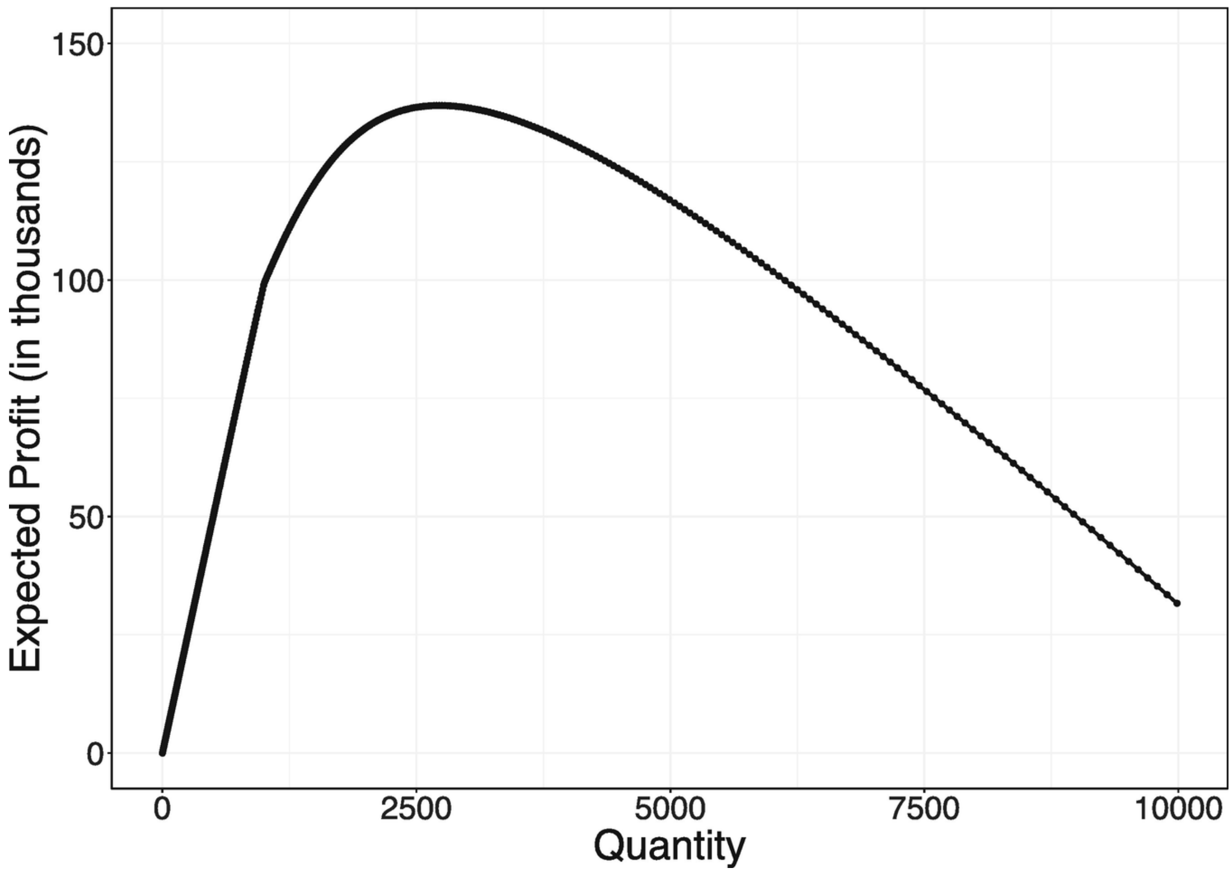

Fig. 4.4 Expected profit as a function of order quantity

4.4 Demand Regularization

Demand regularization is another uncertainty modelling approach that combines statistical modelling with the dynamics of data generation to characterize demand uncertainty correctly. It would be highly effective in practice when the demand lead time varies for different orders. In such cases, companies collect advance orders from some customers. However, some customers may still place urgent orders, for which the demand lead time would be very short. There are two benefits of advance demand information. First, it correlates with the actual demand, so it can be used as an explanatory variable to predict demand. Second, the actual demand can never be less than the advance demand because customers have already committed to purchasing the advance demand. Therefore, it should be used as a lower bound for the demand estimates.

In Fig. 4.5, we present an example of a manufacturer that collects advance orders from customers (i.e. some retailers). The figure shows the actual and advance demand values for the last 25 months. The solid curve represents the actual monthly demand values, while the dashed line shows the advance demand values. The manufacturer plans the production on a monthly basis. Customers are restricted to placing their purchase orders by the last day of the month before their requested delivery date. If a customer wants the ordered items delivered on July 20, for example, the purchase order should be placed on June 30 at the latest. The manufacturer also induces customers to place their orders well in advance by informing them that orders will be replenished on a first-come, first-served basis if there is a product shortage. The manufacturer plans the production schedule at the beginning of June with an intention of meeting the demand for July. The sum of the orders placed in May and earlier months to be delivered in July constitutes the advance demand. The sum of the orders placed in June to be delivered in July constitutes the urgent demand. The sum of the advance and urgent demand gives the actual demand for July. When the production is planned (i.e. the beginning of June), the manufacturer knows the advance demand. She can use this information as an explanatory variable and a lower bound constraint to better predict demand and characterize its uncertainty. The figure shows that there is a high correlation between the actual and advance demand values. Demand values in months 11, 13, 18, 19 and 22 are also equal to the advance demand values.

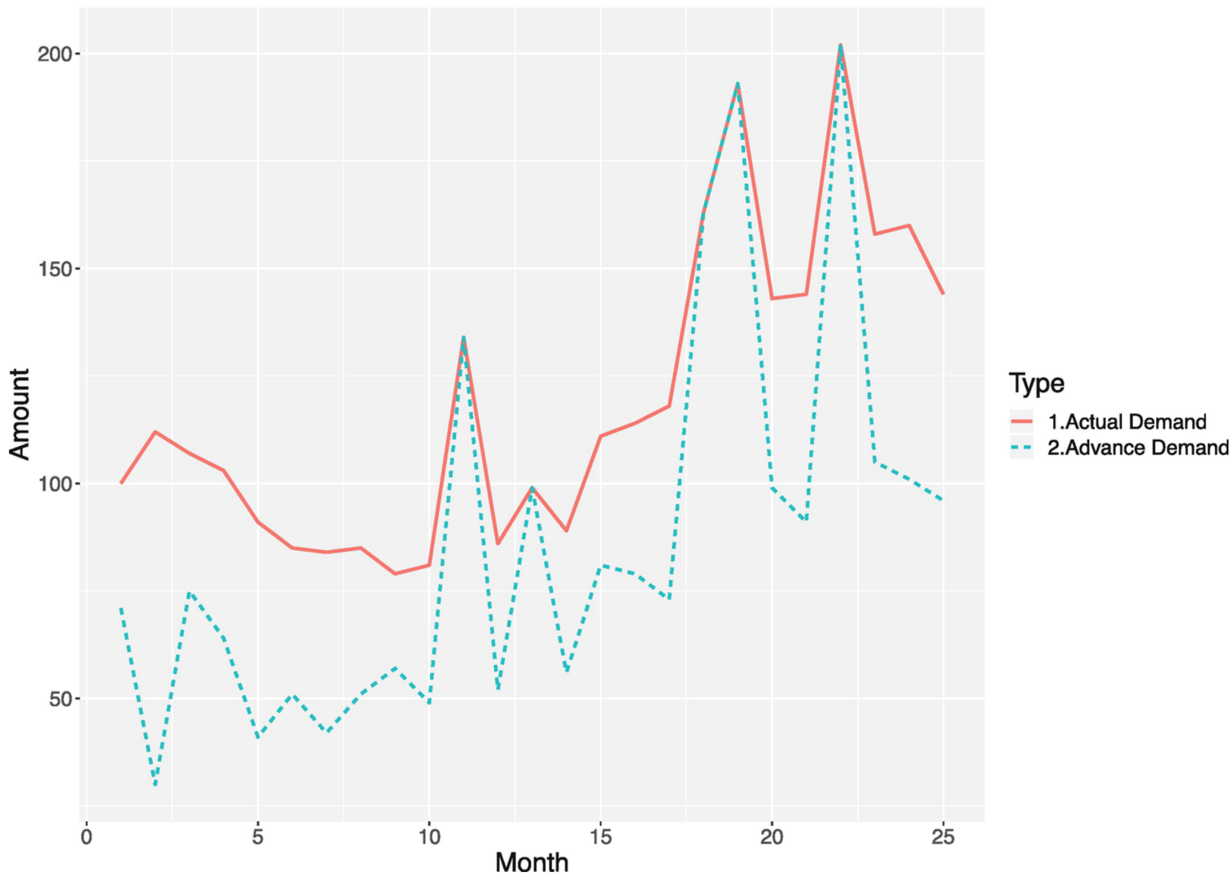

Fig. 4.5 Historical values of actual monthly demand and advance demand when the order quantity is determined

If demand planners ignore the advance demand information and regress demand on other exploratory variables, they would use a linear regression model. Suppose that a dataset contains *n* demand values and there are *k* explanatory variables. We consider a regression model to forecast demand. For notational simplicity, we present the model in the matrix form:

$$Y = XB + E,$$

with:

$$Y = \begin{vmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{vmatrix}, \qquad \mathbf{X} = \begin{vmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \ddots & & \vdots & \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{vmatrix}, \qquad B = \begin{vmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{vmatrix}, \qquad E = \begin{vmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{vmatrix},$$

where Y and X are the vectors of demand values and the matrix of explanatory variables, respectively. For example, if data is collected for n months, y_1 in the Y vector indicates the demand value for the first month. Likewise, $x_{11}, x_{12}, \cdots, x_{1k}$ indicate the values of the exploratory variables for the first month. Those exploratory variables can be economic factors (e.g. economic indices) or intrinsic factors (e.g. online traffic of the company website). The terms B and E denote the vectors of coefficients and error terms, respectively. The first column of the X vector includes only ones, so β_0 gives the intercept of the linear model.

It follows from Chap. 2 that the OLS estimation of B is given by:

$$B_{OLS} = (X^T X)^{-1} X^T Y.$$

Advance demand information determines the lower bound for the actual demand so it should be added to the forecasting model as a constraint. Suppose that the amount of advance orders is equal to L_i for the ith observation. Then, the advance demand vector is defined such that each entry indicates the amount of advance orders for each observation:

$$L = \begin{vmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{vmatrix}.$$

Then, the least squares problem is written as follows:

$$\begin{aligned} &\text{Minimize:} \qquad && (Y - XB)^T (Y - XB), \\ &\text{such that:} \qquad && XB \geq L. \end{aligned}$$

The problem can be solved using Lagrangian optimization, which is reviewed in Chap. 2. The Lagrangian model is written as follows:

$$J(B, \lambda) = (Y - XB)^T (Y - XB) + \lambda^T (L - XB),$$

where λ is the Lagrange multiplier vector. This multiplier should be considered as a penalty cost when the advance demand constraint in the least squares problem is violated. A positive value of an entry in the λ vector indicates that the constraint $XB \geq L$ for that entry is binding.

Therefore, the predicted value of OLS for that entry is less than the advance demand so the demand prediction should be replaced by the advance demand.

In this expression, we have two sets of variables, that is, B and λ, which we would like to estimate. We take the first derivative of the Lagrangian model with respect to these two variables to find the values that minimize $J(B, \lambda)$:

$$\begin{aligned}
\frac{\partial J(B,\lambda)}{\partial B} &= -2X^TY + 2X^TXB - X^T\lambda = 0, \\
B_{REG} &= (X^TX)^{-1}(X^TY + \frac{1}{2}X^T\lambda), \\
\frac{\partial J(B,\lambda)}{\partial \lambda} &= L - XB = 0.
\end{aligned}$$

Combining the last two expressions,

$$\lambda = 2(X(X^TX)^{-1}X^T)^{-1}L - 2Y$$

The formulas for B_{REG} and λ are used together with the inputs X, Y and L to calculate the coefficients.

The λ vector gives us useful insights regarding the value of advance demand in predicting the actual demand. Demand predictions from the regression model are found by XB for each value of Y. Those with a positive λ value should be replaced by the advance demand, which occurs when the advance demand is unexpectedly high. Given that the process of replacing some demand predictions with advance demand is part of the demand regularization approach, the estimates ofB_{REG} are expected to be different from the OLS estimates B_{OLS}.

We now turn back to our example given in Fig. 4.5. The values of the actual and advance demand are given in Table 4.3. We regress the demand values on the previous months' demand and the current months' advance demand values. For this reason, there should be three columns in the X matrix. The first column is the column of ones. The second column is the column of the previous months' demand values, which includes the demand values from the 1st month until the 24th month. The last column shows the advance demand values including the values from the 2nd month to the 25th month. The Y vector includes the demand values from the 2nd month to the 25th month. Likewise, the L vector includes the advance demand values from the

2nd month until the 25th month. Therefore, based on the data given in Table 4.3, these parameters look like the following:

$$Y = \begin{vmatrix} 112 \\ 107 \\ \vdots \\ 144 \end{vmatrix}, \qquad X = \begin{vmatrix} 1 & 100 & 30 \\ 1 & 112 & 75 \\ \vdots & \vdots & \vdots \\ 1 & 160 & 96 \end{vmatrix}, \qquad L = \begin{vmatrix} 30 \\ 75 \\ \vdots \\ 96 \end{vmatrix}.$$

The reason we start from the second month in the Y vector is that the first month's value is used as an explanatory variable. We then construct the X matrix and the L vector accordingly. We present the results of this example in the online web application, which also includes the Python codes. The mean absolute percentage error (MAPE) is 5.96% for the demand regularization model, whereas it is 6.67% for the OLS model. Therefore, the demand regularization model performs better than the OLS model.

Table 4.3 Values of the actual and advance demand

Month	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Demand	100	112	107	103	91	85	84	85	79	81
Advance demand	71	30	75	64	41	51	42	51	57	49
Month	11	12	13	14	15	16	17	18	19	20
Demand	134	86	99	89	111	114	118	163	193	143
Advance demand	134	52	99	56	81	79	73	163	193	99
Month	21	22	23	24	25					
Demand	144	202	158	160	144					
Advance demand	91	202	105	101	96					