卡尔曼滤波简介

一种利用线性系统状态方程，通过系统输入输出观测数据对系统状态进行最优估计的算法，由于观测数据中包括系统中的噪声和干扰的影响，所以最优估计也可以看作是滤波过程。

卡尔曼滤波由一系列递归数学公式描述。它们提供了一种高效可计算的方法来估计过程的状态，并使估计均方误差最小。卡尔曼滤波器应用广泛且功能强大，可以估计信号过去和当前的状态，甚至能估计将来的状态，即使不知道模型的确切性质。

其过程可由以下两个方程描述：

$$x_k = Ax_{k-1} + Bu_{k-1} + \omega_{k-1}$$

$$z_k = Hx_k + v_k$$

其中ω_k和v_k是过程噪声和观测噪声，可假设为相互独立的正态分布的白色噪声。A 为状态转移矩阵，H 为观测矩阵。

$$p(\omega) \sim N(0, Q)$$

$$p(v) \sim N(0, R)$$

定义先验估计与先验估计误差$\hat{x}_k^-$与$\hat{e}_k^-$，令$\hat{e}_k^- = x_k - \hat{x}_k^-$。定义$\hat{x}_k$为已知$z_k$的后验状态估计，以及后验误差$\hat{e}_k = x_k - \hat{x}_k$。

所以可以得到先验估计的协方差$P_k^- = E[\hat{e}_k^- \hat{e}_k^{-T}]$和后验误差的协方差$P_k = E[\hat{e}_k \hat{e}_k^T]$。故可以算出 Kalman 增益 K，使得$\hat{x}_k = \hat{x}_k^- - K(z_k - H\hat{x}_k^-)$。

K 的目的是使后验估计误差的协方差矩阵最小，带入后对 K 求导，使导数为 0，可以的出

$$K_k = \frac{P_k^- H^T}{HP_k^- H^T + R}$$

所以 Kalman 滤波由两部分组成，即时间更新方程：

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1}$$

$$P_k^- = AP_{k-1}A^T + Q$$

状态更新方程：

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1}$$

$$\hat{x}_k = \hat{x}_k^- - K(z_k - H\hat{x}_k^-)$$

$$P_k = (I - K_k H)P_k^-$$

实际应用：

在本例中，添加了温度因素与前一天同一时刻的电负荷因素，与当天的的电负荷做多元线性回归，形成观测矩阵 H，并且令状态转移矩阵 A 为单位阵。

$$x_k = (v_1, T_k, L_{k-1})^T$$

$$H_k = (1, H_1, H_2)$$

$$\mathrm{A} = I_3$$

$$z_k = L_k$$

其中v_1为一基础值，有回归分析得出，T_k为当天的温度，L_{k-1}为前一天同时刻的电负荷，H_1与H_2为基于温度与前一天额同时刻电负荷通过回归分析得出的系数。

观测值z_k为当天的电负荷观测值，噪声与观测误差的方差 Q、R 分别定为 0.01 与 0.0001。预测方程为$y_k = H_k x_k$

所以实际应用 Kalman 滤波由两部分组成，即时间更新方程：

$$x_k = A\hat{x}_{k-1}$$

$$P_k^- = AP_{k-1}A^T + Q$$

状态更新方程：

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1}$$

$$\hat{x}_k = x_k - K(z_k - Hx_k)$$

$$P_k = (I - HK_k)P_k^-$$

而每一个预测值由预测方程$y_k = H_k x_k$得出。

运行结果：

由于中国节假日以及周末调休情况比较复杂，故假日的预测误差较大，而工作日具有较为稳定，每天趋势基本相同的特点，故工作日的误差较小。

利用周袁与卢川所提供的数据 dataDAILY.csv 中的温度与每日最大负荷对 2011 至 2015 年，五年内的最大负荷进行预测，工作日的绝对平均误差为 2.76%，总绝对平均误差为 4.03%，利用 dataDAILY.csv 中的温度与傅长青处理过的 load_before_2013_clean.csv 中每天 48 点电负荷(注：其中有个别时间数据缺失，已补全)，进行 2011-2013，两年内 48 点电负荷预测，工作日的绝对平均误差为 3.21%，总的绝对平均误差为 4.29%。