IEOR 151 – Lecture 6
Data-Driven Newsvendor

1 Nonparametric Newsvendor

Recall that in the setting of the nonparametric newsvendor problem, we do not know what our distribution on the demand is. Instead, we have iid samples of our demand $X_1, \ldots, X_n$, which is drawn from some distribution with unknown cdf $F(u)$. A nonparametric formulation of the newsvendor problem is to compute δ^* to solve the following optimization problem

$$\min_{\delta} \frac{1}{n} \sum_{i=1}^{n} (c_f + c_v \cdot \delta - p \cdot (\delta - X_i)^- + q \cdot (\delta - X_i)^+).$$

The intuition is that this is the sample average of the loss. Because the objective function of the nonparametric newsvendor problem is the expectation of the newsvendor loss under the empirical distribution

$$\hat{F}(z) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(z \leq X_i),$$

the solution is immediate: We should choose an inventory level δ^* such that

$$\hat{F}(\delta^*) = \frac{p - c_v}{p + q}.$$

If we sort the X_i into increasing order $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$, then the solution corresponds to the $\lceil n \cdot (\frac{p - c_v}{p + q}) \rceil$-th value. In other words, the solution is the value

$$X_{(\lceil n \cdot (\frac{p - c_v}{p + q}) \rceil)}.$$

2 Linear Parametric Newsvendor

Suppose we have samples of demand $X_1, \ldots, X_n \in \mathbb{R}$ along with samples of input variables $Z_1, \ldots, Z_n \in \mathbb{R}^p$ that can predict demand. For instance, suppose we are looking at the demand for ice cream. Then the input variables Z_i might be a vector with a component for the outside temperature and another component for the amount of rainfall in the day. Given this setting, we could potentially use a parametric model that takes the input variables and

predicts the corresponding demand. For this model, we would like to choose the parameters of the model to made demand predictions to minimize the expected newsvendor loss. Mathematically, we can pose this problem as the following optimization problem

$$\min_{\beta} \ \frac{1}{n} \sum_{i=1}^{n} (c_f + c_v \cdot Z_i'\beta - p \cdot (Z_i'\beta - X_i)^- + q \cdot (Z_i'\beta - X_i)^+),$$

where β is a vector of parameters, Z_i' denotes the transpose of Z_i, and the decision rule for demand is $\delta = Z'\beta$. In particular, for this decision rule the inventory level depends upon the most recent data we have available.

Computing the β vector is tractable. This problem can be reformulated as a linear program (LP):

$$\min \ \frac{1}{n} \sum_{i=1}^{n} (c_f + c_v \cdot Z_i'\beta + p \cdot s_i + q \cdot t_i)$$
$$\text{s.t. } s_i \geq -(Z_i'\beta - X_i)$$
$$t_i \geq (Z_i'\beta - X_i)$$
$$s_i, t_i \geq 0.$$

2.1 REGULARIZATION

Just as is done in regression modeling, we can regularize this problem using L1-regularization, L2-regularization, or both. The statistical advantages of L1- and L2-regularization translate to this setting. Recall that L1-regularization helps when we expect most of the input variables do not affect demand, and L2-regularization helps when there is a significant amount of correlation between the input variables.

The L1-regularized version of the problem is

$$\min_{\beta} \ \frac{1}{n} \sum_{i=1}^{n} (c_f + c_v \cdot Z_i'\beta - p \cdot (Z_i'\beta - X_i)^- + q \cdot (Z_i'\beta - X_i)^+) + \lambda \|\beta\|_1,$$

where $\lambda > 0$ is a tuning parameter (which can be chosen using cross-validation) and $\|\beta\|_1 = \sum_{k=1}^{p} |\beta^k|$ is the ℓ_1 norm of a p-dimensional vector. This problem can be formulated as an LP, but requires additional exposition.

3 LP Lift of the L1 Ball

Suppose $x \in \mathbb{R}^2$, and consider the constraint $\|x\|_1 = |x_1| + |x_2| \leq 1$. Note this constraint is indicated by the shaded green square on the left side of the image below. One way to represent this constraint set is as the intersection of four half-spaces. For instance, the half-space $-x_1 + x_2 - 1 \leq 0$ is shown on the right side of the image below.
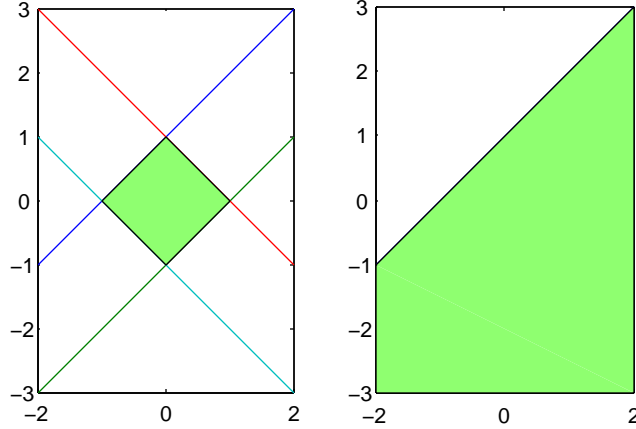
Hence, the constraint $|x_1| + |x_2| - 1 \leq 0$ can be written as the intersection of four half-spaces:

$$x_1 + x_2 - 1 \leq 0$$
$$x_1 - x_2 - 1 \leq 0$$
$$-x_1 + x_2 - 1 \leq 0$$
$$-x_1 - x_2 - 1 \leq 0.$$

Another representation of the constraint $|x_1| + |x_2| - 1 \leq 0$ is as the projection of a set. For example, we can write $|x_1| \leq 1$ as the projection of a polytope by introducing an additional variable t_1:

$$-t_1 \leq x_1 \leq t_1$$
$$t_1 = 1$$

Thus, the constraint $|x_1| + |x_2| - 1 \leq 0$ can be written as the projection of a polytope by introducing two additional variables t_1, t_2. We can write this constraint as

$$-t_1 \leq x_1 \leq t_1$$
$$-t_2 \leq x_2 \leq t_2$$
$$t_1 + t_2 = 1.$$

Now consider the constraint $\sum_{k=1}^{p} |x_k| - 1 \leq 0$. We can write this constraint as the intersection of half-spaces:

$$\sum_{k=1}^{p} \pm x_k - 1 \leq 0,$$

where the $\pm$ indicates that every possible permutation of signs is included in this set of constraints. There are p variables $x_1, \ldots, x_p$ and 2^p constraints/inequalities. However, we can also write this constraint as the projection of a polytope:

$$-t_k \leq x_k \leq t_k, \forall k = 1, \ldots, p$$
$$\sum_{k=1}^{p} t_k = 1.$$

3

There are $2p$ variables $x_1, \ldots, x_p, t_1, \ldots, t_p$ and $2p+1$ constraints/inequalities. The constraint $\sum_{k=1}^{p} |x_k| - 1 \leq 0$ requires significantly less inequalities (as compared to repesenting the set by the intersection of half-spaces) when represented as the projection of a polytope.

4 LP Formulation of L1-Regularized Linear Parametric Newsvendor

We are now ready to reformulate this problem as an LP. In particular, the reformulation is

$$\min \frac{1}{n} \sum_{i=1}^{n} (c_f + c_v \cdot Z_i'\beta + p \cdot s_i + q \cdot t_i) + \lambda \cdot U$$
$$\text{s.t. } s_i \geq -(Z_i'\beta - X_i)$$
$$t_i \geq (Z_i'\beta - X_i)$$
$$s_i, t_i \geq 0$$
$$-v_k \leq \beta^k \leq v_k, \forall k = 1, \ldots, p$$
$$\sum_{k=1}^{p} v_k = U.$$