Statistical properties of time series (continued)

Solution of homework: Is $\sigma_{est}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i-\bar{X})^2$ an unbiased estimator for σ^2 in autocorrelated case?

$$\left\langle \sum_{i=1}^{N}(X_i-\bar{x})^2\right\rangle = \frac{1}{N^2}\sum_i\left\langle\left(NX_i-\sum_{j=1}^{N}X_j\right)^2\right\rangle$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\left\langle\left(N\Delta X_i-\sum_{j=1}^{N}\Delta X_j\right)^2\right\rangle \quad (\Delta X_i \equiv X_i - \langle x\rangle)$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\left\langle N^2\Delta X_i^2 + \left(\sum_{j=1}^{N}\Delta X_j\right)^2 - 2N\Delta X_i\sum_{j=1}^{N}\Delta X_j\right\rangle$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\left\langle N^2\Delta X_i^2 + \sum_{j=1}^{N}\sum_{k=1}^{N}\Delta X_j\Delta X_k - 2N\sum_{j=1}^{N}\Delta X_i\Delta X_j\right\rangle$$

$$= N\sigma^2 - \underbrace{\frac{1}{N}\sum_{i,j=1}^{N}\langle\Delta X_i\,\Delta X_j\rangle}_{\approx\tau\sigma^2 \text{ (see below)}} \approx (N-\tau)\,\sigma^2$$

$\tau \geq 1$

$\Rightarrow$ above estimator systematically underestimates σ^2 for autocorrelated data ($\frac{1}{N}$-effect). Better: $\sigma_{est,imp}^2 = \frac{\sum_i (X_i-\bar{X})^2}{N-\tau}$

Notes:
- no exact unbiased estimator (but bias can be minimized using good approximation for τ)
- σ_{est}^2 may be „good enough"
- In general $\langle(X_i-\bar{x})^2\rangle \neq \langle(X_j-\bar{x})^2\rangle$ for $i\neq j$ and autocorrelated data (less deviations for inner values)

Step (ii) variance of mean value $\bar{x}$:

$$\langle (\bar{x} - \langle x\rangle)^2\rangle = \langle (\tfrac{1}{N}\sum_{i=1}^{N} x_i - \langle x\rangle)^2\rangle$$
$$= \langle [\tfrac{1}{N}\sum_{i=1}^{N} (x_i - \langle x\rangle)]^2\rangle$$
$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\langle (x_i-\langle x\rangle)(x_j-\langle x\rangle)\rangle$$

$$\left(\text{gen. Def.: } c_{A,B} = \frac{\langle (A-\langle A\rangle)(B-\langle B\rangle)\rangle}{\sigma_A \quad \sigma_B}\right)$$

covariance

$$= \frac{\sigma^2}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} c_{i,j}$$

$$\overset{\text{translation inv.}}{=} \frac{\sigma^2}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} c_{i-j}$$

$k = i - j$

$j = i - k$

$$= \frac{\sigma^2}{N^2}\sum_{i=1}^{N}\sum_{k=i-N}^{i-1} c_k$$

$$\overset{c_k\approx 0 \text{ for } |k|>K\ll N}{\approx} \frac{\sigma^2}{N^2}\sum_{i=1}^{N}\sum_{k=-\infty}^{\infty} c_k = \frac{\sigma^2}{N}\tau$$

autocorrelation time $\displaystyle \tau = \sum_{k=-\infty}^{\infty} c_k \overset{c_k=c_{-k}}{=} 1 + 2\sum_{k=1}^{\infty} c_k$

$$c_k \approx \frac{\frac{1}{N-k-1}\sum_{i=1}^{N-k}(x_i-\bar{x})(x_{i+k}-\bar{x})}{\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{x})^2}$$

autocorrelation function

practical computation of autocorrelation time:

$$\tau \approx 1 + 2\sum_{k=1}^{k_c} c_k$$

k_c determined by: $c_k > 0 \ \forall k \le k_c$; $c_{k_c+1} \le 0$

not discussed in lecture

Estimate for $S_{ij} = \langle (x_i - \langle x \rangle)(x_j - \langle x \rangle) \rangle$? $\qquad (i \neq j)$

How good is $\tilde{S}_{ij} = \langle (x_i - \bar{x})(x_j - \bar{x}) \rangle$?

$$\tilde{S}_{ij} = \langle (x_i - \frac{1}{N}\sum_{k=1}^{N} x_k)(x_j - \frac{1}{N}\sum_{m=1}^{N} x_m) \rangle$$

$$= \langle [(x_i - \langle x \rangle) - \frac{1}{N}\sum_{k=1}^{N}(x_k - \langle x \rangle)][(x_j - \langle x \rangle) - \frac{1}{N}\sum_{m=1}^{N}(x_m - \langle x \rangle)] \rangle$$

$$= S_{ij} - \frac{1}{N}\sum_{m=1}^{N} S_{im} - \frac{1}{N}\sum_{k=1}^{N} S_{kj} + \frac{1}{N^2}\sum_{k,m=1}^{N} S_{km}$$

$$= (1 - \frac{1}{N})^2 S_{ij} - \frac{1}{N}(S_{ii} + S_{jj}) + \frac{1}{N^2}\sum_{k} S_{kk} + \text{offdiag terms}$$

$$= (1 - \frac{1}{N})^2 S_{ij} - \frac{\sigma^2}{N} + \text{offdiag terms}$$

$$\xrightarrow[\text{data}]{\text{uncorr}} -\frac{\sigma^2}{N}$$

$$\Rightarrow \boxed{S_{ij} \approx \frac{\tilde{S}_{ij} + \frac{\sigma^2}{N}}{(1 - \frac{1}{N})^2}}$$

(practical data analysis using examples on course page)