

Chapter 1

The Base Case

This section describes the simplest case in which there is a single fixed tree based on the classical CART algorithm of Breiman et. al. [**regtree**] and a first-order polynomial model. The case to be argued here is that combining two well known models into a single framework has benefits to the analyst, even in its most simple formulation. This section also serves as a basis for the construction of more complicated models.

The CART regression tree models the following relationship:

$$Y_t = f(X_t, v_t)$$

where X_t are explanatory covariates, and v_t are the errors in estimating Y_t and f is the tree mapping between the observations and explanatory variables.

A fixed tree model implies:

1. K_T leaves and $K_T - 1$ internal nodes,
2. which implies a fixed number number of known covariates with fixed threshold values,
3. and a fixed number of Kalman filters (KFs).
4. If it is additionally assumed that $\lambda_a = \lambda_s$ (that the rate of arrival of inputs is that same as the rate of computation) then *a priori* it can be assumed that the leaves have probability $\frac{1}{K_T}$ of updating at each iteration. We also assume discrete time iterations.

The model at each leaf k of the tree is:

$$y_t = z_{t,k} + v_{t,k}; \quad v_{t,k} \sim N(0, V_t) \tag{1.1a}$$

$$z_{t,k} = z_{t-1,k} + w_{t,k}; \quad w_{t,k} \sim N(0, W_t) \tag{1.1b}$$

with initial conditions, $z_{0,k} \sim N(\mu_{0,k}, \Sigma_{0,k})$. These parameters, $\mu_{0,k}$, $\Sigma_{0,k}$, are set to the sample mean and variance of the leaves to which the filters correspond but could be set to any appropriate value. Some final assumptions in this simplest of models is that $y_{t,k}$, $z_{t,k}$, V_t and W_t are one-dimensional and that both V_t and W_t are constant over t , known and the same for each leaf, and that the $E[v_{t,k}, w_{t,k}] = 0$, i.e. no autocorrelation.

The combination of these models then explains:

$$Y_t = f(X_t, \Theta, \epsilon_t)$$

where now Θ is a vector of parameters that include $z_{1:t}, y_{1:t-1}, V_t, W_t, \Upsilon_t$ and ϵ_t is the overall error of the model which is some combination of the errors of regression tree, the measurement equation and the state equation. These three different errors can be explained by considering that a process $y_{t,k}$ can be observed (and measured) with some error $v_{t,k}$, one can attempt to explain this process via a regression tree f of the possible states of the process at each time t with some error v but even then, there may be some latent process $z_{t,k}$ that also explains $y_{t,k}$ but because it is both latent and evolving in time, it cannot be included as a covariate in the tree model. This latent process or signal has some uncertainty associated with it, $w_{t,k}$, in part caused by the simplifying Markov assumption.

The link between the two models is that they both model $E[Y_t | X_t, \Theta] = \bar{y}_k = z_{t,k}$ where \bar{y}_k is the sample average at each leaf, which can also be considered as an unobserved parameter, and $z_{t,k}$ is the mean estimate of the process $y_{t,k}$. In brief, a tree is being used at each t to model various possible states of the process $y_{t,k}$.

1.0.1 Updating Equations

At each t a node is first selected and then inference on the leaf is performed. Every leaf is processed according to the intermittent Kalman filter [**sinopoli**] and as such only the selected leaf will perform the forecast and update step (adapted from [**west**]) while the other leaves only propagate the prior under the assumption that the next value of the state in that partition is the same as the previous state but with a linear increase in uncertainty. $\eta_{x,T}$ is an indicator variable over the partitions. If $\eta_{x,T} = k$ then y_t is observed at the leaf selected by covariates x_t from tree T at time t . Else $\eta_{x,T} \neq k$ hence y_t is not observed and distributions propagate under the constant assumption.

Using the equations at 1.1 we have the following:

$$\text{Prior for } z_t : \quad z_{t,k} | z_{t-1,k} \sim N(z_{t-1,k}, R_t); \quad R_t = \Sigma_{t-1,k} + W_t \quad (1.2a)$$

$$(1.2b)$$