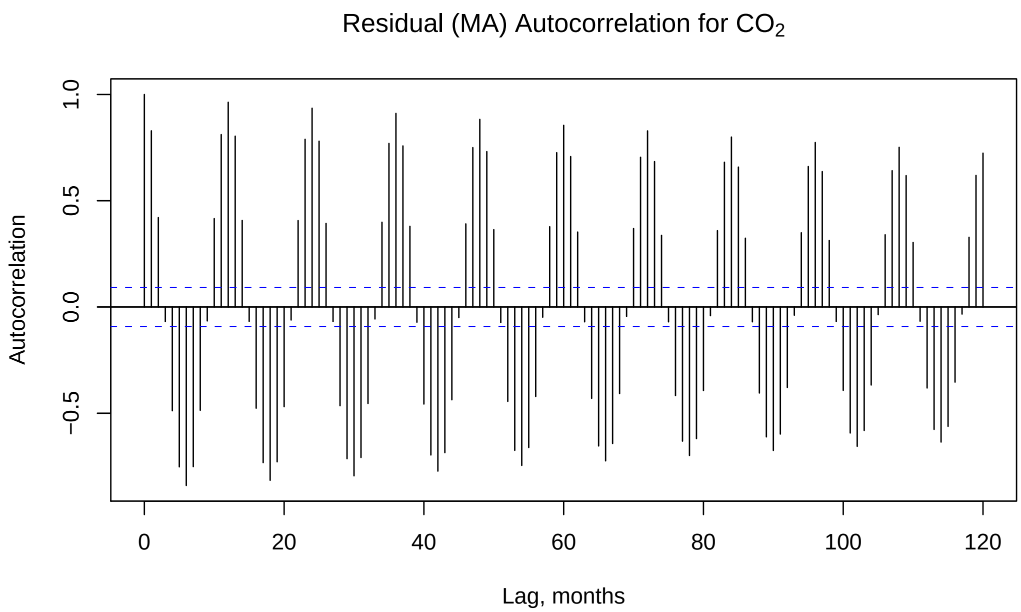
How can simple models for autocovariance arise? Related to the "moving window" statistics –
Moving average (MA): for a window size $2M + 1$, let

$$m_t = \sum_{i=-M}^{M} \omega_i X(t+i), \qquad M \leq t \leq n - M$$

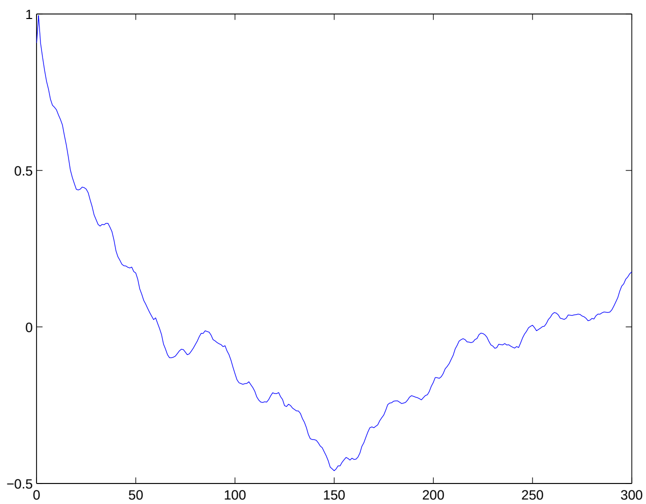
where ω_i are weights such that $\omega_{-M} + \omega_{-M+1} + ... + \omega_0 + ... + \omega_M = 1$. (The weights are usually required to add up to 1 to preserve the mean.)
For time series data, MA may work as a *filter*, that is, it can filter out periodicity and random "noise" to produce a *trend* estimate:



This leads to an improved trend estimate, for example, now the residual autocorrelation function has less artifacts:



Next, in standard time series analysis, we would extract the periodicity etc. The detrended "signal" itself possesses some long-range dependency (we can see that our time span is not enough to fully assess it):



MA models

On the other hand, MA can work as a simple model introducing spatial dependency.
Let $X_1, X_2, ..., X_n$ be independent, Normal (μ, σ^2) r.v.'s. Let Y_t be moving average with the window size, say, 3 (with $M = 1$) and $\omega_i \equiv 1/3$. That is,

$$Y_t = \frac{X_{t-1} + X_t + X_{t+1}}{3}$$

Since X_t are independent, their autocovariance function is a scaled delta function: $Cov(X_t, X_s) = \sigma^2$ when $s = t$ and 0 otherwise.
However, for Y's,

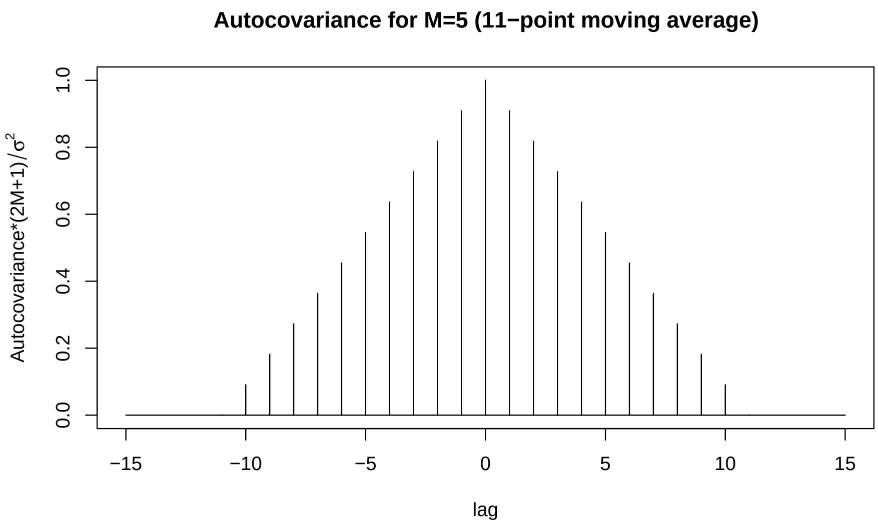
$$C(0) \equiv Var(Y_t) = Cov(X_{t-1}+X_t+X_{t+1}, X_{t-1}+X_t+X_{t+1})/9 = \sigma^2/3, \quad \text{and}$$

$$C(1) \equiv Cov(Y_t, Y_{t+1}) = Cov(X_{t-1} + X_t + X_{t+1}, X_t + X_{t+1} + X_{t+2})/9 = 2\sigma^2/9$$

(use the distributive law, Luke)

Note that Y's are stationary.
For which k does $C(k) = 0$?

You can obtain similar expressions for $C(k)$ in the case of general ω_i and different M (see Problem Set).



Extension to Spatial

In the same vein, we can define moving average for spatial, e.g. 2-d processes. Let V be observed over a discrete grid (for simplicity, integer x, y). Let

$$m(x,y) = \sum_{i,j=-M}^{M} \omega_{ij} V(x+i, y+j)$$

We can also define **covariance function** at a **vector lag h**

$$C(\mathbf{h}) = Cov(V(\mathbf{x}), V(\mathbf{x}+\mathbf{h}))$$

Second-order stationary, or **statistically homogeneous** (some books: WSS, or wide-sense stationary, or weakly stationary) if $C(\cdot)$ doesn't depend on $\mathbf{x}$.

Example: Let $i, j = -1, 0, 1$ (window size 3). Compute $C(\mathbf{h})$ for different $\mathbf{h}$. When does $C(\mathbf{h}) = 0$?