A 2D Gaussian model (akin to Brigo & Mercurio Section 4.2)

Suppose

\[ r(t) = \delta_0 + X_1(t) + X_2(t) \]

where

\[ dX(t) = - \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} dt + \begin{pmatrix} \sigma_1 & 0 \\ \rho \sigma_2 & \sqrt{1 - \rho^2} \sigma_2 \end{pmatrix} \begin{pmatrix} dW_Q^1(t) \\ dW_Q^2(t) \end{pmatrix} \]

In this case we find (BLACKBOARD) that

\[ B_i(\tau) = \frac{1 - e^{-\kappa_i \tau}}{\kappa_i} \]

and that \( A \) is a rather lengthy expression that we may or may not need. (Brigo & Mercurio Lemma 4.2.1 + Thm. 4.2.1, p. 135.)
The same short rate level may give different yield curves, i.e. $P(t, T) \neq f(r(t), T - t)$. 
Quick & dirty estimation: Calibrate to yield (difference) covariance matrix.

Note that with $\tilde{B}(\tau, \kappa) = \frac{1}{\tau}(B(\tau, \kappa_1), B(\tau, \kappa_2))^\top$ we have

$$\frac{\text{cov}(\Delta y(t, \tau_i)), \Delta y(t, \tau_j))}{\Delta t} \approx \tilde{B}(\tau_i, \kappa)(\begin{pmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \rho \\ \sigma_1 \sigma_2 \rho & \sigma_2^2 \end{pmatrix}) \tilde{B}(\tau_j, \kappa)$$

With a guess of the 5 parameters (forget about $\delta_0$ for a moment) we get a theoretical (approximate, unconditional instantaneous) covariance matrix.

We may try to estimate parameters by getting as close as possible to the empirical covariance matrix.
With yields of 7 maturities, the empirical covariance matrix has effectively \((6 \times 7)/2 = 21\) entries.

A simple least squares fit to 50 years of US data gives (R-code and data on homepage)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>(\kappa_1)</th>
<th>(\kappa_2)</th>
<th>(\sigma_1)</th>
<th>(\sigma_2)</th>
<th>(\rho)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated value</td>
<td>0.631</td>
<td>0.194</td>
<td>0.033</td>
<td>0.031</td>
<td>−0.834</td>
</tr>
</tbody>
</table>
And that gives a picture like this for the standard deviations (calibrate to covariance, show standard deviations and correlations in graphs)
And for the correlations:
Observations:

- Not the worst fit, you’ll ever see.

- We need a high negative correlation between factors to make yields as uncorrelated as they are empirically.

- We can use $\delta_0$ to calibrate to today’s observed yield curve as earlier.
More observations:

- Parameters aren't really identified; just switch indices.

- “Proper” inference: Do maximum likelihood; it’s just a Gaussian first-order vector auto-regression.

- Problem: Factors are not observable. Solution: Invert to express in terms of yields. Problem: Parameter dependent transform \( \sim \) Jacobian.

- If we want to use all observed yields, we get some kind of filtering problem.

- Models are affine in data — not in parameters. (This non-linearity is *menier Meinung nach* the main complication. Can we reparametrize?)
• The whole \( \mathbb{P} \) vs. \( \mathbb{Q} \) or parameter risk-premium question pops up again — with a vengeance!

• In the empirical covariance matrix we averaged out any conditional information. Consistent w/ a Gaussian model; not necessarily w/ data.
Messing with your head, I (Rotation, or $Ar$ models in the language of Dai & Singleton)

Suppose that somebody (messr’s Hull & White for instance) comes along with a model like this:

$$dr(t) = (\theta + u(t) - ar(t))dt + \sigma_1 dW_1$$

where

$$du(t) = -bu(t)dt + \sigma_2 dW_2$$

where $dW_1dW_2 = \rho dt$.

Looks “sexy”: It’s Vasicek with stochastic mean reversion level. And correlation. And they can even find ZCB prices.

It is, however, just the 2D Gaussian model in disguise!

BLACKBOARD Or Brigo & Mercurio Section 4.2.5, p. 149.
Messing with your head, II

That $\beta$’s are all 0 is because we want a Gaussian model. Fair enough. But:

- Why is $\delta^\top = (1, 1)$?
- Why is $\theta = 0$?
- Why is $K$ diagonal?
- Why is $\Sigma_{1,2} = 0$?
- Why is $\alpha^\top = (1, 1)$?

Are they real restrictions or just needed for identification, or for us to obtain closed-form solutions?
• The variable $\tilde{X}_i = \delta_i X_i$ has same $\kappa_i$, and just scaled volatility.

• The variable $\tilde{X}_i = X_i - \theta_i$ is a Gaussian process that mean reverts to 0. Shift absorbed by $\delta_0$. (Aside: “CIR + constant” isn’t CIR. This so-called displacement can come in handy.)

• If $K$ can be diagonalized (note: $K$ is not symmetric), say by $M$ ie.

$$MKM^{-1} = D,$$

then with $Y = MX$ we have

$$dY = d(MX) = -MKX dt + M\Sigma dW = -DMY dt + M\Sigma dW$$

$$= -DY dt + \tilde{\Sigma} dW,$$
“and we’re good”.

- At least $K$ can be made lower triangular, by defining $\widetilde{X}_i$’s in a “Gaussian elimination” way. We get $B$ ODEs with a simple recursive structure. (To avoid degenerate cases, diagonal elements are non-0.)

- Volatility terms enter only through the symmetric matrix $\Sigma \Sigma^T$, so 3 free parameters are enough.

- Given some $\Sigma$, we can diagonalize $\Sigma \Sigma^T$ by $M$ and then use $M$ to rotate and get diagonal volatility — but ruin a diagonal $K$.

- In short: This is the 2D Gaussian model.

Here we’ve actually proven Dai & Singleton’s characterization (section B.1) of $\mathbb{A}_0(N)$-models. (They use $\Sigma = I$, rather than $\delta^T = (1, \ldots, 1)$.)

Independent CIRs

Suppose

\[ r(t) = \delta_1 X_1(t) + \delta_1 X_2(t) \]

where the \( X \)'s are independent CIR-type processes

\[ dX_i(t) = \kappa_i (\theta_i - X_i(t)) dt + \sqrt{X_i(t)} dW_i(t) \]

Fits the general framework. But the ZCB price formula immediately reduces to a product of CIR-formulas.
Can we make correlated CIRs just saying $dW_1dW_2 = \rho dt$?

Yes, but we can't solve for ZCB prices (with the ODEs here, at least), because it's not an affine model:

$$[\Sigma \Sigma^\top] = \rho \sqrt{X_1} \sqrt{X_2} \neq a + b^\top X$$

(Chen (1994) actually has something on this.)

CIRs can be made correlated through the drift, but only positively — otherwise we get well-definedness (admissibility) problems ($\sqrt{<0}$).
Making Independent CIRs Look Good

Rewrite to Longstaff/Schwartz stochastic volatility. An exercise?

We get a richer (state-variable dependent) conditional variance, but “loose on correlation”.
Dai & Singleton’s Canonical Representation

BLACKBOARD
Some Named Models

Pure Gaussian: Langetieg. Can find closed-form ZCB-solutions. I usually use diagonal $K$ and non-diagonal $\Sigma$ (for ease).

2 Gaussian, 1 CIR: Das, Balduzzi, Foresi & Sundaram. ZCB-solutions w/ special functions. Not the most flexible model i $A_1(3)$.

1 Gaussian, 2 CIR: Chen. ZCB-solutions w/ special functions. Not the most flexible model i $A_2(3)$.