

Nonlinear responses

Tyler Sotomayor

2024

Worked examples of transition timing, likelihood whitening, endogenous state changes, and asymmetric quadratic-shock responses.

Based on the 2024 BSE course taught by Luca Gambetti and Konstantin Boss. I have revised the classroom code and added the accompanying explanations. The notes identify the papers, data, and assumptions used in each application. These are my course notes, not an official BSE publication. A classroom replication should not be read as a replication of every result in its underlying paper.

1 State weights and timing

Because $e^{1.5} \simeq 4.4817$ and $e^{-1.5} \simeq 0.2231$,

$$F(-1) \simeq 0.8176, \quad F(0) = 0.5, \quad F(1) \simeq 0.1824.$$

Low growth receives the high recession weight, and opposite standardized states have complementary weights. This is an immediate direction check for the sign in the exponential. The implementation's stable formula also returns finite weights at states of magnitude 1000.

With three lags, fitted equations are dated $t = 4, 5, 6, 7, 8$. Their state observations are s_3, s_4, s_5, s_6, s_7 , respectively. Using s_4 for the equation at $t = 4$ would use the contemporaneous state; using s_2 would introduce an extra unintended lag.

Full-sample standardization uses observations later than some historical forecast origins to estimate the mean and standard deviation. It is a well-defined retrospective normalization, but not the same information set available at each origin. Moreover, the raw state itself must have a documented timing construction. Lagging a centered moving average once does not necessarily remove all future observations from that average.

2 Whitening a likelihood

The standardized residuals are $2/2 = 1$ and $3/3 = 1$. The negative Gaussian log likelihood is therefore

$$\mathcal{L} = \log(2\pi) + \log 2 + \log 3 + \frac{1}{2}(1 + 1) \simeq 4.6296.$$

Dropping the constant subtracts $\log(2\pi) \simeq 1.8379$, giving about 2.7918. The minimizing parameter values do not change when the same fixed-sample constant is omitted, but reported objective levels cannot be compared until the convention is reconciled. This is precisely the adjustment used in the classroom likelihood audit.

For a common $T \times k$ design X and constant $n \times n$ covariance Σ , the GLS normal equations for the coefficient matrix D are $X'(Y - XD)\Sigma^{-1} = 0$. Multiplying on the right by Σ gives $X'Y - X'XD = 0$,

hence $D = (X'X)^{-1}X'Y$. The production implementation uses QR to solve the equivalent system rather than forming that inverse. Equal regime covariances make weights irrelevant to the covariance step; they do not remove state interactions from a mean design that already contains them.

3 A state that changes after impact

Initially $s_{t-1} = -1 < 0$, so both paths use coefficient 0.8 at impact. The unshocked value is $y_t^0 = -0.8$; the shocked value is $y_t^1 = -0.8 + 1 = 0.2$. Their impact difference is one.

The histories then imply different next states. The unshocked path remains in the high-persistence regime, giving $y_{t+1}^0 = 0.8(-0.8) = -0.64$. The shocked path has crossed zero and uses coefficient 0.2, giving $y_{t+1}^1 = 0.2(0.2) = 0.04$. The horizon-one difference is

$$0.04 - (-0.64) = 0.68.$$

Freezing the initial coefficient at 0.8 on both paths would instead give the difference $0.8 \times 1 = 0.8$. The disagreement is not a simulation error. One calculation lets the shock change the future regime; the other deliberately prevents that channel. In a generalized response with random future innovations, the comparison would also integrate over future innovation paths under a stated conditional distribution.

4 Shock size and asymmetry

The response function at this horizon is $r(\delta) = -0.5\delta - 0.2\delta^2$. Thus

Shock	Linear part	Quadratic part	Total
+1	-0.5	-0.2	-0.7
-1	+0.5	-0.2	+0.3
+2	-1.0	-0.8	-1.8
-2	+1.0	-0.8	+0.2

Doubling the positive shock changes the response from -0.7 to -1.8, not -1.4. Both the positive and negative shocks have a negative quadratic component. A large negative shock can therefore have a smaller positive effect than a small negative shock in this example.

Adding and subtracting unit-shock responses recovers

$$C_h a = \frac{r(1) - r(-1)}{2} = \frac{-0.7 - 0.3}{2} = -0.5, \quad C_h b = \frac{r(1) + r(-1)}{2} = \frac{-0.7 + 0.3}{2} = -0.2.$$

Each response compares the chosen shock with zero, holding the history and other innovations fixed. It is not a deviation from the unconditional mean of the squared innovation. For a unit-variance shock, $E(\eta^2) = 1$, whereas the zero-shock baseline has $\eta^2 = 0$.