Practice questions

1) Given two independent variables x, y with a Gaussian PDF, define a new set of variables

$$\begin{cases} \hat{x} = a_{11} x + a_{12} y \\ \hat{y} = a_{21} x + a_{22} y \end{cases}$$

(a) What is the correlation between $\hat{x}$ and $\hat{y}$?

(b) What is the joint PDF $P_{\hat{x}\hat{y}}$?

(c) Is the estimate of the covariance $\langle \hat{x} \hat{y} \rangle$ consistent with the $P_{\hat{x}\hat{y}}$? (Show it)

2) Given a model $y = \alpha x + n$

(a) Compute α using the statistics of x, y

(b) Show that the estimate of α is equivalent to minimizing the error variance using LSQ

(c) Show that the square of the correlation coefficient $\rho = \dfrac{\langle x y \rangle}{(\langle x^2 \rangle \langle y^2 \rangle)^{1/2}}$ is the fraction of variance of $\langle y^2 \rangle$ that cannot be explained by x

3) Given the two independent variable x, y define a set

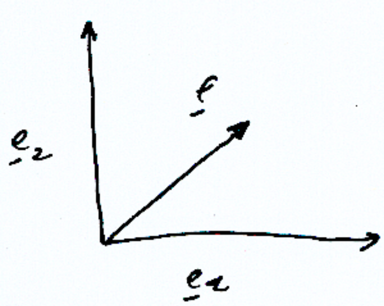
$$\begin{cases} \hat{x} = x\cos\phi + y\sin\phi \\ \hat{y} = -x\,\cancel{\sin(\phi)}\sin\phi + y\cos\phi \end{cases}$$

(a) Find the angle ϕ for which $\hat{x}$ and $\hat{y}$ are uncorrelated. ~~other~~

(b) What is the analytical formula for ϕ?

(c) Can you apply this rotation to any two time series to render them uncorrelated?

4) Consider the Cartesian basis in 2D



$$\underline{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \underline{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Any vector $\underline{f}$ can be expressed acording to the projection on the two axis defined by $\underline{e}_1$ $\underline{e}_2$

$$\underline{f} = \underline{e}_1\alpha_1 + \underline{e}_2\alpha_2$$ or more generally $$\underline{f} = \sum_{i=1}^{N} \alpha_i \underline{e}_i$$

The set $\underline{e}_i$ is said to be orthogonal if $\underline{e}_i$ has the property that no $\underline{e}_i$ can be represented using another $\underline{e}_j$

e.g. $e_1 * \alpha_1 \neq e_2$ because $\underline{e}_1$ and $\underline{e}_2$ is orthogonal 90°

more generally

$$\underline{e}_i^T \underline{e}_j = \delta_{ij}$$

(a) Show that $\underline{e}_1$ and $\underline{e}_2$ are indeed orthogonal by direct multiplication.

(b) How can you use the orthogonality to estimate α_1 projection for the vector $\underline{f}$?

5) Assume that $f(t)$ is a continous function.

(a) What are the projections of $f(t)$ on the basis

$$e_1(t) = a_1 \cos(2\pi s_1 t)$$

$$e_2(t) = b_2 \cos(2\pi s_2 t)$$

(b) Are e_1 and e_2 orthogonal (NOTE: $\underline{e}_1^T \underline{e}_2 = \int_{-\infty}^{\infty} e_1(t)\, e_2(t)\, dt$)

(c) What does this tell you about the Fourier series?

6) Assume $\underline{y} = \underline{\underline{A}}\,\underline{m} + \underline{n}$

where $\langle \underline{n}\,\underline{n}^T \rangle = \sigma_n^2 \underline{\underline{I}}$ ← Identity matrix

$\underline{\underline{A}}$ is a known $M \times N$ matrix and $\underline{y}$ a measured quantity.

(a) when is the system $\underline{y} = \underline{\underline{A}}\,\underline{m} + \underline{n}$ ~~det~~ overdetermined and underdetermine? (write down the conditions)

(b) Find $\hat{\underline{m}}$ using LSQ. (NOTE: Define the cost function to begin)

(c) Now assume I told you $\langle \underline{m}\,\underline{m}^T \rangle = \underline{\underline{P}} = \begin{bmatrix} \sigma_m^2 & & \emptyset \\ & \ddots & \\ \emptyset & & \sigma_m^2 \end{bmatrix}$

Find $\hat{\underline{m}}$ using this information.

(d) Estimate the uncertainty on your estimate $\hat{\underline{m}}$ by computing $\hat{\underline{\underline{P}}} = \langle \hat{m}\,\hat{m}^T \rangle$. Is this matrix diagonal and equal to $\underline{\underline{P}}$?

Does it depend on $\underline{y}$? (If not try to explain why)

7) ~~No~~. Assume $\underline{y} = \underline{\underline{A}}\,\underline{m} + \underline{n}$ like in #6.

Suppose I also told you that $\underline{\underline{B}}\,\underline{m} = \underline{x}$ exactly.

(a) Write down a cost function to include this information. Include also $\underline{\underline{P}}$

(b) Derive the equations that will minimize you cost function. (These are called the normal equations)

8) Given $\underline{y} = \underline{\underline{E}}\,\underline{x} + \underline{n}$, $\underline{\underline{E}}^{T}$ is said to be the adjoint of $\underline{\underline{E}}$, your model.

(a) What is the meaning of the adjoint $\underline{\underline{E}}^{T}$? (In words)

(b) Based on the statistics of $\underline{x}$ and $\underline{y}$ show that $\underline{\underline{E}} = \langle \underline{y}\,\underline{x}^{T} \rangle \langle \underline{x}\,\underline{x}^{T} \rangle^{-1}$

(c) Compute the Adjoint $\underline{\underline{E}}^{T}$ and discuss its meaning. based on what you find. Does it match what you said in (a)

8) Given a series $y(t)$ with stationary statistics.

ⓐ Define what it means to be stationary.

ⓑ Derive the autocorrelation function $\alpha(\tau)$ where τ is the time lag.

ⓒ Show that the Fourier Transform of $\alpha(\tau)$ is equal to the spectrum $\hat{y}(s)\,\hat{y}^*(s)$

9) Given the following equation

$$\begin{cases} \text{I.} \quad \dfrac{dy}{dt} = -s_b\, y + f(t) & (f(t) \text{ is white noise forcing}) \\ \text{II.} \quad \dfrac{dy}{dt} = -s_b\, y + A\cos(2\pi s_a t) \end{cases}$$

with $s_b < s_a$

ⓐ Compute the analytical spectrum of $y(t)$ in I. and II.

ⓑ Draw the spectra and label the frequency s_b, s_a

ⓒ Discuss the shape of the spectra.

10) Suppose you have a time series

$$y(t_n) \qquad t_n = n\,\Delta T \qquad n = 1 \dots N$$

with an autocorrelation

$$\alpha(\Delta T) = e^{-\Delta T / 4\Delta T}$$

(a) Plot $\alpha(\Delta T)$ as a function of ΔT

(b) What are the degrees of freedom of $y(t_n)$

(c) Explain why the actual degrees of freedom $N^* < N$ and make a plot of

(d) Use $\alpha(\Delta T)$ to write down ~~an~~ the coefficients a_1 and a_2 of an AR-2 (autoregressive process of order 2)

$$x(t_n) = a_1 x(t_{n-1}) + a_2 x(t_{n-2}) + n(t_n)$$

11) Assume you have the following dynamics

$$\frac{d^2 y}{dt^2} = -s_b y + f(t) + A\cos(2\pi s_a t)$$

and you computed the spectrum of $y(t)$.
You see some peaks and you want to know if they are significant.

ⓐ Which AR process would you use as your null hypothesis (NOTE: $\frac{d^2y}{dt^2} = \frac{y_{t+1} + y_{t-1} - 2y_t}{\Delta t^2}$)

ⓒ Without solving for the spectra how many significant peaks would you guess and at what frequency?

ⓑ If the spectral shape of the AR process you picked is given by the function $P_{AR}(s)$ and you had 10 realizations of $y(t)$ to use for computing $Y(s)Y^*(s)$, how would you estimate the 95% confidence limit?

12) The Discrete Fourier Transform

$$y(t) = A_0 + \sum_{k=1}^{N/2-1} \left\{ A_k \cos\left(\frac{2\pi k t}{T}\right) + B_k \sin\left(\frac{2\pi k t}{T}\right) \right\}$$

$$+ A_{N/2} \cos\left(\frac{\pi N t}{T}\right)$$

(a) What is the Nyquist frequency and its amplitude.

(b) Which is the mean?

(c) What is the variance explained by the periodicity $s_a = \frac{10}{T}$?

(d) Using the orthogonality relationship how can you compute A_{10} and B_{10}?

(Assume $A_0 = \emptyset$)

13) Draw the autocorrelation function and spectrum of the following processes

(a) white noise

(b) red noise

(c) $y = at$

(d) $y = a\cos(2\pi s a t)$

14) (a) Explain the concept of Tapering?

(b) when we have a finite time series and we take a spectrum what kind of tapering is implicit? (If we do none)

what does this do to the spectral peaks estimate? ~~[illegible]~~

15) The test of variance says

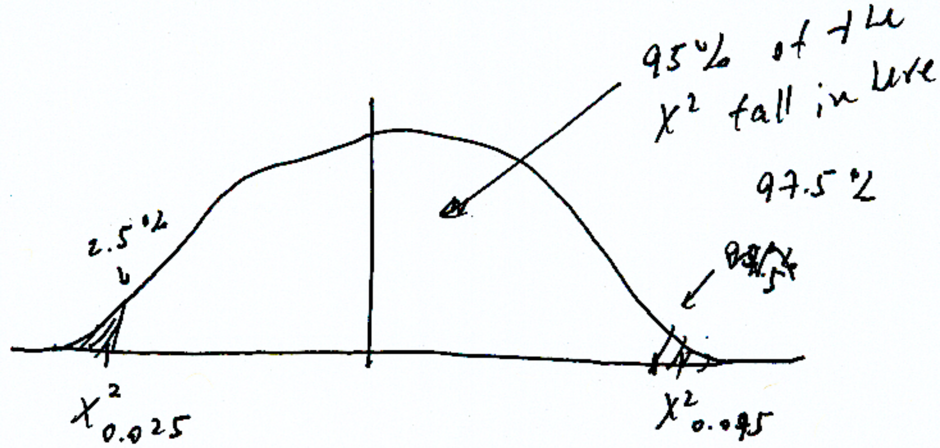
confidence limit → $\dfrac{\nu s^2}{\chi^2_{0.025,\nu}} < \sigma^2 < \dfrac{\nu s^2}{\chi^2_{0.975,\nu}}$

σ^2 = true variance

s^2 = sample variance

where r = degree of freedom in the estimate of s^2

and



ⓐ Given that the sample variance is computed from a variable $x(t_n)$ with $n = 1 \dots N$

derive the identity

$$r s^2 = \sigma^2 \chi^2$$

where r are the degrees of freedom in computing s^2

$r = N-1$

ⓑ suppose $x(t_n)$ has an autocorrelation

$\alpha(\Delta T) = e^{-\Delta T/4\Delta T}$ and lenght N. How does that affect the confidence limit? Will they be higher of smaller?