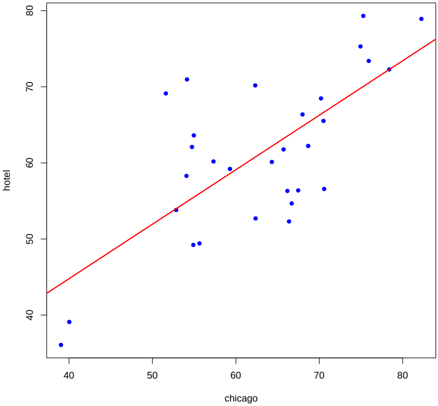
State Space Models and FFBS

Time-Varying Coefficients

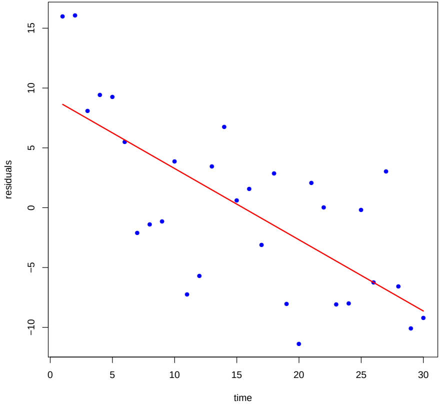
State Space Models

Time-Varying Coefficients

Recall the hotels example where we regressed mothly observations of one hotel's occupancy rate on the overall downtown Chicago occupancy rate:



Here is the time series plot of the residuals.



The trend line is fit to the residuals using

$$r_t = \alpha + \beta\, t + \epsilon$$

The hotel might argue that, based on the plot, there could be some doubt about this simple specification.

To think about a more general model let

$$r_t = \theta_t + \epsilon_t$$

The trend model uses the very "tight" specification:

$$\theta_t = \alpha + \beta\, t.$$

We could be more flexible by transforming t:

$$\theta_t = \alpha + \beta\, t + \gamma\, t^2.$$

Clearly we have to impose some kind of "restriction" on the $\{\theta_t\}$.

We do not what the "perfect" fit: $r_t = \theta_t$.

But how can we avoid the nuisance of picking the transformations?

We can put a random-walk prior on the $\{\theta_t\}$:

$$\theta_t = \theta_{t-1} + W_t, \quad W_t \sim N(0, W^2).$$

If we pick W "small", then we can say each the θ_t can be anything, but successive ones cannot be too different.

Our model (for the residuals) is:

$$p(\theta_0, \theta, r) = p(\theta_0)\, p(\theta \mid \theta_0)\, p(r \mid \theta),$$

where

$$\theta = (\theta_1, \theta_2, \ldots, \theta_T), \;\; r = (r_1, r_2, \ldots, r_T),$$

and,

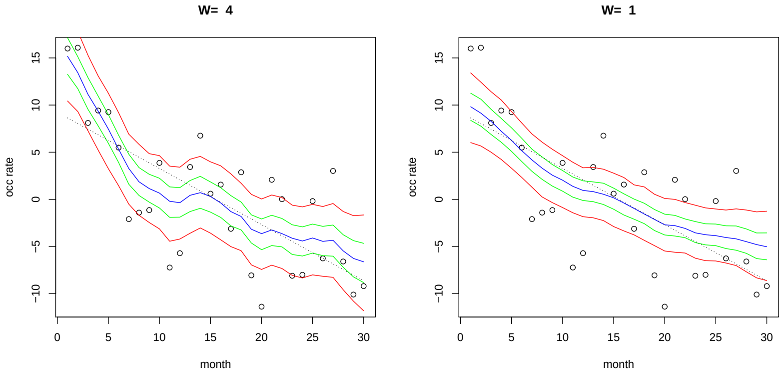
$$p(\theta \mid \theta_0) = \Pi_{t=1}^T \; p(\theta_t \mid \theta_{t-1}), \;\; p(r \mid \theta) = \Pi_{t=1}^T \; p(r_t \mid \theta_t).$$

Using FFBS (forward filtering, backward sampling)
we can get draws:

$$(\theta_0, \theta) \mid r.$$

I ran FFBS and got a nd $\times$ T matrix where each row is a draw of θ.

blue: median of $\{\theta_t\}$ draws.
green: 25% and 75% quantiles of $\{\theta_t\}$ draws.
red: 5% and 95% quantiles of $\{\theta_t\}$ draws.



We can write a comprehensive model for the hotel data (rather than just the residuals):

$$H_t = \theta_t + \beta C_t + v_t, \quad v_t \sim N(0, V).$$
$$\theta_t = \theta_{t-1} + W_t, \quad W_t \sim N(0, W).$$

With priors:

$$p(\theta_0), \quad p(\beta), \quad p(V), \quad p(W).$$

and draw:

$$\begin{aligned} (\theta_0, \theta) \quad &| \quad \ldots \\ \beta \quad &| \quad \ldots \\ V \quad &| \quad \ldots \\ W \quad &| \quad \ldots \end{aligned}$$

We can think of this as a *time-varying parameter* model.

We can start with

$$H_t = \theta + \beta\, C_t + v_t,$$

and then let the intercept vary over time.

It is also very common to let the slope vary over time.

State Space Models

We observe a time series $\{X_t\}$.

We imagine that the distribution of X_t depends on some unobserved "latent" state θ_t which is evolving over time.

Our model consists of the *observation equation*:

$$p(X_t \mid \theta_t),$$

and the *state equation*:

$$p(\theta_t \mid \theta_{t-1}).$$

In addition, we need a prior on the initial state: $p(\theta_0)$.

The general picture:

$$
\begin{array}{cccccccccc}
\theta_0 \rightarrow & \theta_1 \rightarrow & \theta_2 \rightarrow & \cdots & \theta_{t-1} \rightarrow & \theta_t \rightarrow & \theta_{t+1} \rightarrow & \cdots & \rightarrow & \theta_T \\
 & \downarrow & \downarrow & & \downarrow & \downarrow & \downarrow & & & \downarrow \\
 & X_1 & X_2 & & X_{t-1} & X_t & X_{t+1} & & & X_T
\end{array}
$$

Each X is a "peek" at the corresponding θ.

If you margin out the θ's get a model in which future X's depend on past X's.