KTH Matematik

Motiv för exponentiell utjämning

Detta är menat att komplementera presentationen av exponentiell utjämning som prognosmetod i Axsäter “Lagerstyrning”.
Vi modellerar efterfrågan i tidsperiod t som en stokastisk variabel $X_t = a + \epsilon_t$, där a är en (okänd) konstant och ϵ_t är en stokastisk variabel sådan att $E(\epsilon_t) = 0$ och sådan att alla $\epsilon_t, t = 1, 2, \ldots$ är oberoende.
Baserat på informationen upp till och med tidsperiod t, dvs $\{X_s\}_{s=1}^t$, låter vi prognosen för tidsperiod $t + \tau$ vara $\hat{X}_{t,t+\tau} = \hat{a}_t$, där $\hat{a}_t$ beräknas ur rekursionen

$$\hat{a}_t = (1-\alpha)\hat{a}_{t-1} + \alpha X_t, \quad \alpha \in (0,1) \tag{1}$$

med något begynnelsevärde $\hat{a}_0$.
Motivet till detta förfaringssätt kan härledas till följande minimeringsproblem:

$$\min_{\hat{a}_t} \sum_{s=0}^{t-1} \beta^s (X_{t-s} - \hat{a}_t)^2,$$

där $\beta \in (0,1)$ är konstant och $\hat{a}_t$ betecknar estimatet av den okända storheten a. Vi vill alltså minimera en viktad summa av prognosfelen i kvadrat. Viktningen är sådan att nyare data får högre vikt än gamla data.
Första ordningens nödvändiga villkor för optimalitet ger då

$$\sum_{s=0}^{t-1} \beta^s (X_{t-s} - \hat{a}_t) = 0,$$

eller

$$\sum_{s=0}^{t-1} \beta^s X_{t-s} = \hat{a}_t \sum_{s=0}^{t-1} \beta^s = \hat{a}_t \frac{1-\beta^t}{1-\beta},$$

dvs

$$\hat{a}_t = \frac{1-\beta}{1-\beta^t} \sum_{s=0}^{t-1} \beta^s X_{t-s}.$$

Om vi gör motsvarande analys för data till och med $t-1$ så får vi

$$\hat{a}_{t-1} = \frac{1-\beta}{\beta(1-\beta^{t-1})} \sum_{s=1}^{t-1} \beta^s X_{t-s}.$$

Vi kan utveckla $\hat{a}_t$ till

$$\hat{a}_t = \frac{1-\beta}{1-\beta^t} \sum_{s=1}^{t-1} \beta^s X_{t-s} + \frac{1-\beta}{1-\beta^t} X_t = \frac{1-\beta^{t-1}}{1-\beta^t} \beta \hat{a}_{t-1} + \frac{1-\beta}{1-\beta^t} X_t$$

För stora t, gäller att $1-\beta^t \approx 1$ och $(1-\beta^{t-1})/(1-\beta^t) \approx 1$ vilket ger rekursionen

$$\hat{a}_t = \beta \hat{a}_{t-1} + (1-\beta) X_t.$$

Om vi nu ersätter β med $1-\alpha$ fås rekursionen (1).
Vi skall nu visa att asymptotiskt kommer $\hat{a}_t$ vara en väntevärdesriktig skattning av den okända storheten a. Dvs oavsett begynnelsevärde $\hat{a}_0$ så kommer $E(\hat{a}_t) \to a$.
Vi får om vi utvecklar $\hat{a}_t$

$$\begin{aligned} \hat{a}_t &= \alpha X_t + (1-\alpha)[\alpha X_{t-1} + (1-\alpha)\hat{a}_{t-2}] \\ &= \alpha X_t + (1-\alpha)\alpha X_{t-1} + (1-\alpha)^2 \alpha X_{t-2} + (1-\alpha)^3 \hat{a}_{t-3} \end{aligned}$$

och alltså

$$\hat{a}_t = \alpha \sum_{s=0}^{t-1} \beta^s X_{t-s} + \beta^t \hat{a}_0, \tag{2}$$

där $\beta = 1-\alpha$. Vi får

$$E(\hat{a}_t) = \alpha \sum_{s=0}^{t-1} \beta^s E(X_{t-s}) + \beta^t \hat{a}_0 = \alpha \sum_{s=0}^{t-1} \beta^s a + \beta^t \hat{a}_0 = \alpha \frac{1-\beta^t}{1-\beta} a + \beta^t \hat{a}_0 = (1-\beta^t)a + \beta^t \hat{a}_0$$

och det önskade resultatet genom att låta $t \to \infty$.

Prognosfel över ledtiden

Låt efterfrågan X_t beskrivas av en stokastisk process i diskret tid. Antag att vi vid tiden t gör prognosen $\hat{X}_{t,t+\tau}$ av $X_{t+\tau}$, för $\tau = 1, 2$, osv, baserat på informationen $X_s, s = 1, \ldots, t$ och låt $\hat{X}_t^L = \sum_{\tau=1}^{L} \hat{X}_{t,t+\tau}$ vara prognosen för efterfrågan under ledtiden.
Antag att vi mäter prognosfelen över en period,

$$Z_t = \hat{X}_{t-1,t} - X_t.$$

Om nu $\hat{X}_{t-1,t}$ är en väntevärdesriktig skattning av X_t så gäller att $E(Z_t) = 0$. Vi är nu intresserade av att skatta $V(Z_t)$. Ett sätt att skatta $V(Z_t)$ är att beräkna det genomsnittliga felet i kvadrat (mean squared error),

$$\text{MSE} = \frac{1}{t} \sum_{s=1}^{t} Z_s^2,$$

men man kan också använda exponentiell utjämning, på följande sätt:

$$\text{MSE}_t = (1-\gamma)\text{MSE}_{t-1} + \gamma Z_t^2, \quad \gamma \in (0,1),$$

med någon lämplig begynnelse-gissning MSE_0. Jämför med bokens förslag att skatta det genomsnittliga absoluta felet (mean absolute deviation), MAD, med hjälp av exponentiell utjämning (se (2.25) i boken) och sedan skatta $V(Z_t)$ med $\pi \text{MAD}_t^2/2$.
Låt nu Z_t^L beteckna prognosfelet under ledtiden,

$$Z_t^L = \hat{X}_t^L - \sum_{\tau=1}^{L} X_{t+\tau} = \sum_{\tau=1}^{L} (\hat{X}_{t,t+\tau} - X_{t+\tau}).$$

Om vi har väntevärdesriktiga prognoser gäller förstås att $E(Z_t^L) = 0$.
Vi skall nu göra en analys av variansen på prognosfelet under ledtiden. Antag att $X_t = a + \epsilon_t$ där ϵ_t är oberoende stokastiska variabler med väntevärde 0 och varians σ_ϵ^2.

$X_{t+\tau}$ är oberoende av $\hat{X}_{t,t+\tau}$, ty $\hat{X}_{t,t+\tau}$ beror bara på vad som hänt till och med tidsperiod t. Därför gäller att

$$V(Z_t^L) = V(\hat{X}_t^L) + V(\sum_{\tau=1}^{L} X_{t+\tau}) = V(\hat{X}_t^L) + L\sigma_\epsilon^2. \tag{3}$$

För att analysera $V(\hat{X}_t^L)$ måste vi också veta hur själva prognosen utförs. Antag att prognosen tas fram med hjälp av enkel exponentiell utjämning, dvs att

$$\hat{X}_{t,t+\tau} = \hat{a}_t = \alpha X_t + (1-\alpha)\hat{a}_{t-1}.$$

Vi får ur (2) att för stora t

$$V(\hat{a}_t) \approx \alpha^2(1+\beta^2+\beta^4+\beta^6+\cdots)\sigma_\epsilon^2 = \frac{\alpha^2}{1-\beta^2}\sigma_\epsilon^2 = \frac{\alpha}{2-\alpha}\sigma_\epsilon^2.$$

Vilket då ger

$$V(\hat{X}_t^L) = V(L\hat{a}_t) \approx \frac{\alpha}{2-\alpha}L^2\sigma_\epsilon^2.$$

Om vi kombinerar detta med (3) får vi

$$V(Z_t^L) \approx \left(\frac{\alpha}{2-\alpha}L^2 + L\right)\sigma_\epsilon^2.$$

Variansen på en-periodersfelet är alltså

$$V(Z_t^1) = V(Z_t) \approx \left(\frac{\alpha}{2-\alpha} + 1\right)\sigma_\epsilon^2,$$

som vi skattar med MSE_t enligt ovan. Det verkar då rimligt att skatta $V(Z_t^L)$ med

$$\frac{\left(\frac{\alpha}{2-\alpha}L^2 + L\right)}{\left(\frac{\alpha}{2-\alpha} + 1\right)}\mathrm{MSE}_t.$$

Jämför detta med bokens enklare men grövre skattning (2.26). Liknande skattningar kan göras för andra prognosmetoder, t.ex. för exponentiell utjämning med trend, men de algebraiska räkningarna tenderar att bli ganska långa och komplicerade.