Solving Stochastic Dynamic Equilibrium Models: A k-Order Perturbation Approach

(Preliminary version)

Michel Juillard and Ondra Kamenik

Autumn 2004

Abstract
Among the many approaches currently used for solving stochastic dynamic equilibrium models, the most promising for solving medium size models is the perturbation method. Several papers have shown how a second order method provides approximations of the level effect of future volatility on current decision. Higher order approximations take into account how future volatility modifies also the curvature of the decision rules.

This paper presents an algorithm for using an arbitrary k-order perturbation method for solving stochastic dynamic equilibrium models. A first order approximation is equivalent to solving a linearized version of the original model and encounters the usual difficulties arising from the matrix polynomial equation at the heart of linear rational expectation models.

However, for all orders larger than one, the mathematical problem is always the same and involves only solving linear problems. The recurrence is put in evidence by expressing the model as a combination of multivariate functions and exploiting the properties of partial derivatives of combination of function and of the implicit function theorem.

Some of the linear systems encountered in deriving the k-order approximation aren't standard and require the adaptation of an algorithm used for Sylvester equations in order to be solved efficiently.

There has been recently a lot of interest in stochastic dynamic equilibrium models. There is however no unanimity on the best way of solving and simulating them. The perturbation approach seems promising because it is possible to handle easily a larger state space than with other approaches (**?**). Several papers have analysed second order approximation (**?????**). **?** discuss higher order approximations. In this article, we present recursive formulas to compute approximations at an arbitrary order. This is made possible by a concise formulation of the problem and a recursive formula for the k–order partial derivatives of the combination of two multivariate functions.

In the first section, I present a general version of the problem to be solved. The second section deals with the main features of the perturbation approach and discusses the Taylor expansion of the model. I the third section, I recall briefly the computation of a linear approximation of the model. The recursive computation of k–order approximation is developped in the fourth section. The fifth section discusses some issues in the programming of the algorithm.

1 Model

We consider a general non–linear rational expectation model for a vector of n variables, y_t:

$$E_t\left(f(y^*_{t-1}, y_t, y^{**}_{t+1}, u_t)\right) = 0.$$

There are n_1 predetermined variables, y^* and $n_2 = n - n_1$ forward looking variables y^{**}[1]. The system is affected by m stochastic shocks, u_t:

$$u_t = \sigma \eta_t$$

where σ is a scale parameter and the moments of η are noted[2]

$$E\left\{\eta_{i_1}\eta_{i_2}\dots\eta_{i_k}\right\} = \left[\Sigma^{(k)}\right]_{i_1 i_2 \dots i_k} \qquad\qquad i_1, i_2, \dots, i_k = 1, \dots, m$$

It is further assumed that stochastic shocks u_t are known at the beginning of period t and $E_t()$ means the expectation conditional on the realization of u in period t and all previous periods. By construction, the shocks u_t have no serial correlation, but time dependency of the shocks can be modeled with adding engogenous variables.

[1] A more general version of this model, with variables that are both predetermined and forward–looking as well as with static variables can always be written in the form of the model discussed here with the addition of auxiliary variables and equations.

[2] Throughout the paper, I use a tensor based notation: $[A]_{i_1 i_2 \dots i_k}$ refers to element $i_1, i_2, \dots, i_k$ of the k–dimension object A.

Finding an dynamic equilibrium solution for this model is equivalent to finding the following function:

$$y_t = g(y^*_{t-1}, u_t, \sigma).$$

In what follows, it will be usefull to distinguish, among the solution functions, the transition functions for the predetermined variables:

$$y^*_t = g^*(y^*_{t-1}, u_t, \sigma)$$

and the decision functions for the forward looking variables:

$$y^{**}_t = g^{**}(y^*_{t-1}, u_t, \sigma).$$

The future value of forward looking variables is then given by

$$\begin{aligned} y^{**}_{t+1} &= g^{**}\left(y^*_t, u_{t+1}, \sigma\right) \\ &= g^{**}\left(g^*\left(y^*_{t-1}, u_t\right), u_{t+1}, \sigma\right). \end{aligned}$$

The original model can be rewritten as function of state variables, y^*_{t-1}, u_t, σ, and of future shocks, u_{t+1}:

$$F\left(y^*_{t-1}, u_t, \sigma, u_{t+1}\right) = f\left(y^*_{t-1}, g\left(y^*_{t-1}, u_t, \sigma\right), g^{**}\left(g^*\left(y^*_{t-1}, u_t, \sigma\right), u_{t+1}, \sigma\right), u_t\right).$$

2 The perturbation approach and the Taylor expansion of the model

The basic idea of the perturbation approach is to use a Taylor expansion of the model above to recover the coefficients of the Taylor expansion of the solution functions, $g(y^*_{t-1}, u_t, \sigma)$.

The Taylor expansion is taken around the non–stochastic and static equilibrium of the system, where $u_t = u_{t+1} = 0$ and $\sigma = 0$. The non–stochastic, static equilibrium, $\bar{y}$, satisfies

$$f(\bar{y}^*, \bar{y}, \bar{y}^{**}, 0) = 0.$$

In the same way, the Taylor expansion for the solution functions, $g(y^*_{t-1}, u_t, \sigma)$, will be expressed around $g(\bar{y}^*, 0, 0)$.

It is simpler to write the expansion of the model, if one collects y^*_{t-1} and u_t in a single vector s_t as they are all known at period t. Furthermore, still to make the expression simpler, we will get rid of the time subscript, using $s = s_t$, $u = u_t$ and

$u' = u_{t+1}$. The p order Taylor expansion of the ith equation of the model around $(\bar{s}, 0, 0)$, where $\bar{s} = \begin{bmatrix} \bar{y} \\ 0 \end{bmatrix}$, is written

$$\begin{aligned} F_i^{(p)}(s, \sigma, u') &= F_i(\bar{s}, 0, 0) + \sum_{j=1}^{p} \frac{1}{j!} \Bigg(\left[F^i_{s^j}\right]_{\alpha_1 \ldots \alpha_j} [\hat{s}]^{\alpha_1} \ldots [\hat{s}]^{\alpha_j} \\ &\quad + \sum_{k=1}^{j-1} \binom{j}{k} \left[F^i_{s^k \sigma^{j-k}}\right]_{\alpha_1 \ldots \alpha_k} [\hat{s}]^{\alpha_1} \ldots [\hat{s}]^{\alpha_k} \sigma^{j-k} \\ &\quad + \sum_{k=1}^{j-1} \binom{j}{k} \left[F^i_{s^k u'^{j-k}}\right]_{\alpha_1 \ldots \alpha_k \beta_1 \ldots \beta_{j-k}} [\hat{s}]^{\alpha_1} \ldots [\hat{s}]^{\alpha_k} [u']^{\beta_1} \ldots [u']^{\beta_{j-k}} \\ &\quad + \sum_{k=1}^{j-1} \binom{j}{k} \left[F^i_{u'^k \sigma^{j-k}}\right]_{\beta_1 \ldots \beta_k} [u']^{\beta_1} \ldots [u']^{\beta_k} \sigma^{j-k} \\ &\quad + \sum_{k=1}^{j-2} \sum_{m=k+1}^{j-1} \binom{j}{k, m-k, j-m} \left[F^i_{s^k u'^{m-k} \sigma^{j-m}}\right]_{\alpha_1 \ldots \alpha_k \beta_1 \ldots \beta_{m-k}} [\hat{s}]^{\alpha_1} \ldots [\hat{s}]^{\alpha_k} [u']^{\beta_1} \ldots \\ &\quad + \left[F^i_{\sigma^j}\right] \sigma^j + \left[F^i_{u'^j}\right]_{\beta_1 \ldots \beta_j} [u']^{\beta_1} \ldots [u']^{\beta_j} \Bigg) \end{aligned}$$

where $\hat{s} = s - \bar{s}$ and $\left[F^i_{s^j}\right]_{\alpha_1 \ldots \alpha_j k} = \frac{\partial^j F}{\partial s_{\alpha_1} \ldots \partial s_{\alpha_j}}$. In tensor notation, the same index used first as subscript and then superscript of two tensors implies summation of the products[3].

At time t, the variables in the state vector s and the scale variable σ are known, therefore the only random quantity entering the conditional expectation is u' which is, by definition, uncorrelated with previous information. Because of the linear form of the Taylor expansion, it is possible to express the conditional

[3] i.e.

$$[A]_{\alpha\beta} [B]^{\alpha} [C]^{\beta} = \sum_i \sum_j A_{ij} B_i C_j.$$

expectation of the model as functions of the moments of future shocks:

$$
\begin{aligned}
\bar{F}_i^{(p)}(s,\sigma) &= E_t\left(F_i^{(p)}(s,\sigma,u')\right)\\
&= F_i(\bar{s},0,0) + \sum_{j=1}^{p} \frac{1}{j!}\left(\left[F^i_{s^j}\right]_{\alpha_1\ldots\alpha_j}[\hat{s}]^{\alpha_1}\ldots[\hat{s}]^{\alpha_j}\right.\\
&\quad + \sum_{k=1}^{j-1}\binom{j}{k}\left[F^i_{s^k\sigma^{j-k}}\right]_{\alpha_1\ldots\alpha_k}[\hat{s}]^{\alpha_1}\ldots[\hat{s}]^{\alpha_k}\,\sigma^{j-k}\\
&\quad + \sum_{k=1}^{j-1}\binom{j}{k}\left[F^i_{s^k u'^{j-k}}\right]_{\alpha_1\ldots\alpha_k\beta_1\ldots\beta_{j-k}}[\hat{s}]^{\alpha_1}\ldots[\hat{s}]^{\alpha_k}\,[\Sigma]^{\beta_1\ldots\beta_{j-k}}\,\sigma^{j-k}\\
&\quad + \sum_{k=1}^{j-1}\binom{j}{k}\left[F^i_{u'^k\sigma^{j-k}}\right]_{\beta_1\ldots\beta_k}[\Sigma]^{\beta_1\ldots\beta_k}\,\sigma^{j}\\
&\quad + \sum_{k=1}^{j-2}\sum_{m=k+1}^{j-1}\binom{j}{k,m-k,j-m}\left[F^i_{s^k u'^{m-k}\sigma^{j-m}}\right]_{\alpha_1\ldots\alpha_k\beta_1\ldots\beta_{m-k}}[\hat{s}]^{\alpha_1}\ldots[\hat{s}]^{\alpha_k}\,[\Sigma]^{\beta_1\ldots\beta_{m-k}}\\
&\quad \left. + \left[F^i_{\sigma^j}\right]\sigma^j + \left[F^i_{u'^j}\right]_{\beta_1\ldots\beta_j}[\Sigma]^{\beta_1\ldots\beta_j}\,\sigma^j\right)\\
&= 0,
\end{aligned}
$$

$i = 1,\ldots,n.$

All the terms involving the moments of future shocks represent the deterministic effects of taking into account future uncertainty. They entail departure from certainty equivalence.

Regrouping terms, one finds that necessary and sufficient conditions for the above system of equations to be satisfied for any value of $\hat{s}$ and σ are

$$
\begin{aligned}
\left[F^i_{s^j}\right]_{\alpha_1\ldots\alpha_j} &= 0\\
\left[F^i_{\sigma^j}\right] + \left[F^i_{u'^j}\right]_{\beta_1\ldots\beta_j}[\Sigma]^{\beta_1\ldots\beta_j} + \sum_{k=1}^{j-1}\binom{j}{k}\left[F^i_{u'^k\sigma^{j-k}}\right]_{\beta_1\ldots\beta_k}[\Sigma]^{\beta_1\ldots\beta_k} &= 0\\
\left[F^i_{s^k\sigma^{j-k}}\right]_{\alpha_1\ldots\alpha_k} + \left[F^i_{s^k u'^{j-k}}\right]_{\alpha_1\ldots\alpha_k\beta_1\ldots\beta_{j-k}}[\Sigma]^{\beta_1\ldots\beta_{j-k}} + &\\
\sum_{m=1}^{j-k-1}\binom{j-k}{m}\left[F^i_{s^k u'^m\sigma^{j-k-m}}\right]_{\alpha_1\ldots\alpha_k\beta_1\ldots\beta_m}[\Sigma]^{\beta_1\ldots\beta_m} &= 0\\
i = 1,\ldots,n \qquad j = 1,\ldots,p \qquad k = 1,\ldots,j-1 &
\end{aligned}
$$

3 Computation of first order partial derivatives

Computing the first order approximation using the perturbation approach is nothing else but solving a linearized version of the model, a problem for which solutions exist since a long time. I use here the solution advocated by **?** and **?**. This section, although not innovative, is necessary for fixing notations and as a starting block for computing approximations at higher orders.

As we will see shortly, the first order approximation is more difficult from a mathematical point of view because it entails solving a matrix polynomial equation when, for higher orders, it is only necessary to solved large linear problems.

At the first order, only the first moment of future shocks enters the equations. As, in the case of a normal distribution, it equals zero, the equations to be satisfied are only

$$\left[F_s^i\right]_{\alpha_1} = 0$$

The computation of the partial derivatives of the solution functions is done in sequence, starting with the derivatives with respect to the dynamic variables of the model. then, computing the partial derivatives with respect to the exogenous stochastic shocks.

At the first order, it is easier to use a matrix notation:

$$\begin{aligned} [F_{y^*}] &= \left[f_{y^*_-}\right] + [f_{y^*}]\left[g^*_{y^*}\right] + [f_{y^{**}}]\left[g^{**}_{y^*}\right] + \left[f_{y^{**}_+}\right]\left[g^{**}_{y^*}\right]\left[g^*_{y^*}\right] \\ &= 0 \end{aligned}$$

where $[F_{y^*}]$ represents the Jacobian of the system of equations $F()$ with respects to the state variables y^*_{t-1}, $\left[f_{y^*_-}\right]$ is the Jacobian of the original system of equations with respect to the lagged value of the predetermined variables, $[f_{y^*}]$, with respect to the current value of the predetermined variabels, $[f_{y^{**}}]$, with respect to the current value of the forward–looking variables, and $\left[f_{y^{**}_+}\right]$ with the future value of forward–looking variables. This is a matrix polynomial equation where the unkowns are the matrices $\left[g^*_{y^*}\right]$ and $\left[g^{**}_{y^*}\right]$).

The above equation can be rewritten as

$$\begin{bmatrix} f_{y^*} & f_{y^{**}_+} \end{bmatrix} \begin{bmatrix} I \\ g^{**}_{y^*} \end{bmatrix} g^*_{y^*} = \begin{bmatrix} -f_{y^*_-} & -f_{y_{**}} \end{bmatrix} \begin{bmatrix} I \\ g^{**}_{y^*} \end{bmatrix}$$

This matrix polynomial equation can then be solved using a real generalized Schur decomposition:

$$\begin{bmatrix} f_{y^*} & f_{y_+^{**}} \end{bmatrix} = Q \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} Z^T$$
$$\begin{bmatrix} -f_{y_-^*} & -f_{y_{**}} \end{bmatrix} = Q \begin{bmatrix} S_{11} & S_{12} \\ 0 & S_{22} \end{bmatrix} Z^T$$

where Q and Z are real, orthogonal matrices, T_{11} and T_{22} are triangular matrices and S_{11} and S_{22} are block–triangular matrices. The generalized eigenvalues $\lambda_i = S_{ii}/T_{ii}$ are ordered such that to have a modulus smaller than one in the first block and larger than one (and possibly infinite, if $T_{ii} = 0$)[4] in the second one.

Premultiplying by Q^T, one obtains

$$\begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} Z^T \begin{bmatrix} I \\ g_{y^*}^{**} \end{bmatrix} g_{y^*}^* = \begin{bmatrix} S_{11} & S_{12} \\ 0 & S_{22} \end{bmatrix} Z^T \begin{bmatrix} I \\ g_{y^*}^{**} \end{bmatrix}$$

Exploding trajectories are excluded when

$$\begin{bmatrix} Z_{12}^T & Z_{22}^T \end{bmatrix} \begin{bmatrix} I \\ g_{y^*}^{**} \end{bmatrix} = 0$$

or

$$g_{y^*}^{**} = -\left(Z_{22}^T\right)^{-1} Z_{12}^T$$

It follows that

$$g_{y^*}^* = \left(T_{11}\left(Z_{11}^T + Z_{21}^T g_{y^*}^{**}\right)\right)^{-1} S_{11}\left(Z_{11}^T + Z_{21}^T g_{y^*}^{**}\right)$$

$\left[g_{y^*}^*\right]$ and $\left[g_{y^*}^{**}\right]$ let us form $[g_{y^*}] = \begin{bmatrix} g_{y^*}^* \\ g_{y^*}^{**} \end{bmatrix}$. The derivatives with respect to u_t can be directly obtained from

$$\begin{aligned} \left[F_u^i\right] &= [f_u] + [f_y]\,[g_u] + \left[f_{y_+^{**}}\right]\left[g_{y^*}^{**}\right]\left[g_u^*\right] \\ &= 0. \end{aligned}$$

And,

$$[g_u] \;=\; -\left([f_y] + \begin{bmatrix} \left[f_{y_+^{**}}\right]\left[g_{y^*}^{**}\right] & I \end{bmatrix}\right)^{-1}[f_u]$$

[4]The case $s_{ii} = t_{ii} = 0$ is excluded by requiring that the model should have multiple equilibria only in the presence of a unit root.

All first derivatives with respect to σ are null from

$$\begin{aligned}\left[F^i_\sigma\right] &= [f_y]\,[g_\sigma] + \left[f_{y^{**}_+}\right][g^{**}_\sigma] + \left[f_{y^{**}_+}\right]\left[g^{**}_{y^*}\right][g^*_\sigma] \\ &= 0.\end{aligned}$$

The fact that, at the first order, the derivatives of $g()$ with respect to σ are null is the manifestation of the certainty equivalence principle: the volatility of future shocks doesn't matter for today's decisions.

4 kth order derivatives

After the first derivatives, all derivatives of higher order are computed recursively using the same formulas.

Most of what follows depends on the kth order partial derivatives of the composite of two functions, given the partial derivatives of each function. If $F(r) = f(z(r))$,

$$\left[F^i_{r^j}\right]_{\alpha_1\ldots\alpha_j} = \sum_{l=1}^{j}\left[f^i_{z^l}\right]_{\beta_1\ldots\beta_l}\sum_{c\in\mathcal{M}_{l,j}}\prod_{m=1}^{l}\left[z_{r^{|c_m|}}\right]^{\beta_m}_{\boldsymbol{\alpha}(c_m)}$$

where $\mathcal{M}_{l,j}$ is the set of all partitions of the set of j indices with l classes, $|.|$ is the cardinality of a set, c_m is m-th class of partition c, and $\boldsymbol{\alpha}(c_m)$ is a sequence of α's indexed by c_m. Note that $\mathcal{M}_{1,j} = \{\{1,\ldots,j\}\}$ and $\mathcal{M}_{j,j} = \{\{1\},\{2\},\ldots,\{j\}\}$.[5] In order to keep the formulas compact, we will use $\boldsymbol{\alpha}_n$ for $\alpha_1\ldots\alpha_n$.

Let's define

$$r = \begin{bmatrix} y^*_{t-1} \\ u_t \\ \sigma \\ u_{t+1}\end{bmatrix} \quad \text{and} \quad z(r) = \begin{bmatrix} y^*_{t-1} \\ u_t \\ y_t \\ y^{**}_{t+1}\end{bmatrix} = \begin{bmatrix} y^*_{t-1} \\ u_t \\ g(y^*_{t-1},u_t,\sigma) \\ g^{**}(g^*(y^*_{t-1},u_t,\sigma),u_{t+1},\sigma)\end{bmatrix}$$

Indeed, $F(r) = f(z(r))$ and we can use the formula to retrieve derivatives of g.

[5] Example for the third partial derivatives

$$\begin{aligned}\left[F^i_{r^3}\right]_{\alpha_1\alpha_2\alpha_3} &= \left[f^i_z\right]_{\beta_1}[z_{r^3}]^{\beta_1}_{\alpha_1\alpha_2\alpha_3} + \left[f^i_{z^2}\right]_{\beta_1\beta_2}\Big([z_r]^{\beta_1}_{\alpha_1}[z_{r^2}]^{\beta_2}_{\alpha_2\alpha_3} + [z_r]^{\beta_1}_{\alpha_2}[z_{r^2}]^{\beta_2}_{\alpha_1\alpha_3} \\ &\quad + [z_r]^{\beta_1}_{\alpha_3}[z_{r^2}]^{\beta_2}_{\alpha_1\alpha_2}\Big) + \left[f^i_{z^3}\right]_{\beta_1\beta_2\beta_3}[z_r]^{\beta_1}_{\alpha_1}[z_r]^{\beta_2}_{\alpha_2}[z_r]^{\beta_3}_{\alpha_3}\end{aligned}$$

Note that the first order partial derivatives of z with respect to y^*_{t-1} are given by

$$\left[z_{y^*}\right]_{\alpha_1} = \begin{bmatrix} [I]_{\alpha_1} \\ 0 \\ \left[g_{y^*}\right]_{\alpha_1} \\ \left[g^{**}_{y^*}\right]_{\gamma_1} \left[g^*_{y^*}\right]^{\gamma_1}_{\alpha_1} \end{bmatrix},$$

and, for higher order, $k > 1$,

$$\left[z_{y^*k}\right]_{\boldsymbol{\alpha}_k} = \begin{bmatrix} 0 \\ 0 \\ \left[g_{y^*k}\right]_{\boldsymbol{\alpha}_k} \\ \left[g^{**}_{y^*}\right]_{\gamma_1} \left[g^*_{y^*k}\right]^{\gamma_1}_{\boldsymbol{\alpha}_k} + \\ \sum_{l=2}^{k-1} \left[g^{**}_{y^*l}\right]_{\boldsymbol{\gamma}_l} \sum_{c\in\mathcal{M}_{l,k}} \prod_{m=1}^{l} \left[g^*_{y^*|c_m|}\right]^{\gamma_m}_{\boldsymbol{\alpha}(c_m)} + \\ \left[g^{**}_{y^*k}\right]_{\boldsymbol{\gamma}_k} \prod_{m=1}^{k} \left[g^*_{y^*}\right]^{\gamma_m}_{\alpha_k} \end{bmatrix}.$$

4.1 Recovering $\left[g_{y^*k}\right]$

In order to calculate derivatives $\left[g_{y^*k}\right]$ we use $\left[F_{y^*k}\right] = 0$. Putting the unknown derivatives, $\left[g_{y^*k}\right]$, on the left handside, one gets

$$\left[f_z\right]_\beta \begin{bmatrix} 0 \\ 0 \\ \left[g_{y^*k}\right]_{\boldsymbol{\alpha}_k} \\ \left[g^{**}_{y^*}\right]_{\gamma} \left[g^*_{y^*k}\right]^{\gamma}_{\boldsymbol{\alpha}_k} + \left[g^{**}_{y^*k}\right]_{\boldsymbol{\gamma}_k} \prod_{m=1}^{k} \left[g^*_{y^*}\right]^{\gamma_m}_{\alpha_m} \end{bmatrix}^{\beta} = \left[B_{y^*k}\right]_{\boldsymbol{\alpha}_k}$$

with

$$\begin{aligned} \left[B_{y^*k}\right]_{\boldsymbol{\alpha}_k} &= -\left[f_{y^{**}}\right]_\beta \left[\sum_{l=2}^{k-1} \left[g^{**}_{y^*l}\right]_{\boldsymbol{\gamma}_l} \sum_{c\in\mathcal{M}_{l,k}} \prod_{m=1}^{l} \left[g^*_{y^*|c_m|}\right]^{\gamma_m}_{\boldsymbol{\alpha}(c_m)}\right]^{\beta} \\ &\quad - \sum_{l=2}^{k} \left[f_{z^l}\right]_{\boldsymbol{\beta}_l} \sum_{c\in\mathcal{M}_{l,k}} \prod_{m=1}^{l} \left[z_{y^*|c_m|}\right]^{\beta_m}_{\boldsymbol{\alpha}(c_m)}. \end{aligned}$$

All the partial derivatives of $g()$ on the RHS are of order lower than k and have already been computed. Finding the unknown derivatives at order k entails solving a linear problem.

The above expression can be put back in matrix form. The higher order partial derivatives are arranged in a matrix, each row corresponding to a different function and all the partial cross–derivatives unfolding in columns[6]

$$[f_{y^*}]\left[g^*_{y^{*k}}\right]+[f_{y^{**}}]\left[g^{**}_{y^{*k}}\right]+\left[f_{y^{**}_+}\right]\left[g^{**}_{y^*}\right]\left[g^*_{y^{*k}}\right]+\left[f_{y^{**}_+}\right]\left[g^{**}_{y^{*k}}\right]\left[G^*_{y^{*k}}\right]=\left[B_{y^{*k}}\right]$$

where $\left[G^*_{y^{*k}}\right]=\bigotimes_{m=1}^k\left[g^*_{y^*}\right]$. This can be rewritten more compactly:

$$\left[\begin{array}{ccc}[f_{y^*}]+\left[f_{y^{**}_+}\right]\left[g^{**}_{y^*}\right] & \vdots & [f_{y^{**}}]\end{array}\right]\left[\begin{array}{c}\left[\begin{array}{c}g^*_{y^{*k}}\\ \hline g^{**}_{y^{*k}}\end{array}\right]\end{array}\right]+\left[\begin{array}{ccc}0 & \vdots & \left[f_{y^{**}_+}\right]\end{array}\right]\left[\begin{array}{c}\left[\begin{array}{c}g^*_{y^{*k}}\\ \hline g^{**}_{y^{*k}}\end{array}\right]\end{array}\right]\left[G^*_{y^{*k}}\right]=\left[B_{y^{*k}}\right].$$

It is a system of equations of the form

$$AX+BXC=D.$$

Such systems, when they are small, can be solved by vectorizing the system[7]. It is however more efficient to solve them using an algorithm for the Sylvester equation (see Appendix B).

4.2 Recovering $\left[g_{y^{*i}u^j}\right]$

The other k-order derivatives with respect to the state variables and to the exogenous variables are obtained from $\left[F^k_s\right]=0$ which is rewritten as:

$$[f_z]_\beta\left[\begin{array}{c}0\\0\\ \left[g_{s^k}\right]_{\boldsymbol{\alpha}_k}\\ \left[g^{**}_{y^*}\right]_\gamma\left[g^*_{s^k}\right]^\gamma_{\boldsymbol{\alpha}_k}\end{array}\right]^\beta=\left[B_{s^k}\right]_{\boldsymbol{\alpha}_k}$$

---

[6]For example, one would write the third order derivatives of two functions of two variables: $[f_{y^3}]=\left[\begin{array}{cccc}\frac{\partial^3 f_1}{\partial y_1^3} & \frac{\partial^3 f_1}{\partial y_1^2\partial y_2} & \frac{\partial^3 f_1}{\partial y_1\partial y_2^2} & \frac{\partial^3 f_1}{\partial y_2^3}\\ \frac{\partial^3 f_2}{\partial y_1^3} & \frac{\partial^3 f_2}{\partial y_1^2\partial y_2} & \frac{\partial^3 f_2}{\partial y_1\partial y_2^2} & \frac{\partial^3 f_2}{\partial y_2^3}\end{array}\right]$.

[7]They are put in the form

$$\left[\begin{array}{ccc}I\otimes A & \vdots & C^T\otimes B\end{array}\right]\mathrm{vec}\,X=\mathrm{vec}\,B.$$

with

$$
\begin{aligned}
[B_{s^k}]_{\boldsymbol{\alpha}_k} &= -[f_{y^{**}}]_\beta \left[\sum_{l=2}^{k} \left[g^{**}_{y^{*l}} \right]_{\boldsymbol{\gamma}_l} \sum_{c \in \mathcal{M}_{l,k}} \prod_{m=1}^{l} \left[g^{*}_{s^{|c_m|}} \right]^{\gamma_m}_{\boldsymbol{\alpha}(c_m)} \right]^{\beta} \\
&\quad - \sum_{l=2}^{k} \left[f^i_{z^l} \right]_{\boldsymbol{\beta}_l} \sum_{c \in \mathcal{M}_{l,k}} \prod_{m=1}^{l} \left[z_{s^{|c_m|}} \right]^{\beta_m}_{\boldsymbol{\alpha}(c_m)},
\end{aligned}
$$

where $s = \begin{bmatrix} y^* \\ u \end{bmatrix}$. This expression obviously includes also the cross-derivatives among variables included in y^*, already computed at the previous step and is only used here to keep things compact. In practical programming, one would exlude from enumeration the cross-derivatives already computed.

The system to be solved is then simply

$$
\left[\begin{array}{ccc} [f_{y^*}] + \left[f_{y^{**}_+} \right] [g^{**}_{y^*}] & \vdots & [f_{y^{**}}] \end{array} \right] \left[\begin{array}{c} \left[g^*_{s^k} \right] \\ \left[g^{**}_{s^k} \right] \end{array} \right] = [B_{s^k}].
$$

4.3 Recovering $\left[g_{y^{*k}\sigma^j} \right]$

First we put $G = g^{**}(g^*(y^*, u, \sigma), u', \sigma)$, and calculate all derivatives of order $k + j$ in $\left[G_{y^{*k}\sigma^j} \right]$. Putting

$$
w = \begin{bmatrix} g^*(y^*, u, \sigma) \\ u' \\ \sigma \end{bmatrix}
$$

we get:[8]

$$
\left[G_{y^{*k}\sigma^j} \right]_{\boldsymbol{\alpha}_k} = \sum_{l=1}^{j+k} [g^{**}_{w^l}]_{\boldsymbol{\beta}_l} \sum_{c \in \mathcal{M}_{l,j+k}} \prod_{m=1}^{l} \left[w_{y^{*k}\sigma^j(c_m)} \right]^{\beta_m}_{\boldsymbol{\alpha}(c_m)}
$$

The terms involving $k + j$ order derivative of g will occur only for $l = 1$ and $l = k + j$. For $l = 1$ we get $\left[g^{**}_{y^*} \right]_\beta \left[g^*_{y^{*k}\sigma^j} \right]^\beta$. For $l = k + j$ we get

$$
\sum_{n=0}^{l} \left[g^{**}_{y^{*n}\sigma^{l-n}} \right]_{\boldsymbol{\beta}_n} \sum_{c \in \mathcal{M}_{l,l}} \prod_{m=1}^{n} \left[g^*_{y^{*k}\sigma^j(c_m)} \right]^{\beta_m} \prod_{m=n+1}^{l} \left[\sigma_{y^{*k}\sigma^j(c_m)} \right]
$$

[8] Here, $w_{y^{*k}\sigma^j(c_m)}$ is a derivative of w with respect to variables selected by partition of indices c_m from variables $y^{*k}\sigma^j$. Also, selection of $\boldsymbol{\alpha}(c_m)$ ignores indices in c_m behind k since they correspond to univariate σ.

Now let us note that for $n < k$, there will be at least one term $[\sigma_{y*}] = 0$ in the product, and for $n > k$, there will be at least one term $[g^*_\sigma] = 0$ in the product. Thus, the only remaining term of the sum is:

$$\left[g^{**}_{y^{*k}\sigma^l}\right]_{\boldsymbol{\beta}_k} \prod_{m=1}^{k} \left[g^*_{y^*}\right]^{\beta_m}$$

To complete the task, we first put

$$\begin{aligned}
[D_{kj}]_{\boldsymbol{\alpha}_k} &= \left[F_{y^{*k}u'^j}\right]_{\boldsymbol{\alpha}_k\boldsymbol{\beta}_j} [\Sigma]^{\boldsymbol{\beta}_j} \\
[E_{kj}]_{\boldsymbol{\alpha}_k} &= \sum_{m=1}^{j-1} \binom{j}{m} \left[F_{y^{*k}u'^m\sigma^{j-m}}\right]_{\boldsymbol{\alpha}_k\boldsymbol{\beta}_m} [\Sigma]^{\boldsymbol{\beta}_m}
\end{aligned}$$

and from the constraint

$$\left[F_{y^{*k}\sigma^j}\right]_{\boldsymbol{\alpha}_k} + [D_{kj}]_{\boldsymbol{\alpha}_k} + [E_{kj}]_{\boldsymbol{\alpha}_k} = 0$$

we obtain

$$\begin{aligned}
&[f_y]_\beta \left[g_{y^{*k}\sigma^j}\right]^\beta_{\boldsymbol{\alpha}_k} + \left[f_{y^{**}_+}\right]_\beta \left[G_{y^{*k}\sigma^j}\right]^\beta_{\boldsymbol{\alpha}_k} = \\
&-\sum_{l=2}^{j+k} [f_{z^l}]_{\boldsymbol{\beta}_l} \sum_{c\in\mathcal{M}_{l,j+k}} \prod_{m=1}^{l} \left[z_{y^{*k}\sigma^j(c_m)}\right]^{\beta_m}_{\boldsymbol{\alpha}(c_m)} - [D_{kj}]_{\boldsymbol{\alpha}_k} - [E_{kj}]_{\boldsymbol{\alpha}_k}
\end{aligned}$$

Note that the only term of order $k+j$ in D_{kj} is $g^{**}_{y^{*k}u^j}$ and the only terms of order $k+j$ in E_{kj} are $g^{**}_{y^{*k}u^i\sigma^{j-i}}$ for $i = 1, \ldots, j-1$. So, after substituing for $G_{y^{*k}\sigma^j}$ we get $g_{y^{*k}\sigma^j}$ as a solution of Sylvester equation:

$$\begin{aligned}
&\left[[f_{y^*}] + \left[f_{y^{**}_+}\right]_\gamma \left[g^{**}_{y^*}\right]^\gamma\right]_\beta \left[g^*_{y^{*k}\sigma^j}\right]^\beta_{\boldsymbol{\alpha}_k} + [f_{y^{**}}]_\beta \left[g^{**}_{y^{*k}\sigma^j}\right]^\beta_{\boldsymbol{\alpha}_k} + \\
&\left[f_{y^{**}+}\right]_\beta \left[g^{**}_{y^{*k}\sigma^j}\right]^\beta_{\boldsymbol{\alpha}_k} \prod_{m=1}^{k} \left[g^*_{y^*}\right]^{\beta_m}_{\alpha_m} = \mathrm{RHS}
\end{aligned}$$

where

$$\begin{aligned}
\mathrm{RHS} \;=\; &-\sum_{l=2}^{j+k} [f_{z^l}]_{\boldsymbol{\beta}_l} \sum_{c\in\mathcal{M}_{l,j+k}} \prod_{m=1}^{l} \left[z_{y^{*k}\sigma^j(c_m)}\right]^{\beta_m}_{\boldsymbol{\alpha}(c_m)} \\
&- \left[f_{y^{**}_+}\right]_\gamma \left[\sum_{l=1}^{j+k-1} \sum_{n=0}^{l} \left[g^{**}_{y^{*n}\sigma^{l-n}}\right]_{\boldsymbol{\beta}_n} \sum_{c\in\mathcal{M}_{n,j+k+n-l}} \prod_{m=1}^{n} \left[g^*_{y^{*k}\sigma^j(c_m)}\right]^{\beta_m}_{\boldsymbol{\alpha}(c_m)}\right]^\gamma \\
&- [D_{kj}]_{\boldsymbol{\alpha}_k} - [E_{kj}]_{\boldsymbol{\alpha}_k}
\end{aligned}$$

4.4 Recovering $\left[g_{y^{*k}u^i\sigma^j}\right]$

As in the previous section, we first calculate $k+i+j$ order terms in $\left[G_{y^{*k}u^i\sigma^j}\right]$. In the very simmilar way, we get the sum

$$\left[g^{**}_{y^*}\right]_\gamma \left[g^{*}_{y^{*k}u^i\sigma^j}\right]^\gamma +$$
$$\left[g^{**}_{y^{*k+i}\sigma^j}\right]_{\boldsymbol{\gamma}_{k+i}} \prod_{m=1}^{k} \left[g^*_{y^*}\right]^{\gamma_m} \prod_{m=k+1}^{k+j} \left[g^*_u\right]^{\gamma_m}.$$

Now we put

$$\begin{aligned}
\left[D_{kij}\right]_{\boldsymbol{\alpha}_k\boldsymbol{\beta}_i} &= \left[F_{y^{*k}u^iu'^j}\right]_{\boldsymbol{\alpha}_k\boldsymbol{\beta}_i\boldsymbol{\gamma}_j} \left[\Sigma\right]^{\boldsymbol{\gamma}_j}\\
\left[E_{kij}\right]_{\boldsymbol{\alpha}_k\boldsymbol{\beta}_i} &= \sum_{m=1}^{j-1} \binom{j}{m} \left[F_{y^{*k}u^iu'^m\sigma^{j-m}}\right]_{\boldsymbol{\alpha}_k\boldsymbol{\beta}_i\boldsymbol{\gamma}_m} \left[\Sigma\right]^{\boldsymbol{\gamma}_m}
\end{aligned}$$

and from the constraint

$$\left[F_{y^{*k}u^i\sigma^j}\right]_{\boldsymbol{\alpha}_k\boldsymbol{\beta}_i} + \left[D_{kij}\right]_{\boldsymbol{\alpha}_k\boldsymbol{\beta}_i} + \left[E_{kij}\right]_{\boldsymbol{\alpha}_k\boldsymbol{\beta}_i} = 0$$

we easily solve for $\left[g_{y^{*k}u^i\sigma^j}\right]$

$$\left[\left[f_{y^*}\right] + \left[f_{y^{**}_+}\right]_\delta \left[g^{**}_{y^*}\right]^\delta\right]_\gamma \left[g^*_{y^{*k}u^i\sigma^j}\right]^\gamma_{\boldsymbol{\alpha}_k\boldsymbol{\beta}_i} + \left[f_{y^{**}}\right]_\gamma \left[g^{**}_{y^{*k}u^i\sigma^j}\right]^\gamma_{\boldsymbol{\alpha}_k\boldsymbol{\beta}_i} = \mathrm{RHS}$$

where

$$\begin{aligned}
\mathrm{RHS} = &-\sum_{l=2}^{N} \left[f_{z^l}\right]_{\boldsymbol{\gamma}_l} \sum_{c\in\mathcal{M}_{l,N}} \prod_{m=1}^{l} \left[z_{y^{*k}u^i\sigma^j(c_m)}\right]^{\gamma_m}_{\boldsymbol{\alpha}_k\boldsymbol{\beta}_i}\\
&-\left[f_{y^{**}_+}\right]_\delta \left[\sum_{l=2}^{N} \left[g^{**}_{y^{*l}}\right]_{\boldsymbol{\gamma}_l} \sum_{c\in\mathcal{M}_{l,N}} \prod_{m=1}^{l} \left[g^*_{y^{*k}u^i\sigma^j(c_m)}\right]^{\gamma_m}_{\boldsymbol{\alpha}\boldsymbol{\beta}(c_m)}\right]^\delta\\
&-\left[f_{y^{**}_+}\right]_\delta \left[\sum_{l=1}^{N}\sum_{n=1}^{l-1} \left[g^{**}_{y^{*n}\sigma^{l-n}}\right]_{\boldsymbol{\gamma}_n} \sum_{c\in\mathcal{M}_{n,N+n-l}} \prod_{m=1}^{n} \left[g^*_{y^{*k}u^i\sigma^j(c_m)}\right]^{\gamma_m}_{\boldsymbol{\alpha}\boldsymbol{\beta}(c_m)}\right]^\delta\\
&-\left[D_{kij}\right]_{\boldsymbol{\alpha}_k\boldsymbol{\beta}_i} - \left[E_{kij}\right]_{\boldsymbol{\alpha}_k\boldsymbol{\beta}_i},
\end{aligned}$$

and $N=k+i+j$. Note that the righthand side involves the derivative $g^{**}_{y^{*k+i}\sigma^j}$. Also, note that the only term of order N in D_{kij} is $g^{**}_{y^k_* u^{i+j}}$, and the only terms of order N in E_{kij} are $g^{**}_{y^k_* u^{i+m}\sigma^{j-m}}$ for $m=1,\ldots,j-1$.

4.5 Recovering $[g_{\sigma^k}]$

The derivatives with respect to the scale parameter σ are retrieved from the constraint

$$\left[F^i_{\sigma^k}\right] + \left[F^i_{u'^k}\right]_{\boldsymbol{\beta}_k} [\Sigma]^{\boldsymbol{\beta}_k} + \sum_{j=1}^{k-1} \binom{k}{j} \left[F^i_{u'^j\sigma^{k-j}}\right]_{\boldsymbol{\beta}_j} [\Sigma]^{\boldsymbol{\beta}_j} = 0.$$

Developing $\left[F^i_{\sigma^k}\right]$ and rearranging terms gives

$$\left[f^i_z\right]_\beta \begin{bmatrix} 0 \\ 0 \\ [g_{\sigma^k}] \\ \left[g^{**}_{y^*}\right]_\gamma \left[g^*_{\sigma^k}\right]^\gamma + \left[g^{**}_{\sigma^k}\right] \end{bmatrix}^\beta = B$$

with

$$\begin{aligned} B \quad = \quad & -\left[f^i_{y^{**}}\right]_\beta \left[\sum_{l=2}^{k} \left[g^{**}_{y^{*l}}\right]_{\boldsymbol{\gamma}_l} \sum_{c\in\mathcal{M}_{l,k}} \prod_{m=1}^{l} \left[g^*_{\sigma^{|c_m|}}\right]^{\gamma_m} \right]^\beta \\ & -\sum_{l=2}^{k} \left[f^i_{z^l}\right]_{\boldsymbol{\beta}_l} \sum_{c\in\mathcal{M}_{l,k}} \prod_{m=1}^{l} \left[z_{\sigma^{|c_m|}}\right]^{\beta_m} - \left[F^i_{u'^k}\right]_{\boldsymbol{\beta}_k} [\Sigma]^{\boldsymbol{\beta}_k} \\ & -\sum_{j=1}^{k-1} \left[F^i_{u'^j\sigma^{k-j}}\right]_{\boldsymbol{\beta}_j} [\Sigma]^{\boldsymbol{\beta}_j} \end{aligned}$$

The partial derivatives with respect to future shocks, $\left[F^i_{u'^k}\right]$, can in turn be obtained by

$$[F_{u'^k}]_{\boldsymbol{\beta}_k} = \sum_{l=1}^{k} \left[f_{y^{**l}_+}\right]_{\boldsymbol{\gamma}_l} \sum_{c\in\mathcal{M}_{l,k}} \prod_{m=1}^{l} \left[g^{**}_{u^{|(c_m)|}}\right]^{\gamma_m}_{\boldsymbol{\beta}_{c_m}}$$

However, the term $[F_{u'^j\sigma^{k-j}}]$ is not so easy, since

$$\begin{aligned} & [F_{u'^j\sigma^{k-j}}]_{\boldsymbol{\alpha}_j} = \\ & \sum_{l=1}^{k} \left[f_{y^{**l}_+}\right]_{\boldsymbol{\beta}_l} \sum_{c\in\mathcal{M}_{l,k}} \prod_{m=1}^{l} \left[G_{u'^j\sigma^{(k-j)(c_m)}}\right]^{\beta_m}_{\boldsymbol{\alpha}(c_m)} + \\ & \sum_{l=1}^{k} \sum_{n=1}^{l} \left[f_{y^n y^{**(n-l)}_+}\right]_{\boldsymbol{\beta}_n\boldsymbol{\gamma}_{n-l}} \sum_{c\in\mathcal{M}_{l,k}} \prod_{m=1}^{n} \left[g_{u'^j\sigma^{k-j}(c_m)}\right]^{\beta_m}_{\boldsymbol{\alpha}(c_m)} \prod_{m=1}^{l-n} \left[G_{u'^j\sigma^j(c_m)}\right]^{\gamma_m}_{\boldsymbol{\alpha}(c_m)} \end{aligned}$$

The sum in the last term cannot be further simplified since the derivative $\left[g_{u'^j\sigma^{k-j}(c_m)}\right]$ is nonzero only if all indices in c_m fall behind j and are not isolated. This condition on c_m is too irregular and it is better to calculate $[F_{u'^j\sigma^{k-j}}]$ directly. The same holds for derivatives of the form $\left[F_{y^k u^i \sigma^j}\right]$.

4.6 Order of calculations

Let us suppose, that all the derivatives of g and G are calculated for orders less than N. Here we describe the order of calculations of derivatives of g for order equal to N.

The calculation begins with $g_{y^{*N}}$ since the righthand side of its equation depends only on lower order derivatives. Then all the derivatives with respect to $y^k u^i$ for $k + i = N$ can be calculated since their righthand sides depend only on $g_{y^{*N}}$.

Now we can retrieve $g_{y^{N-1}\sigma^1}$, since its righthand side depends only on $g_{y^{*N}}$, and $g_{y^{*N-i}u^i}$. Therefore we let j go from 1 to $N-1$. In each cycle, before we can retrieve $g_{y^{*N-j}\sigma^j}$, we must retrieve all $g_{y^{*N-j}u^i\sigma^{j-i}}$ for all $i = 1, \ldots, j-1$ since these are needed by $g_{y^{*N-j}\sigma^j}$. However, there is a dependency among these terms, the terms with lower i depend on those with higher i, so the calculations are done with decreasing i.

After the outer loop is done, we have to recover all $g_{u^i\sigma^{N-i}}$ for $i = N - 1, \ldots, 1$, and then g_{σ^N}. This corresponds to the loop body for $j = N$.