Macroeconometria 1

Mestrado em Economia Monetária e Financeira

Mestrado em Economia

ISCTE-IUL, Dep. de Economia

MÉTODOS TRADICIONAIS DE PREVISÃO

Luís Filipe Martins

luis.martins@iscte.pt

http://iscte.pt/~lfsm

Departamento de Métodos Quantitativos,

ISCTE-IUL, Escola de Gestão

Lisboa, Setembro de 2009

1 Introdução: Medidas de erro de previsão

- Diferentes métodos de previsão produzem diferentes resultados, pelo que um dos critérios possíveis (e fundamental) para a selecção de um modelo é a precisão relativa das suas previsões.
- Erro de previsão (diferença entre a observação real/efectiva de y no período genérico t e o valor previsto para a variável nesse período $\hat{y}_t$, obtido pelo modelo e com base nas observações passadas):
$$e_t = y_t - \widehat{y}_t.$$
- $\widehat{y}_t \equiv \widehat{y}_{t|t-h}$, previsão para o período t obtida no período $t - h$ $(h = 1, 2, ...)$.
- Notação alternativa: $\widehat{y}_{t+h} \equiv \widehat{y}_{t+h|t}$ ou mesmo $\widehat{y}_{t+h-1} \equiv \widehat{y}_{t+h-1|t-1}$.
- Normalmente, usa-se o erro a um passo $\widehat{y}_{t|t-1}$, onde $h = 1$. A sequência $\left\{\widehat{y}_{t|t-1}\right\}_{t=1}^{T}$ gera $\{e_t\}_{t=1}^{T}$.
- $e_t \approx 0$ se o modelo for razoável (o modelo descreve bem o padrão comportamental da série, restando apenas flutuações aleatórias/erráticas pouco significativas que têm elas próprias origem em fenómenos não previsíveis).
- Na estimação do modelo, bem como a comparação entre modelos de previsão, usam-se critérios de minimização da

função dos erros de previsão. Entre outras medidas...

- Erro Absoluto Médio (EAM)

$$EAM = \frac{\sum_{t=k}^{T} |e_t|}{T-k+1}, k = 1, 2, ..., T$$

- Erro Percentual Absoluto Médio (EPAM)

$$EPAM = \frac{\sum_{t=k}^{T} \frac{|e_t|}{y_t}}{T-k+1} \times 100$$

Este critério relativo permite ainda avaliar a precisão de um dado método relativamente a séries diferentes.

- Erro Percentual Médio (EPM)

$$EPM = \frac{\sum_{t=k}^{T} \frac{e_t}{y_t}}{T-k+1} \times 100$$

Pode ser utilizado como medida do enviesamento previsional (sub ou sobre-previsão). Assume-se e_t com igual sinal (caso contrário, podem-se anular erros de diferente sinal mas de igual magnitude. Eg: $-10 + 10 = 0!$)

- Erro Quadrático Médio

$$EQM = \frac{\sum_{t=k}^{T} e_t^2}{T-k+1}$$

A raíz quadrada deste valor, $REQM = \sqrt{\sum_{t=k}^{T} e_t^2 / (T-k+1)}$ pode ser utilizada como estimativa do desvio padrão do erro de previsão a um passo, quando representado por e_t.

- Estatística U de Theil

$$U = \frac{REQM \text{ do modelo de previsão}}{REQM \text{ do modelo simplista}}$$

em que o modelo simplista (ou naive) utiliza como previsão para o período $t+1$ o valor observado em t ($\widehat{y}_{t+1} = y_t$). A interpretação da estatística é a seguinte:

- $U = 1$: previsões do modelo são equivalentes à do modelo simplista

- $U < 1$: o modelo considerado é superior ao modelo naive

- $U > 1$: o modelo prevê pior que o método simplista

- $U = 0$ (qual a interpretação?)

Esta estatísitca pode ainda ser utilizada para comparar diferentes métodos alternativos, para além do naive.

- Teste do Sinal

Dados dois métodos A e B, considere-se o diferencial da medida do erro (também pode ser EQM ou EAM, para além do erro de previsão simples e_{t+j-1})

$$d_j = e_{t+j-1}^{(A)} - e_{t+j-1}^{(B)}, \qquad j = 1, ..., m.$$

Pode-se efectuar um teste para H_0 : Não há diferenças significativas entre A e B, através da estatística

$$S = \frac{2}{\sqrt{m}} \sum_{j=1}^{m} [I(d_j > 0) - 0.5] \xrightarrow{d} N(0,1), \text{ sob } H_0$$

onde $I(.)$ é uma função indicatriz que assume os valores 0 e 1 consoante a condição entre (.) é verificada ou não.

- Teste de Diebold-Mariano (DM)

Este teste utiliza igualmente os diferencias das medidas de erro de previsão. Considere-se o diferencial médio

$$\overline{d} = \frac{1}{m} \sum_{j=1}^{m} d_j$$

e a hipótese nula de que os diferenciais são nulos $(H_0 : d = 0)$. Em rigor, a hipotese nula é $E\left(d_j\right) = 0$ A estatística de teste é

$$DM = \frac{\overline{d}}{\sqrt{\omega}} \xrightarrow{d} N(0,1), \text{ sob } H_0,$$

sendo ω estimado através de

$$\widehat{\omega} = \sum_{i=-(h-1)}^{h-1} \widehat{\gamma}_i(d),$$

sendo $\widehat{\gamma}_i(d)$ as autocovariâncias amostrais dos d_j. No caso de erros a 1 passo $(h = 1)$, $\widehat{\omega} = \widehat{\gamma}_0$, a variância amostral de d_j.

- Esquemas alternativos para estimar os parametros θ dos modelos usados para construir as previsões (para o caso, $h = 1)$:
  - Recursive: Estimar θ com $t = 1, ..., R$ $\rightarrow$prever $\widehat{y}_{R+1}$ $\rightarrow$ Re-estimar θ com $t = 1, ..., R + 1$ usando $\widehat{y}_{R+1}$ $\rightarrow$prever $\widehat{y}_{R+2} \rightarrow \ldots$

- Rolling: Estimar θ com $t = 1, ..., R$ $\rightarrow$prever $\widehat{y}_{R+1}$ $\rightarrow$ Re-estimar θ com $t = 2, ..., R+1$ usando $\widehat{y}_{R+1}$ $\rightarrow$prever $\widehat{y}_{R+2}$ $\rightarrow$Re-estimar θ com $t = 3, ..., R+2$ usando $\widehat{y}_{R+2}$ $\rightarrow$prever $\widehat{y}_{R+3}$ $\rightarrow$ ...
- Fixed: Estimar θ com $t = 1, ..., R$ $\rightarrow$prever $\widehat{y}_{R+1}$ $\rightarrow$ prever $\widehat{y}_{R+2}$ usando a estimativa obtida$\rightarrow$ prever $\widehat{y}_{R+3}$ usando a estimativa obtida$\rightarrow$ ...

2 Análise de Decomposição

- É usual, quando se analisa uma série temporal, considerá-la como o produto de um conjunto de efeitos/padrões/componentes. Importa então identificar esses efeitos (ou combinação de efeitos) e a partir daí construir previsões.
- Uma série temporal pode, em geral, ser descrita através dos seguintes componentes:

(1) Tendência: Efeitos persistentes/permanentes sobre o nível da série durante um período longo de tempo. Esses efeitos traduzem-se num movimento ascendente ou descendente da série, ou seja, de crescimento/decrescimento de longo prazo da série. Quando não existe uma tendência claramente definida, diz-se que a série é estacionária (embora seja mais correcto referirmos uma série sem tendência)

(2) Componente sazonal: Referente às oscilações em relação à tendência, mas com periodicidade anual ou infra-anual. Só

é detectável em séries com uma frequência inferior ao ano, tendo origem em causas naturais (clima, por exemplo) ou sociais (hábitos de consumo, etc.)

(3) Componente cíclica: Flutuações recorrentes com prazo superior a um ano, medindo-se normalmente de pico a pico (recessão a recessão, por ex.). Em séries económicas, a periodicidade é pouco definida e geralmente é complicada de analisar formalmente.

(4) Componente residual/irregular/errática: Designa movimentos erráticos que não têm um padrão bem definido. As causas são atribuídas a influências não controláveis, não mensuráveis ou desconhecidas, pelo que imprevisíveis. Em suma, são todos os efeitos não explicados pelos componentes anteriores.

- Objectivo: Procurar identificar os padrões 1, 2 e 3, sempre que estes estejam presentes, ”decompondo” a série de forma a que o que fica por explicar sejam apenas movimentos irregulares e sem padrão, com propriedades estatísticas adequadas (próximas do ”ruído branco”).

- Esta perpectiva é apresentada de uma forma sistematizada e intuitiva. Formalmente, esta análise traduz-se no modelo genérico
$$y_t = f(T_t, S_t, C_t, \varepsilon_t),$$
em que a previsão de y_t é feita através de $\widehat{y}_t = f(\widehat{T}_t, \widehat{S}_t, \widehat{C}_t)$, ou seja, numa primeira fase procede-se a uma decomposição,

e depois a uma ”recomposição” na fase de previsão:

$$y_1, y_2, ..., y_T \left\{ \begin{array}{c} T_1, ..., T_T \\ S_1, ..., S_T \\ C_1, ..., C_T \\ \varepsilon_1, ..., \varepsilon_T \end{array} \right.$$

- Dado que nenhuma das componentes é directamente observável, é necessário produzir estimativas, isto é,

$$y_1, y_2, ..., y_T \left\{ \begin{array}{ccc} \hat{T}_1, ..., \hat{T}_T & \longmapsto & \hat{T}_{T+h} \\ \hat{S}_1, ..., \hat{S}_T & \longmapsto & \hat{S}_{T+h} \\ \hat{C}_1, ..., \hat{C}_T & \longmapsto & \hat{C}_{T+h} \\ \hat{\varepsilon}_1, ..., \hat{\varepsilon}_T & \longmapsto & F(\varepsilon_{T+h}) \\ \text{Decomposição} & & \text{Previsão} \end{array} \right\} \begin{array}{c} \hat{y}_{T+h} \\ \\ \text{Recomposição} \end{array},$$

  onde $F(.)$ designa a eventual distribuição de ε_t, isto é, assume-se que o erro é uma variável aleatória (métodos não determinísticos).

- Em geral, o erro ε_t deverá ter as propriedades de ”ruido branco” (white noise=processo puramente aleatório e imprevisível): $E(\varepsilon_t) = 0$ e $Cov(\varepsilon_t, \varepsilon_{t+s}) = \left\{ \begin{array}{ll} \sigma_\varepsilon^2, & s = 0 \\ 0, & s \neq 0 \end{array} \right., \forall t.$

- Esta componente é essencial em todo o processo de modelação e previsão. Assim, é através de testes às hipóteses/premissas relativas ao comportamento de ε_t que se avalia a adequação do modelo estimado. Dado que a componente residual não é observada, socorremo-nos das suas estimativas (os resíduos do modelo) para efectuar esses testes. Se estas estimativas se afastarem dessas hipóteses, significa que algo está a escapar ao analista, pois as flutuações não serão erráticas. Será necessário retomar a fase inicial de modelação. Contudo, é usual concentrar a atenção apenas nos indicadores condensados que nos dão uma medida dos erros de previsão.

- Para operacionalizar este este método, distinguem-se três tipos de modelo:

(1) Modelo Aditivo (os efeitos não são interdependentes, adicionando-se simplesmente): $y_t = T_t + S_t + C_t + \varepsilon_t$

(2) Modelo Multiplicativo (efeitos interdependentes, p.e., sasonalidade crescente/decrescente ao longo do tempo): $y_t = T_t.S_t.C_t.\varepsilon_t$

(3) Modelo Misto: $y_t = T_t.S_t + C_t + \varepsilon_t$

- A escolha do tipo de modelo prende-se sobretudo com o padrão da sazonalidade (constante ou variável com o tempo). Tomar as diferenças entre máximo e mínimo para cada ano, fazendo depois uma regressão sobre as médias anuais da variável. Se a correlação for significativa, então um modelo multiplicativo é o mais adequado.

2.1 Tipos de tendência

(1) Modelo sem tendência (tendência constante):

$$T_t = a + \eta_t$$

sendo a constante e η_t um ruído branco. Como se pode perceber, o nível da série é globalmente constante.

(2) Modelo sem tendência, nível estocástico (”random walk” ou passeio aleatório):

$$T_t = T_{t-1} + \eta_t$$

O nível da série é igual ao do período anterior, mais uma perturbação aleatória de média nula. O nível da série oscila (sendo estocástico), mas a série tende a ser localmente constante.

(3) Tendência linear:

$$T_t = \alpha_0 + \alpha_1 t + \eta_t$$

Adequado para séries em que se detecta um crescimento/decrescimento mais ou menos constante durante o período amostral. Modelo com tendência estocástica

$$\begin{aligned} T_t &= T_{t-1} + b_{t-1} + \eta_t \\ b_t &= b_{t-1} + \xi_t \end{aligned}$$

em que T_t é o nível da sucessão no momento t, sendo b_{t-1} o declive da tendência (que muda com t).

(4) Outros:

- Tendência quadrática: $T_t = b_0 + b_1 t + b_2 t^2 + \eta_t$, quando os acréscimos da série variam com t;

- Tendência exponencial: $T_t = ab^t, b > 0$, para séries com taxas de crescimento constantes;

- Tendência logarítmica: $T_t = b_0 + b_1 \log t + \eta_t$, para processos com crescimento a taxas decrescentes.

3 Previsão para séries sem sazonalidade

3.1 Médias Móveis

3.1.1 Séries sem tendência

- OLS a $y_t = a + \eta_t$:

$$\min \phi(a) = \sum_{1}^{T} (y_t - a)^2 \Longrightarrow \hat{a} = \frac{\sum y_t}{T}.$$

- A função de previsão a h passos é dada por $\hat{y}_{T+h} = \hat{a}, \qquad h = 1, 2, ...$

- Isto equivale a, no processo de estimação, atribuir a todas as observações o mesmo peso $(1/T)$, independentemente da sua ”idade” (localização na amostra). Naguns casos, porém, isso pode não fazer sentido, sobretudo quando estamos perante um modelo localmente constante. Assim, é preferível atribuir maior peso às observações mais recentes, conduzindo ao método das Médias Móveis.

Médias Móveis simples

- Para o momento T,

  $$\hat{a}(T) = M_T = \frac{y_T + y_{T-1} + ... + y_{T-N+1}}{N} = \frac{\sum_{t=T-N+1}^{T} y_t}{N},$$

  onde N é o número de observações incluídas em cada média (período da média móvel). Portanto,

  $$M_T = M_{T-1} + \frac{y_T - y_{T-N}}{N}.$$

  Para o período t genérico, temos

  $$M_t = N^{-1}(y_t + y_{t-1} + ... + y_{t-N+1}),$$

  que varia com t

- Objectivos: (1) Utilizar apenas, como informação relevante, os dados mais recentes; (2) Filtrar a série de flutuações aleatórias, "alisando a série" através do operador média, permitindo assim identificar um padrão dos dados, expurgado de flutuações (mais alisada quando maior for N). Porém, N não deve ser tão grande de modo a que se perca informação relevante sobre o padrão da série. A escolha de N é subjectiva, embora um critério que se pode usar é o da minimização de uma medida do erro de previsão (EQM, REQM, etc). Supõe-se que o melhor N para o período amostral será também razoável para previsão (hipótese de estabilidade da série)!

- Regra: N elevado quando o nível da série for razoavelmente constante; N pequeno se o nível da série não for estável (uma vez que M_t demora N períodos a ser consistente com o novo nível).

- A função de previsão a um passo é dada por $\hat{y}_{t+1} = M_t$, enquanto que a h passos teremos

$$\hat{y}_{T+h} = \hat{a}(T) = M_T, \qquad h = 1, 2, ...,$$

ou seja, previsões constantes para todo o h.

Médias Móveis centradas

- Trata-se de uma técnica utilizada sobretudo para filtrar os dados (usar informação futura), não para previsão.

- A fórmula genérica depende de N ser par ou impar:

  - N impar

$$\begin{aligned} CM_t &= \frac{1}{N}[y_{t+\frac{N-1}{2}} + y_{t+\frac{N-1}{2}-1} + ... + y_t + ... + y_{t-\frac{N-1}{2}}] \\ &= (2q+1)^{-1} \sum_{j=-q}^{q} y_{t-j} \end{aligned}$$

- N par: calculam-se duas médias simples em que a observação central muda de posição. Ex: $N = 4$,

$$CM_t = \left\{ \begin{array}{l} M^{(1)} = \frac{y_{t+1}+y_t+y_{t-1}+y_{t-2}}{4} \\ M^{(2)} = \frac{y_{t+2}+y_{t+1}+y_t+y_{t-1}}{4} \end{array} \right\} CM_t = \frac{M^{(1)} + M^{(2)}}{2}.$$

3.1.2 Séries com tendência linear

Médias Móveis Duplas

- Suponha-se que o processo é gerado por

$$y_t = b_0 + b_1 t + \varepsilon_t,$$

sendo a previsão dada por $\hat{y}_{T+1} = M_T$. Facilmente se prova que $E(M_T) = b_0 + b_1 T - \frac{N-1}{2} b_1$, ou seja, o erro de previsão é $E(y_{T+1} - M_T) = b_1 + \frac{N+1}{2} b_1$, o que demonstra o enviesamento de M_T para prever valores futuros de uma tendência linear.

- Uma forma de corrigir este enviesamento é obter estimativas dos parâmetros b_0 e b_1. Um método simples é considerar também Médias Móveis Duplas, ou seja, uma média móvel das últimas N médias móveis:

$$\begin{aligned} M_T^{[2]} &= \frac{M_T + M_{T-1} + ... + M_{T-N+1}}{N} = \frac{\sum_{t=T-N+1}^{T} M_t}{N} \\ M_T^{[2]} &= M_{T-1}^{[2]} + \frac{M_T - M_{T-N}}{N} \end{aligned}$$

- Mostra-se que $E(M_T^{[2]}) = b_0 + b_1 T - (N-1) b_1$. Juntando $E(M_T) = b_0 + b_1 T - \frac{N-1}{2} b_1$ e resolvendo o sistema em ordem a b_0 e b_1 obtém-se

$$\begin{aligned} b_0 &= 2E(M_T) - E(M_T^{[2]}) - b_1 T \\ b_1 &= \frac{2}{N-1} \left[E(M_T) - E(M_T^{[2]}) \right], \end{aligned}$$

o que nos fornece estimadores centrados:

$$\begin{aligned}\hat{b}_0 &= 2M_T - M_T^{[2]} - \hat{b}_1 T \\ \hat{b}_1 &= \frac{2}{N-1}\left[M_T - M_T^{[2]}\right].\end{aligned}$$

- Função de previsão para y_T é dada por

$$\hat{y}_T = \hat{b}_0(T) + \hat{b}_1(T).T = 2M_T - M_T^{[2]}.$$

Para previsão a h passos, teremos

$$\hat{y}_{T+h} = \hat{y}_T + \hat{b}_1 h = 2M_T - M_T^{[2]} + \frac{2h}{N-1}(M_T - M_T^{[2]}).$$

Extrapola-se a partir de $\hat{a}(T)$ uma recta com declive $\hat{b}_1$:

$$\begin{aligned}\hat{y}_{T+h} &= \hat{a}(T) + \hat{b}_1(T)h, \\ \hat{a}(T) &= 2M_T - M_T^{[2]}\end{aligned}$$

- É possível estabelecer uma fórmula recursiva para as estimativas dos $b's$ com base nas fórmulas recursivas de M_T e $M_T^{[2]}$.

- Um dos principais problemas com esta abordagem é o facto de as médias móveis serem inluenciadas por outliers durante N períodos.

3.2 Alisamento Exponencial

- É mais razoável assumir que a informação mais recente será mais útil para obter previsões, já que nos dá uma uma melhor perspectiva do nível actual da série. Nesse sentido, as observações recentes terão um peso maior que as antigas. O método do alisamento exponencial simples permite-nos isso.

3.2.1 Séries sem tendência

- Em $T-1$, temos a estimativa do nível da série $\hat{a}(T-1)$. Em T, obtemos y_T, pelo que podemos rever $\hat{a}(T-1)$. A nova estimativa do nível da série pode ser obtida como resultado de uma média ponderada entre a informação mais antiga, contida em $\hat{a}(T-1)$, e a informação mais recente, y_T, da seguinte forma:

$$\hat{a}(T) = \alpha y_T + (1-\alpha)\hat{a}(T-1), \qquad 0 < \alpha < 1,$$

  o que nos dá uma expressão recursiva para actualizar a estimativa do nível da série (e consequentemente fazer previsão para o próximo período).

- Como $e_T = y_T - \hat{a}(T-1)$,

$$\hat{a}(T) = \hat{a}(T-1) + \alpha[y_T - \hat{a}(T-1)] = \hat{a}(T-1) + \alpha e_T.$$

  Designando $\hat{a}(T) = S_T$,

$$S_T = S_{T-1} + \alpha(y_T - S_{T-1}) = \alpha y_T + (1-\alpha)S_{T-1},$$

  ou seja, o método ajusta, através de α (a constante de alisamento), S_T (a estatística alisada) tendo em conta o erro de previsão mais recente.

- Como $\hat{y}_{T+1} = S_T = \alpha y_T + (1-\alpha)S_{T-1}$, a previsão para o próximo período (estimativa mais recente do nível da série) é uma média ponderada (por α) da observação mais recente (y_T) e da previsão para o período corrente ($S_{T-1} = \hat{y}_T$).

- S_T é uma sucessão de valores alisados/filtrados (i.e., sem componente errática), tratando-se de uma sequência de médias móveis ponderadas (ponderação decresce exponencialmente/geometricamente com antiguidade, com soma dos pesos 1):

$$\begin{aligned} S_T &= \alpha y_T + (1-\alpha)S_{T-1} \\ &= \alpha y_T + (1-\alpha)\alpha y_{T-1} + (1-\alpha)^2 S_{T-2} = ... \\ &= \alpha \sum_{j=0}^{T-1} (1-\alpha)^j y_{T-j} + (1-\alpha)^T S_0 \end{aligned}$$

- S_0 é a estimativa inicial do nível da série. Este valor pode ser obtido como média das primeiras k observações, ou, mais simplesmente, $S_0 = y_1$.

- Prova-se que, para T grande,

$$\begin{aligned} E(S_T) &= a, \\ V(S_T) &= \frac{\alpha}{2-\alpha}\sigma_\varepsilon^2, \end{aligned}$$

pelo que $\hat{a}(T) = S_T$ é um estimador centrado do nível da série.

- Tal como anteriormente, função de previsão a um passo é dada por

$$\hat{y}_{t+1} = S_t,$$

enquanto que a h passos teremos

$$\hat{y}_{T+h} = \hat{a}(T) = S_T, \qquad h = 1, 2, ...$$

- Quanto maior α, menor o alisamento efectuado sobre a série e, em princípio, melhor a capacidade previsional. Porém, existirão problemas se a alteração se deve apenas a movimentos irregulares ou observações extemporâneas, e não a verdadeiras alterações nos parâmetros. Usualmente, α é escolhido com base na minimização de uma função dos erros de previsão.

- Existe uma correspondência entre o Método das Médias Móveis simples e o alisamento exponencial, ou seja, é sempre possível encontrar o alisamento exponencial correspondente ao alisamento por médias móveis de período N. Contudo, o contrário não se verifica.

3.2.2 Séries com tendência: Alisamento exponencial duplo (método de Brown)

- Se a série em estudo tiver um comportamento global ascendente ou descendente, uma possibilidade é considerar um modelo de tendência linear $y_t = b_0 + b_1 t + \varepsilon_t$, b_0, b_1 constantes. Uma hipótese é estimar o modelo por OLS e efectuar previsões do tipo

$$\begin{aligned} \hat{y}_T &= \hat{b}_0 + \hat{b}_1 T, \\ \hat{y}_{T+h} &= \hat{y}_T + \hat{b}_1 h. \end{aligned}$$

- Contudo, quando o comportamento da série oscila este modelo pode revelar-se demasiado rígido, pelo que se podem utilizar técnicas de alisamento exponencial.O valor de $S_T = \hat{a}(T) = \alpha y_T + (1-\alpha)S_{T-1}$ é um estimador enviesado (prevê sistematicamente atrasado) porque

$$E(S_T) = E(y_T) - \frac{1-\alpha}{\alpha} b_1.$$

Portanto, fazer como no caso das Médias Móveis obtendo estimativas de b_1.

- Como

$$E(y_{T+h}) = b_0 + b_1(T+h) = b_0 + b_1 T + b_1 h,$$

voltando a alisar exponencialmente a série, obtemos

$$S_T^{[2]} = \alpha S_T + (1-\alpha) S_{T-1}^{[2]},$$

onde, se pode mostrar que

$$E(S_T^{[2]}) = b_0 + b_1(T) - 2\frac{1-\alpha}{\alpha} b_1.$$

- Conjugando S_T e $S_T^{[2]}$,

$$E(S_T) - E(S_T^{[2]}) = b_1 \frac{1-\alpha}{\alpha},$$

pelo que

$$\hat{b}_1(T) = \frac{\alpha}{1-\alpha}(S_T - S_T^{[2]}).$$

- O nível actual da sucessão $(b_0 + b_1 T)$ pode ser estimado através de $\hat{a}(T) = 2S_T - S_T^{[2]}$ (note-se que $\hat{a}(T) = \hat{y}_T$), pois

$$2E(S_T) - E(S_T^{[2]}) = b_0 + b_1 T.$$

- Uma estimativa do termo independente b_0 é

  $$\hat{b}_0(T) = \hat{y}_T - T\hat{b}_1(T) = 2S_T - S_T^{[2]} - T\frac{\alpha}{1-\alpha}(S_T - S_T^{[2]}).$$

- Assim, a previsão a um passo far-se-á como

  $$\hat{y}_{T+1} = \hat{a}(T) + \hat{b}_1(T),$$

  enquanto que a h passos teremos

$$\hat{y}_{T+h} = \hat{a}(T) + \hat{b}_1(T)h = \hat{b}_0(T) + \hat{b}_1(T)(T+h), \qquad h = 1, 2, ...$$

- Necessário inicializar a sucessão $S_T^{[2]}$ com os valores S_0 e $S_0^{[2]}$. Estes podem ser obtidos a partir de estimativas iniciais de $\hat{b}_0(0)$ e $\hat{b}_1(0)$ utilizando OLS e depois

  $$S_0 = \hat{b}_0 - \hat{b}_1\frac{1-\alpha}{\alpha};\, S_0^{[2]} = \hat{b}_0 - 2\hat{b}_1\frac{1-\alpha}{\alpha}.$$

3.3 Alisamento exponencial: método de Holt

- Viu-se que para uma tendência linear,

  $$\hat{y}_{T+h} = \hat{y}_T + \hat{b}_1(T).h.$$

- Outro método possível para obter essas estimativas é o método de Holt em que utiliza duas fórmulas/equações recursivas (uma para $\hat{y}_T$ ($\hat{a}(T)$), e outra para $\hat{b}_1(T)$) :

$$\hat{y}_T = \alpha y_T + (1-\alpha)[\hat{y}_{T-1} + \hat{b}_1(T-1)],\ 0 < \alpha < 1,$$

ponderação entre a observação real e a última previsão para o nível da série no período T, e

$$\hat{b}_1(T) = \gamma(\hat{y}_T - \hat{y}_{T-1}) + (1-\gamma)\hat{b}_1(T-1),\ 0 < \gamma < 1,$$

ponderando a previsão anterior e uma outra estimativa do declive dada por $\hat{y}_T - \hat{y}_{T-1}$.

- Duas constantes de alisamento, α e γ, o que aumenta as possibilidades de escolha (mas também a arbitrariedade).

- Prova-se que o método de Brown é um caso particular do método de Holt, ou seja, para uma qualquer inicialização e um dado α, existe uma representação em termos de α e γ de Holt, nomeadamente

$$\alpha = \alpha^*(2-\alpha^*) \text{ e } \gamma = \frac{\alpha^*}{2-\alpha^*}. \qquad [\alpha^* = \text{Método de Brown}]$$

- α e γ são escolhidos minimizando uma função dos erros de previsão.

- A inicialização da recursão poderá ser feita como nos métodos anteriores, ou então fixar $\hat{y}_2 = y_2$ e $\hat{b}_1(2) = y_2 - y_1$, ou seja, a recursão tem início em $t = 3$, até $t = T$.

3.4 Alisamento exponencial: tendência amortecida

- Os métodos propostos consideram apenas duas alternativas de extrapolação no nível de tendência (constante e crescimento linear).

- Contudo, a experiência empírica tem vindo a mostrar que estes métodos tendem a sobrestimar os valores reais observados, sobretudo para séries com tendência e para horizontes médios e longos. Isso significa que a opção de crescimento linear não será a mais correcta, sobretudo se a série apresnta um comportamneto curvilíneo acentuado.

- Uma possibilidade é com modelos de tendência polinomial de ordem superior a 1. Um exemplo é o modelo de tendência quadrática
$$y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon_t.$$
Neste caso aplicar-se-iam alisamentos de terceira ordem (com médias móveis ou alisamento exponencial), ou seja, poder-se-iam considerar as estatísticas $M_T^{[3]} = \sum M_t^{[2]}/N$ e $S_T^{[3]} = \alpha S_T^{[2]} + (1-\alpha) S_{T-1}^{[3]}$. A construção de previsões é em tudo semelhante ao exposto anteriormente.

- Alternativa que permite tornar a metodologia do alisamento exponencial mais robusta e flexível (generalização do método de Holt): Alisamento exponencial com tendência amortecida (damped). Lógica: Extrapolação do nível da série de modo a que o efeito da tendência se vai atenuando consoante o horizonte de previsão cresce. Para tal, introduz-se no método de Holt um parâmetro adicional que actua como constante de ajustamento à estimativa do declive da tendência.

- Estimador recursivo do nível da tendência será

$$\hat{y}_T = \alpha y_T + (1-\alpha)[\hat{y}_{T-1} + \phi \hat{b}_1(T-1)].$$

A ponderação ϕ da última estimativa disponível do declive da tendência pode ser diferente de 1 (Holt).

- Estimador do declive da tendência:

$$\hat{b}_1(T) = \gamma(\hat{y}_T - \hat{y}_{T-1}) + (1-\gamma)\phi \hat{b}_1(T-1).$$

- Função de previsão é

$$\hat{y}_{T+h} = \hat{y}_T + \sum_{j=1}^{h} \phi^j \hat{b}_1(T).$$

- A influência da estimativa do declive da tendência varia com o horizonte e previsão. Assim, teremos quatro tipo de extrapolações, em função do valor de ϕ:

- $\phi = 0$: Neste caso, o efeito do declive da tendência é nulo (alisamento exponencial simples);

- $0 < \phi < 1$: O efeito do declive manifesta-se, mas de forma exponencialmente decrescente, ou seja, os acréscimos/decréscimos de $\hat{b}_1(T)$ vão-se atenuando com h;

- $\phi = 1$: Acréscimos de $\hat{b}_1(T)$ constantes (Método de Holt);

- $\phi > 1$: $\hat{b}_1(T)$ é ponderado de forma exponencialmente crescente com h, pelo que a extrapolação é ”explosiva”. Na prática, as séries económicas raramente apresentarão este comportamento, mas pode acontecer (ex: índices da bolsa americana a partir de 1997).

- A escolha de ϕ é feita da mesma forma que para os parâmetros anteriores, ou seja, minimizando uma função dos erros de previsão a um passo.
- A formulação de tendência amortecida é facilmente extensível ao método de Brown:

$$\hat{y}_T = \hat{y}_{T-1} + \phi\hat{b}_1(T-1).$$

4 Intervalos de Previsão

- Até aqui, viram-se alguns métodos que permitem a construção de previsões pontuais para a série em estudo. Contudo, é também conveniente ter uma medida da incerteza associada à previsão, ou seja, dos possíveis erros de previsão. De facto, a probabilidade de o valor previsto coincidir com o valor real é virtualmente nula.
- Para tal, utilizam-se intervalos de confiança para as previsões, ou seja, intervalos de previsão, em que apresentam valores que se crê irem conter o valor futuro da série, com uma determinada probabilidade. Nalgumas circunstâncias, é possível apresentar intervalos de previsão que terão uma estrutura do tipo

Previsão $\pm$ ”nível de confiança”$\times$raíz da variância do erro de previsão.

- Alisamento exponencial simples ($y_t = a + \varepsilon_t$, $\varepsilon_t \sim n.i.d(0, \sigma_\varepsilon^2)$, ruído branco gaussiano):

$$\hat{y}_{T+h} \pm z_{\gamma/2}\sqrt{\frac{2}{2-\alpha}}\hat{\sigma}_\varepsilon, \qquad h = 1, 2, ...$$

  em que z_γ é o valor crítico retirado da Normal com nível de confiança γ e $\hat{\sigma}_\varepsilon$ é uma estimativa do desvio-padrão da componente residual.

- Uma outra forma de construir intervalos de previsão (a um passo) baseia-se na utilização do Erro Absoluto Médio (EAM). Assim, assumindo erros normais, o intervalo de previsão a $(1-\gamma)\,\%$ pode ser construído como

$$\hat{y}_{T+1} \pm z_{\gamma/2} \times 1.25 \times EAM.$$

  Para previsão a h passos, a constante 1.25 altera-se. O intervalo de previsão é facilmente actualizável à medida que chega nova informação, uma vez que

$$M_{t+1} = \frac{tEAM_t + |e_{t+1}|}{t+1}; e_{t+1} = y_{t+1} - S_t.$$

5 Previsão para séries com sazonalidade

- Séries económicas observadas com uma frequência inferior a um ano (mensais ou trimestrais, normalmente). Efeitos sazonais? Tomá-los em conta nas previsões!

- Modelo multiplicativo $y_t = T_t S_t \varepsilon_t$ Versus Modelo aditivo $y_t = T_t + S_t + \varepsilon_t$.
- Necessário estimar os efeitos/factores sazonais S_t (extender o que vimos anterior, incluindo também S_t).

5.1 Dessazonalização com médias móveis

5.1.1 Modelo Multiplicativo

- Considere-se L o n^o de ”estações” no ano, ou períodos sazonais (12 se forem dados mensais, 4 se trimestrais, etc.).
- Modelo multiplicativo: S_t é do tipo $(1 + j)$, com $j > -1$ definindo a taxa de acréscimo ($j > 0$) ou decréscimo ($j < 0$) do nível da série no respectivo período, devido ao efeito sazonal. Por exemplo, $j = 0.12$ significa $S_t = 1.12$, ou seja, um efeito sazonal expansivo sobre o nível da série (12% deste nível). Se $j = -0.12$, $S_t = 0.88$, ou seja, um efeito sazonal depressivo de 12% sobre o nível da série. Deste modo,

$$\sum_{t=1}^{L} S_t = L.$$

- Médias móveis centradas de período L,

$$CMA_t = L^{-1}[y_{t-(L-1)/2} + ... + y_t + ... + y_{t+(L-1)/2}]$$

ou as correspondentes versões para L par

$$M_{t1} = L^{-1}[y_{t-L/2+1} + ... + y_t + ... + y_{t+L/2}],$$
$$M_{t2} = L^{-1}[y_{t-L/2+2} + ... + y_t + ... + y_{t+L/2+1}],$$
$$CMA_t = (M_{t1} + M_{t2})/2.$$

- Estimador do nível da sucessão no momento t ($\hat{y}_t = CMA_t$) uma vez que filtra as componentes sazonal e irregular; Os rácios $z_t = y_t/CMA_t$ permitem obter estimativas das componentes sazonal e irregular; Para eliminar a componente irregular, volta-se a filtrar a série através das médias relativas a cada período sazonal. Trimestrais:

$$S_1^* = (z_1 + z_5 + ... + z_{T-3})$$
$$S_2^* = (z_2 + z_6 + ... + z_{T-2})$$
$$S_3^* = (z_3 + z_7 + ... + z_{T-1})$$
$$S_4^* = (z_4 + z_8 + ... + z_T)$$

- S_j^* são estimativas não normalizadas dos factores sazonais, já filtrados de efeitos residuais. Por vezes, a soma dos S_j^* não é igual a L, pelo que se procede à sua normalização:

$$\hat{S}_j = S_j^* L / \sum S_j^*.$$

- Dessazonalizar a série y_t :

$$y_t^D = y_t^{(j)} / \hat{S}_j.$$

- Previsões: Utilizar a série dessazonalizada y_t^D para efectuar previsões do nível da tendência (método mais adequado -

constante vs. tendência linear, médias móveis vs. alisamento exponencial). Com base nestas previsões $\hat{y}^D_{T+h}$, controem-se as previsões para a série original y_t repondo a sazonalidade, isto é,

$$\hat{y}_{T+h} = \hat{y}^D_{T+h}.\hat{S}^{(j)}_{T+h},$$

onde $\hat{S}^{(j)}_{T+h}$ representa o factor sazonal referente ao período $T+h$.

5.1.2 Modelo Aditivo

- Coeficientes sazonais S_t representam o número de unidades da variável y_t a adicionar ou a subtrair ao nível da série para ter em conta os efeitos depressivos ou expansivos periódicos que afectam a variável. Desse modo, a soma desses efeitos durante o ano é nula, isto é, os efeitos compensam-se:

$$\sum_{t=1}^{L} S_t = 0.$$

- Estimação dos coeficientes sazonais (ver o procedimento anterior) mas $z_t = y_t - CMA_t\ (= T_t + S_t + \varepsilon_t - \hat{T}_t \simeq S_t + \varepsilon_t)$.

- Filtrando a série através das médias relativas a cada período sazonal, eliminamos a componente errática e obtemos estimativas não standardizadas S_j^*. A standardização é feita através de

$$\hat{S}_j = S_j^* - \sum S_j^*/L,$$

pelo que a série dessazonalizada vem

$$y_t^D = y_t^{(j)} - \hat{S}_j.$$

Depois de efectuada a previsão sobre a série dessazonalizada, as previsões para a série original obtém-se através de

$$\hat{y}_{T+h} = \hat{y}_{T+h}^D + \hat{S}_{T+h}^{(j)}.$$

- Criticas: (1) S_j mantém-se constante sobre a amostra e no futuro! (2) Poderão existir problemas se a série se desviar do comportamento constante ou de tendência linear implícitos no método, uma vez que nessas circunstâncias o estimador CMA_t será pouco fiável.

5.2 Método Holt-Winters

- Extensão do método de alisamento exponencial de Holt aplicado à estimação de factores sazonais, apresentando a vantagem de não considerar um padrão sazonal fixo.

5.2.1 Modelo Multiplicativo

- Considere-se o modelo $y_t = (\beta_0 + \beta_1 t)S_t + \varepsilon_t$, ou seja, retirada a sazonalidade, temos um modelo de tendência linear. Teremos, portanto, equações recursivas para estimar o nível actual da série ($\hat{y}_t$), o declive da tendência ($\beta_1(t)$) e a componente sazonal S_t.

- $$\hat{y}_T = \alpha \frac{y_T}{\hat{S}(T-L)} + (1-\alpha)[\hat{y}_{T-1} + \hat{b}_1(T-1)],$$

ponderando a estimativa deste nível efectuada em $T-1$ $(\hat{y}_{T-1}+\hat{b}_1(T-1))$ e a série dessazonalizada.

- (ver método de Holt)

$$\hat{b}_1(T)=\gamma(\hat{y}_T-\hat{y}_{T-1})+(1-\gamma)\hat{b}_1(T-1),\ 0<\gamma<1.$$

-

$$\hat{S}_T=\delta\frac{y_T}{\hat{y}_T}+(1-\delta)\hat{S}_{T-L},\ 0<\delta<1,$$

ponderando a última estimativa para período homólogo $(\hat{S}_{T-L})$ e uma outra estimativa dessa componente, dada por $\frac{y_T}{\hat{y}_T}$.

- A função de previsão virá

$$\hat{y}_{T+h}=[\hat{y}_T+\hat{b}_1(T).h].\hat{S}_{T+h-L},$$

ou seja, trata-se da função de previsão para a componente de tendência devidamente multiplicada pela previsão do factor sazonal.

5.2.2 Modelo Aditivo

- O modelo subjacente é $y_t=\beta_0+\beta_1 t+S_t+\varepsilon_t$.

-

$$\begin{aligned}
&\hat{y}_T=\alpha[y_T-\hat{S}(T-L)]+(1-\alpha)[\hat{y}_{T-1}+\hat{b}_1(T-1)],\\
&\hat{b}_1(T)=\gamma(\hat{y}_T-\hat{y}_{T-1})+(1-\gamma)\hat{b}_1(T-1),\ 0<\gamma<1,\\
&\hat{S}_T=\delta(y_T-\hat{y}_T)+(1-\delta)\hat{S}_{T-L},
\end{aligned}$$

vindo a função de previsão

$$\hat{y}_{T+h} = \hat{y}_T + \hat{b}_1(T).h + \hat{S}_{T+h-L}.$$

- Para além da inicialização de $\beta_0(0)$ e $\beta_1(0)$, é também necessário inicializar S_0. Esta pode ser feita através de $S_j = y_j - (y_1 - \hat{b}_{1,L+1}(j-1))$, $j = 1, ..., L$.

5.3 Tendência Amortecida

- Viu-se anteriormente uma forma flexível e robusta de prever séries com tendência através da generalização do método de Holt. Esta metodologia pode igualmente ser aplicada a séries com sazonalidade, quer multiplicativa, quer aditiva.

- Modelo Multiplicativo

$$\begin{aligned}
\hat{y}_T &= \alpha \frac{y_T}{\hat{S}(T-L)} + (1-\alpha)[\hat{y}_{T-1} + \phi \hat{b}_1(T-1)], \\
\hat{b}_1(T) &= \gamma(\hat{y}_T - \hat{y}_{T-1}) + (1-\gamma)\phi \hat{b}_1(T-1), \\
\hat{S}_T &= \delta \frac{y_T}{\hat{y}_T} + (1-\delta)\hat{S}_{T-L} \\
\hat{y}_{T+h} &= [\hat{y}_T + \sum_{j=1}^{h} \phi^j \hat{b}_1(T)].\hat{S}_{T+h-L}.
\end{aligned}$$

- Modelo Aditivo

$$
\begin{aligned}
&\hat{y}_T = \alpha(y_T - \hat{S}_{T-L}) + (1-\alpha)[\hat{y}_{T-1} + \phi\hat{b}_1(T-1)],\\
&\hat{b}_1(T) = \gamma(\hat{y}_T - \hat{y}_{T-1}) + (1-\gamma)\phi\hat{b}_1(T-1),\\
&\hat{S}_T = \delta(y_T - \hat{y}_T) + (1-\delta)\hat{S}_{T-L}\\
&\hat{y}_{T+h} = \hat{y}_T + \sum_{j=1}^{h} \phi^j \hat{b}_1(T) + \hat{S}_{T+h-L}.
\end{aligned}
$$