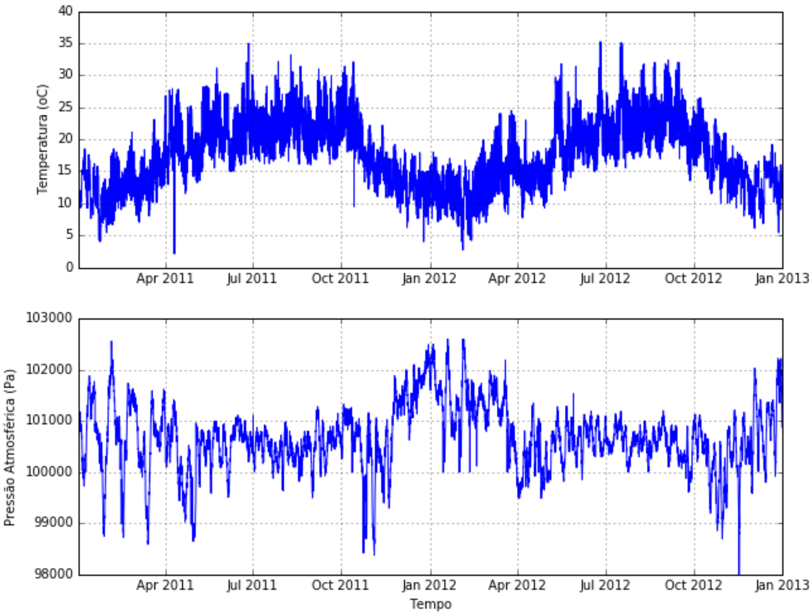
Modelação Numérica (2016/17)
Exame, 2^a Época

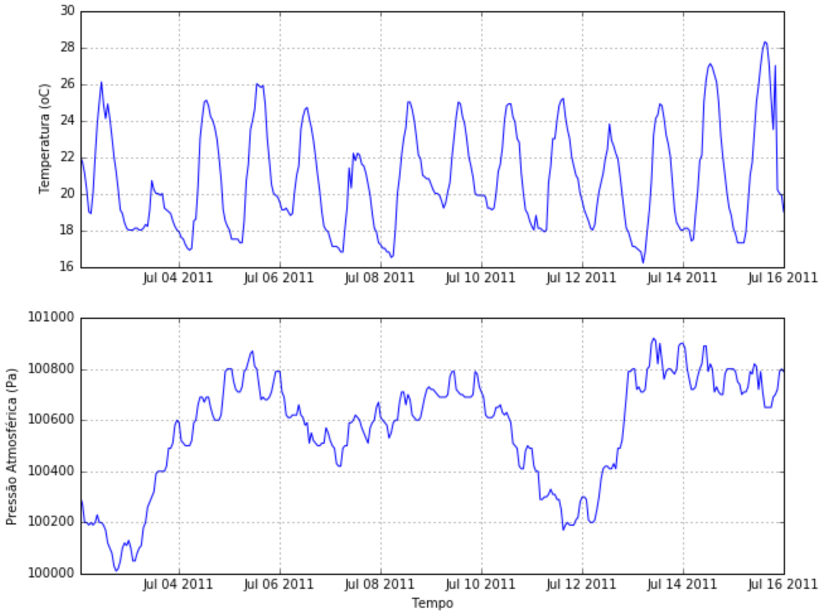
27/Junho/2017
L-MOG, L-EG, MI-EEA
DEGGE, FCUL

1 Exercício 1 - Análise digital de séries temporais (7 valores)

Observa os gráficos seguintes, que mostram a variação da temperatura e pressão atmosférica em Lisboa nos anos de 2011 e 2012.



Observa agora as mesmas séries temporais, com maior resolução num período de duas semanas em Julho de 2011.



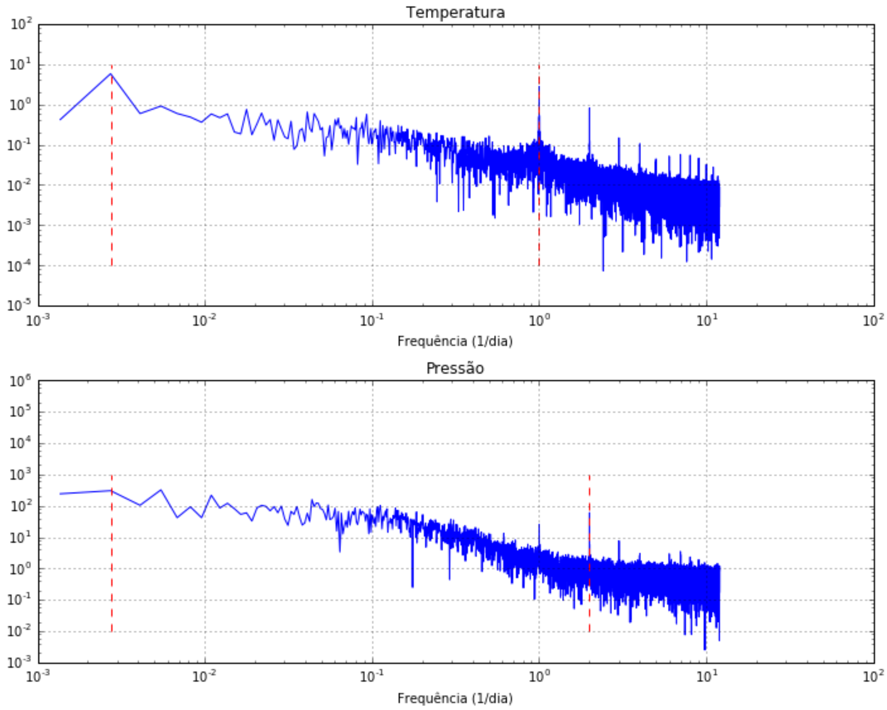
1. **Com base nos gráficos acima, consegues identificar periodicidades nestas séries de temperatura e pressão atmosférica? Se sim, quais são as periodicidades que consegues identificar, e quais as amplitudes correspondentes?**

   A temperatura tem uma variação anual (mais alta no Verão e mais baixa no Inverno), com uma amplitude de aproximadamente $10\,^oC$. Tem também uma variação diurna (valores mais altos durante o dia e mais baixos durante a noite), com uma amplitude de aproximadamente $4\,^oC$ (estimada visualmente a partir do gráfico).

   A pressão tem uma variação anual (valores mais altos no Inverno e mais baixos no Verão), com uma amplitude de aproximadamente $1000\,Pa$. Tem também uma variação semi-diurna (dois ciclos por dia), com uma amplitude de 50–100 Pa.

2. **Desenha o espectro de amplitudes das séries de temperatura e de pressão atmosférica. Lembra-te de identificar ambos os eixos e as suas escalas.**

   No espectro, devemos desenhar picos correspondentes às periodicidades que identificámos acima, com as respectivas amplitudes. O espectro deve ainda ter alguma energia em toda a gama de frequências, dado o elevado nível de ruído do sinal.



As linhas a tracejado vermelho identificam os períodos de 365 dias, 1 dia, e 0.5 dia (frequências de $1/365$, 1 e 2 dia^{-1}, respectivamente).

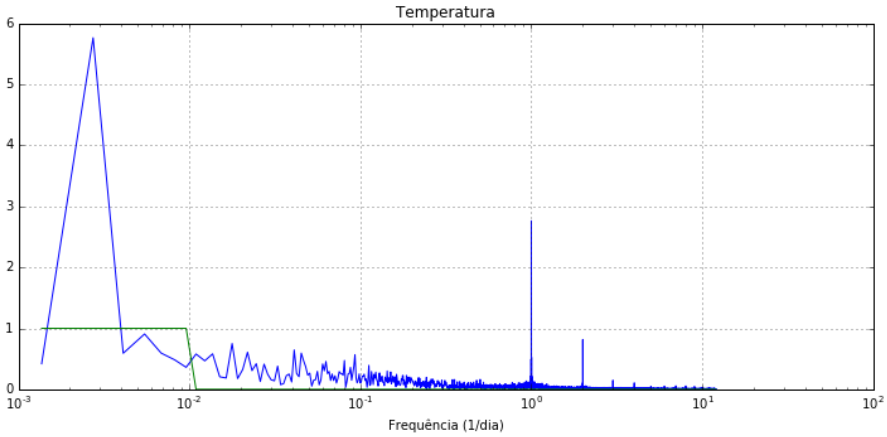
3. **As séries estão amostradas de hora em hora. Conseguirás identificar variações de temperatura com um período de uma hora? Porquê?**

   Como as séries estão amostradas de hora em hora ($dt = 1\,h$), e como precisamos de pelo menos dois pontos para amostrar bem um período, o período mais curto que conseguimos representar correctamente com o intervalo de amostragem $dt = 1\,h$ é $2\,dt = 2\,hr$. Assim sendo, não é possível representar corectamente variações com períodos de 1 hora com esta taxa de amsotragem.

4. **Se quiseres estudar a variação anual destas séries de dados, que tipo de filtro podes usar? Desenha o filtro e diz como o aplicarias.**

   Um dos filtros que poderia utilizar seria um filtro passa-baixo, que deixasse passar as baixas frequências e eliminasse as altas frequências, com uma frequência de corte um pouco acima de $1/365\ dia^{-1}$.

   Um possível exemplo é dado pela função de transferência desenhada a verde na figura abaixo:



Aplicaria o filtro da seguinte forma:

(a) Calcular o espectro (ou seja, a transformada de Fourier) da função: $\mathcal{F}(f(t)) = F(f)$.

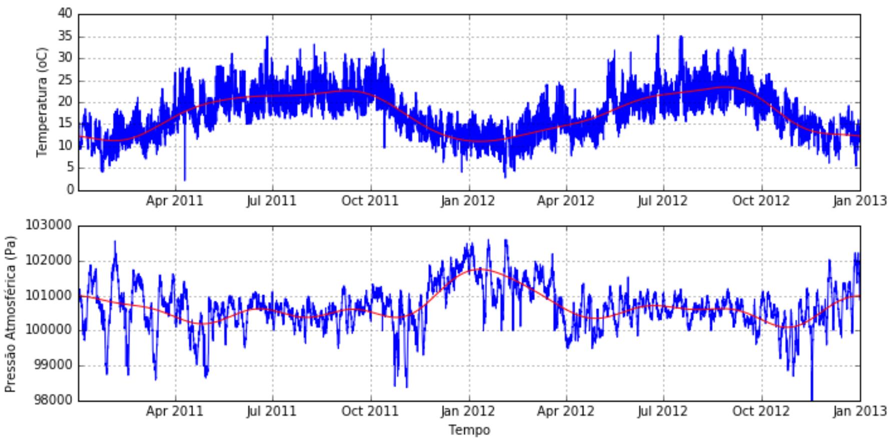
(b) Construir a função de transferência $H(f)$, de modo a deixar passar apenas a frequência desejada ($1/365\,dia^{-1}$, ver desenho acima, curva a verde) e eliminar as restantes frequências do sinal.

(c) Multiplicar o espectro da função pela função de transferência: $F(f) \times H(f)$.

(d) Passar o espectro filtrado de volta para o domínio do tempo, aplicando a transformada inversa de Fourier: $\mathcal{F}^{-1}\left(F(f) \times H(f)\right)$.

5. **Desenha a série que esperarias obter depois de aplicar o filtro.**

Depois de aplicar o filtro passa-baixo acima, iria obter uma série em que retinha apenas as variações de longo período (a vermelho nos gráficos abaixo).



6. **Que resultado esperarias obter se calculasses a correlação cruzada entre as duas séries filtradas? Que informação daí poderias tirar?**

Fazendo a correlação entre os dois sinais, ficaria a conhecer quão semelhantes eles são, e qual o desfasamente entre si. Como uma das séries alcança os seus valores máximos no Inverno e a outra no Verão, esperaria que o extremo da correlação fosse um valor negativo. Ou seja, as duas séries devem ser anti-correlacionadas. O tempo em que o valor máximo absoluto da correlação se regista deve corresponder ao desfasamento entre os dois sinais.

2 Exercício 2 - Equações diferenciais parciais (7 valores)

O som propaga-se como ondas acústicas (perturbações de pressão) em gases, líquidos e sólidos. Considera uma onda de som que se propaga a uma dimensão, sem atenuação:

$$\frac{\partial^2 p}{\partial t^2} = v^2 \frac{\partial^2 p}{\partial x^2}$$

Onde p é a pressão acústica (desvio em relação à pressão ambiente) e v é a velocidade a que o som se propaga (que podemos assumir ser uma constante conhecida).

A perturbação pressão inicial de pressão é $0.01\,Pa$, no local onde o som tem origem. A onda propaga-se ao longo de 5 m de distância, até um receptor.

1. **Discretiza a equação acima.**

$$\frac{p_i^{n+1} - 2p_i^n + p_i^{n-1}}{\Delta t^2} = v^2 \frac{p_{i+1}^n - 2p_i^n + p_{i-1}^n}{\Delta x^2}$$

Onde n é o índice do tempo, e i é o índice do espaço, na direcção x.

2. **Como podes calcular a pressão acústica, para cada ponto ($x = x_i$), no tempo seguinte ($t = t^{n+1}$)? (Isto é, re-escreve a expressão na forma $p_i^{n+1} =$ ...). Trata-se de uma esquema explícito ou implícito? Porquê?**

$$p_i^{n+1} = v^2 \frac{\Delta t^2}{\Delta x^2} \left(p_{i-1}^n - 2p_i^n + p_{i+1}^n\right) - p_i^{n-1} + 2p_i^n$$

Trata-se de um método explícito, porque o ponto que queremos calcular (p_i^{n+1}) depende apenas de pontos atrás no tempo (p^n e p^{n-1}). Assim sendo, podemos calcular p_i^{n+1} explicitamente em função de pontos atrás no tempo, já conhecidos.

3. **Num cenário em que o receptor não tem influência na propagação da onda, que tipo de condição fronteira utilizarias nesse extremo do domínio para resolver o problema? Como a implementarias?**

No extremo em que se encontra o receptor, o domínio é aberto, pelo que aplicamos uma fronteira aberta: p[nx-1]=p[nx-2], onde nx é o número de pontos ao longo da barra (em x, no espaço).

4. **Dá um exemplo de um cenário em que tivesses de aplicar uma condição fronteira diferente no receptor, indicando o tipo de fronteira que utilizarias.**

Se o houvesse uma parede a $5\,m$ da fonte acústica, o som seria reflectido, e teríamos de implementar uma fronteira fechada.

5. **Como discretizarias o mesmo problema, se a propagação da onda se fizesse a duas dimensões?**

$$\frac{\partial^2 p}{\partial t^2} = v^2 \left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2}\right)$$

A equação seria agora discretizada da seguinte forma:

$$\frac{p_{i,j}^{n+1} - 2p_{i,j}^{n} + p_{i,j}^{n-1}}{\Delta t^2} = v^2 \left(\frac{p_{i+1,j}^{n} - 2p_{i,j}^{n} + p_{i-1,j}^{n}}{\Delta x^2} + \frac{p_{i,j+1}^{n} - 2p_{i,j}^{n} + p_{i,j-1}^{n}}{\Delta y^2} \right)$$

Onde adicionámos o índice j, correspondente ao domínio espacial na direcção y.

A equação a implementar no código seria a seguinte:

$$p_{i,j}^{n+1} = v^2 \Delta t^2 \left(\frac{p_{i-1,j}^{n} - 2p_{i,j}^{n} + p_{i+1,j}^{n}}{\Delta x^2} + \frac{p_{i,j-1}^{n} - 2p_{i,j}^{n} + p_{i,j+1}^{n}}{\Delta y^2} \right) - p_{i,j}^{n-1} + 2p_{i,j}^{n}$$

6. **Quais as principais diferenças entre as implementações computacionais do problema a 1D e a 2D?**

   A principal diferença seria que quando estivessemos a trabalhar com um domínio espacial 2D, a matriz de pressões teria de ser uma matriz 2D.

3 Exercício 3 - Optimização de parâmetros (6 valores)

Fez-se uma experiência sísmica para determinar a estrutura do sub-solo. Sabe-se que a estrutura do sub-solo pode ser bem aproximada por um modelo de duas camadas horizontais. A experiência consistiu em emitir ondas sísmicas e medir os seus tempos de chegada a pontos às seguintes distâncias da fonte sísmica:

$$x = (6.0, 10.1, 14.3, 18.4, 22.5, 26.7)$$

Os tempos de chegada das ondas sísmicas aos vários pontos foram os seguintes:

$$t = (3.5, 4.3, 5.1, 5.8, 6.9, 8.1)$$

O tempo de chegada das ondas sísmicas às estações relaciona-se da seguinte forma com a estrutura do sub-solo:

$$t_i = t_A + s_B x_i$$

Onde t_i são os tempo de chegada das ondas sísmicas a cada estação, t_A depende da espessura e da lentidão (inverso da velocidade de propagação das ondas, $1/v$) da camada superior (A), e s_B é a lentidão da camada inferior (B).

1. **Identifica as observações feitas e escreve-as na forma de um vector $\vec{o}$. Faz o mesmo para os parâmetros a inverter/optimizar e escreve o vector $\vec{p}$.**

   As observações são os vários tempos de chegada das ondas às várias estações $(t_1, t_2, ..., t_6)$:

$$\vec{o} = \begin{bmatrix} o_1 \\ o_2 \\ o_3 \\ ... \\ o_N \end{bmatrix} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ ... \\ t_6 \end{bmatrix}$$

   Os parâmetros a optimizar são t_A e s_B:

$$\vec{p} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} t_A \\ s_B \end{bmatrix}$$

2. **Este problema é sub-determinado ou sobre-determinado? Qual o número de parâmetros de que precisarias para obter um problema equi-determinado?**

   Este problema é sobre-determinado, porque temos mais observações do que parâmteros a determinar. Para o problema ser equi-determinado, teríamos de ter 2 observações para constranger os 2 parâmetros.

3. **Podes resolver este problema utilizando um algoritmo linear? Se sim, escreve as equações matriciais do problema a resolver, na forma $\vec{o} = G\vec{p}$. Indica como procederias para determinar os parâmetros ($\vec{p}$).**

   Sim, a relação entre as observações e os parâmetros é linear, por isso podemos resolver o problema utilizando um algoritmo de inversão linear.

$$\begin{bmatrix} o_1 \\ o_2 \\ ... \\ o_6 \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \\ G_{...} & G_{...} \\ G_{61} & G_{62} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

$$\begin{bmatrix} t_1 \\ t_2 \\ ... \\ t_6 \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ ... & ... \\ 1 & x_6 \end{bmatrix} \begin{bmatrix} t_A \\ s_B \end{bmatrix} = \begin{bmatrix} t_A + s_B x_1 \\ t_A + s_B x_2 \\ ... \\ t_A + s_B x_6 \end{bmatrix}$$

   Para determinar quais os parâmetros que constituem o vector $\vec{p}$, basta inverter a matriz G e multiplicá-la pelas observações $\vec{o}$:

$$\vec{o} = G\vec{p} \;\Leftrightarrow\; G^{-1}\vec{o} = G^{-1}G\vec{p} \;\Leftrightarrow\; \vec{p} = G^{-1}\vec{o}$$

   Neste caso, como a matriz G não é quadrada, não podemos invertê-la directamente. Teremos de calcular a sua inversa generalizada/pseudo-inversa.

4. **É possível resolver este problema utilizando um algoritmo não-linear? Se sim, qual seria a função de custo que utilizarias?**

   Sim, os algoritmos não lineares podem ser utilizados para resolver tanto problemas de inversão lineares como não lineares.

   Uma possível função de custo, $C1$, seria:

$$\begin{aligned} C1 &= \sum_{i=0}^{N} |obs_i - mod_i| \\ &= \sum_{i=0}^{N} |t_i - (t_A + s_B x_i)| \end{aligned} \tag{1}$$

5. **A norma que utilizaste acima é L1 ou L2? Porque razão podes preferir utilizar uma ou a outra?**

   A norma acima é L1. A norma L2 para a mesma função de custo seria:

$$
\begin{aligned}
C2 &= \sum_{i=0}^{N} |obs_i - mod_i|^2 \\
&= \sum_{i=0}^{N} |t_i - (t_A + s_B x_i)|^2
\end{aligned}
\tag{2}
$$

A norma L2 penaliza mais as diferencas maiores entre observações e previsões do modelo. Assim sendo, utilizamos a norma L2 quando queremos evitar grandes desajustes entre observações e previsões do modelo.

6. **Em que situações é vantajoso usar um método de inversão linear ou não-linear?**

   É preferível utilizar um método de inversão linear, sempre que o problema a resolver exiba uma relação linear entre observações e parâmetros. As inversões lineares são computacionalmente eficientes, e existe todo um formalismo matemático que permite avaliar as incertezas dos parâmetros.

   As inversões não lineares são computacionalmente mais pesadas, mas permitem explorar eficientemente um grande espaço de parâmetros, mesmo quando o problema a resolver não exibe uma relação linear entre observações e parâmetros.