**Modelos lineares baseadosem filtros Auto Regressivos**

Afonso Araújo – Orientador

Lucas Ribeiro - Aluno

**Motivação**

Muito se tem feito para melhora no setor de distribuição energética no Brasil com diversas frentes de fontes de energia, eólica, termoelétricas, nucleares, solar e hidrelétricas. As eólicas tem um grande potencial ainda a ser explorado, as termo são caras e possuem elevado nível de poluição da mesma forma que a energia nuclear, a solar vêm ganhando muito espaço e ficando cada vez mais barata, mas não há ainda como competir com a hidrelétrica quando se fala em quantidade de energia gerada comparado a poluição gerada e o custo da energia, porém a hidrelétrica depende de algo que não é controlado, chuvas para encher os reservatórios.

Prever chuvas têm sido cada vez mais importante para toda a sociedade, seja para o setor agrícola, enchentes em cidades e também para o setor elétrico. Com uma previsão mais próxima o possível do real é possível saber quando os reservatórios estarão com armazenamento de água adequado para que seja possível gerar energia. Um outro ponto é a previsão de preço da energia, que varia de acordo com a capacidade geradora das hidrelétricas e a necessidade de acionamento das termoelétricas e nucleares (energias mais caras). Saber quanto a energia vai custar é de grande interesse para todo o setor saber quando realizar investimentos como para o consumidor se preparar.

É nessa motivação que temos a proposta de construir um modelo para previsão de cheias e vazias dos reservatórios das bacias hidrográficas, modelo este baseado em filtros lineares auto regressivos, que seja open source e que esteja disponível online.

**Proposta**

* Inicialmente simular vazões de cheias pela da identificação de componentes hidrológicas através de filtros lineares auto regressivos.
* Cada componente será identificado como um modelo auto regressivo com uma variável exógena de ordem p, ARX(p)
* As componentes hidrológicas serão separadas através de um processo de filtragem
* A ordem p e os respectivos coeficientes serão identificados por meio das funções de auto correlação e auto correlação parcial

**Revisão Teórica**

Uma breve revisão teoria do modelo e como ele funciona será apresentada agora.

Existem os modelos lineares auto regressivos estacionários e não estacionários. Os estacionários possuem média, variância e auto covariância são constantes; já os não estacionários, esses valores variam no tempo. Usaremos os modelos Auto regressivos AR(p), denominados de auto regressivos de ordem p e tem a seguinte forma:

Zt = φ1\*Zt−1 + φ2\*Zt−2 + ... + φp\*Zt−p + at

A série 𝑍𝑡 possui valores que são combinações lineares dos p valores passados mais um termo 𝑎𝑡 , no qual incorpora coisas na série até o tempo t que não é explicado pelos valores passados.

O modelo auto regressivo AR(p) não leva em consideração forças externas então caso seja necessário um modelo que trabalhe com forçantes usaremos o modelo ARX(p) onde o termo “X” seria um forçante do sistema, no nosso caso a chuva ou precipitação.

**Processo de filtragem**

* **Codificação e linguagem de programação**

Para iniciar o processo, foi analisado o tipo de arquivo cedido, no caso, .csv. Inicialmente a programação do filtro seria na linguagem C, porém devido a algumas complicações da própria linguagem e com o intuito de agilizar o processo foi adotado o Python 3.6, por ser mais simples e funcional para leitura do arquivo além de ter toda uma gama de bibliotecas prontas para uso.

* **Análise de Recessão**

Os dados cedidos têm as informações das vazões nos postos com uma escala de tempo diária. Com isso o código até o momento só funciona com escalas de tempo diárias.

Para encontrar as recessões foi criado um arquivo.py somente para ter esse retorno de informações e as mesmas são encontradas quando em intervalo de tempo após o atual a vazão diminuiu, ou seja, quando temos uma derivada negativa e para realizar essa análise e saber onde temos uma derivada negativa, foi utilizado a função “gradiant()” da biblioteca “numpy”. A lógica para realização dessa verificação é simples, ao observar que num determinado ponto temos uma derivada negativa indica que o mesmo é um ponto de recessão, entretanto, também é preciso ter atenção a outros parâmetros.

* + **Tempo para início de recessão (1)**

Foi observado que em alguns postos temos “alarmes falsos” de recessão, devido a uma escala de tempo não apropriada para o devido posto. Com isso um foi criado o parâmetro “tempo para inicio de recessão” que faz uma contagem de dias até efetivamente começar uma análise.

* + **Tempo para fim de recessão (2)**

Esse parâmetro tem a mesma intenção do anterior, porém no final de cada recessão (um complementa o outro)

* + **Variação mínima de vazão (3)**

Esse parâmetro tem a função de filtrar variações de vazão muito próximas visto que pode estar dentro da margem de erro da verificação da vazão da bacia.

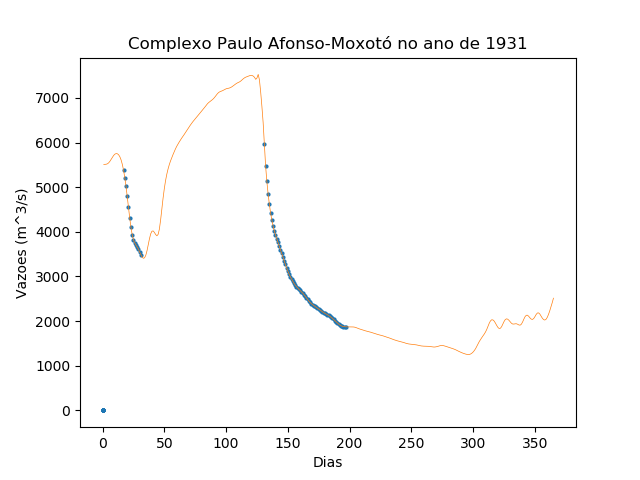
Os parâmetros citados anteriormente foram adquiridos por um arquivo cedido pelo Cepel onde constam constantes de cada bacia que podem ser interpretadas/utilizadas como essas constantes de verificação da analise de recessão. Infelizmente nesse arquivo, nem todas as bacias contam com informações, com isso é feito uma analise mais estimada do valor dessas constantes com base na plotagem da vazão durante o período anual.

Figura : Recessão

A lógica utilizada para a análise de recessão foi: a cada instante de tempo (dia), verifica se a derivada é negativa naquele ponto e também nos pontos seguintes (1)(2), caso seja é adicionado a uma matriz de valores somente compostos por pontos de recessão, no fim de uma recessão é feito a análise de variação de vazão (3), avaliando a margem de erro da leitura da bacia. Caso não seja negativa a derivada, ou não seja um ponto de recessão utilizável para análise não é feito nada.

* **Diferentes constantes de tempo**

Como podemos observar na figura 1, temos diferentes escalas de tempo para a variação da vazão, em certos pontos ela cai rapidamente, mas em outros ela varia de forma lenta, introduzimos então as diferentes componentes que estarão em análise nas bacias.

* + **Componente lenta**

Essa componente é atribuída ao escoamento subterrâneo da bacia

* + **Componente intermediária**

Essa componente é atribuída ao escoamento entre o subterrâneo e o superficial da bacia

* + **Componente rápida**

Essa componente é atribuída ao escoamento superficial da bacia

**Construção do Modelo AR(1)**

A construção do modelo numero de um filtro auto regressivo de primeira ordem foi totalmente baseada da dissertação de ***MESTRADO?*** de Afonso Araújo.

Sob essas circunstâncias temos a forma diferencial da equação da componente lenta do filtro é semelhante à de um filtro passa baixa:

Através da Transformada Z é obtido a equação abaixo do modelo AR (1):

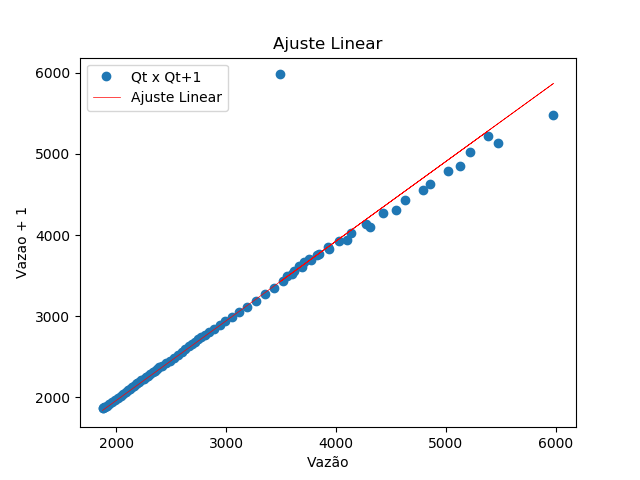
Onde:

K é o coeficiente angular do gráfico

é o intervalo de discretização

A equação por completo fica:

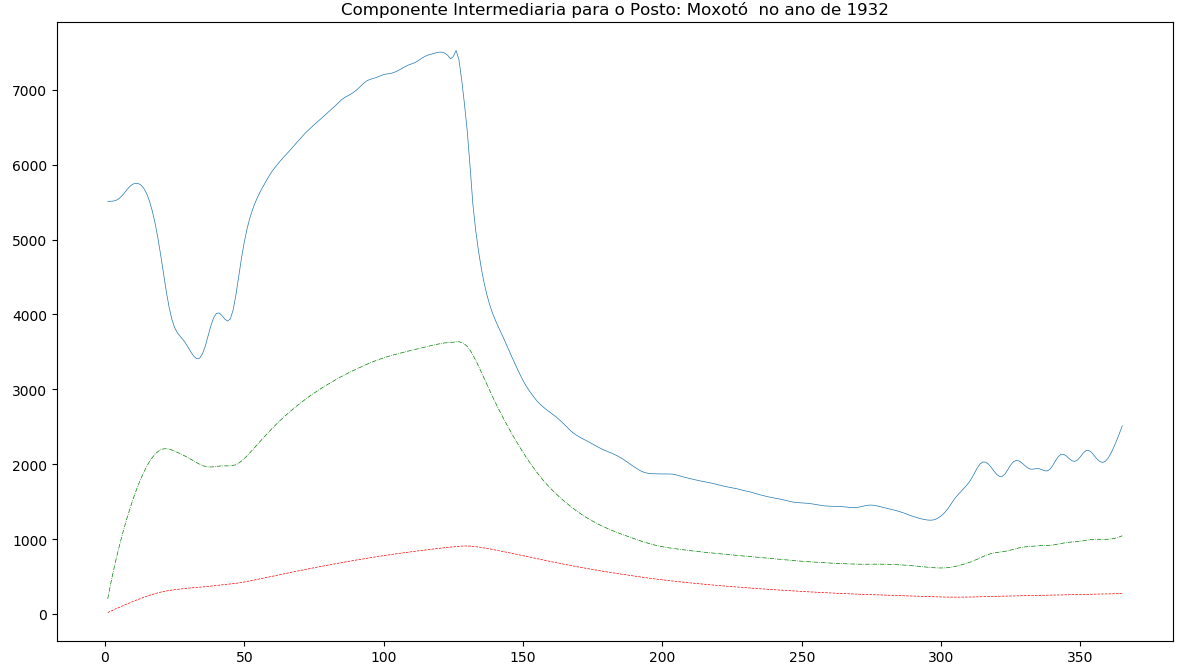
Alpha é ajustado por tentativa e erro, sua função será explicada a seguir.

 A lógica é bem simples, logo após realizar a análise de recessão, é plotado o gráfico e nele feito uma regressão linear para encontrar o valor de K, caso detenhamos mais de um K como na imagem abaixo, a qual notamos uma segunda tendência se formando, é feito um novo ajuste linear naquela região para encontrar outro valor de K.

Quando temos um único K, filtramos o yn com o filtro AR (1) e obtemos o yn(1) que é a componente lenta, a componente intermediária é obtida da subtraindo o a vazão total da recém filtrada.

Quando temos dois K’s, o procedimento inicial é o mesmo até a obtenção de yn(2) pela equação acima, nesse momento, o sinal yn(2) é novamente submetido ao filtro AR (1), porém com o segundo K, obtido, então obtemos um yn(2)’’, componente intermediária atualizada, e a componente rápida é obtida com a mesma lógica, subtraindo yn(2) de yn(2)’’.

Para realização da regressão linear, foi feito um novo arquivo separado e é utilizada biblioteca “numpy” e a função é a “np.linalg.lstsq()”, o importante para utilização desta função é entender que ela faz a regressão com uma matriz linha de valores com uma matriz coluna valores, e para converter a matriz linha para coluna foi utilizado outra função também da biblioteca numpy, “np.asarray()”.  
 Como já foi dito, as bacias podem ter mais constantes de tempo, e quando isso acontece, foi criado uma variável para selecionar se a separação de constantes será feita. Quando é feita o método utilizado foi de input do usuário com a função plt.ginput(), o usuário clica na posição do gráfico que ele julga ser a intercessão entre as duas constantes de tempo, em seguida é separada em duas listas e as regressões são refeitas resultando nos K’s dos dois tipos de reservatórios da bacia.

 Para o filtro auto regressivo, novamente foi criado um novo arquivo.py, e nele é importado todos os arquivos anteriores até o momento e o processo de filtragem é feito de forma interativa em um laço de for, não foi utilizada biblioteca externa.

Temos acima o resultado da filtragem.

Em cor azul temos a vazão total, em vermelho pontilhado temos a componente lenta, e em verde pontilhado temos a componente intermediária/rápida.