Algorithmic Differentiation

Xin (Keira) Shu

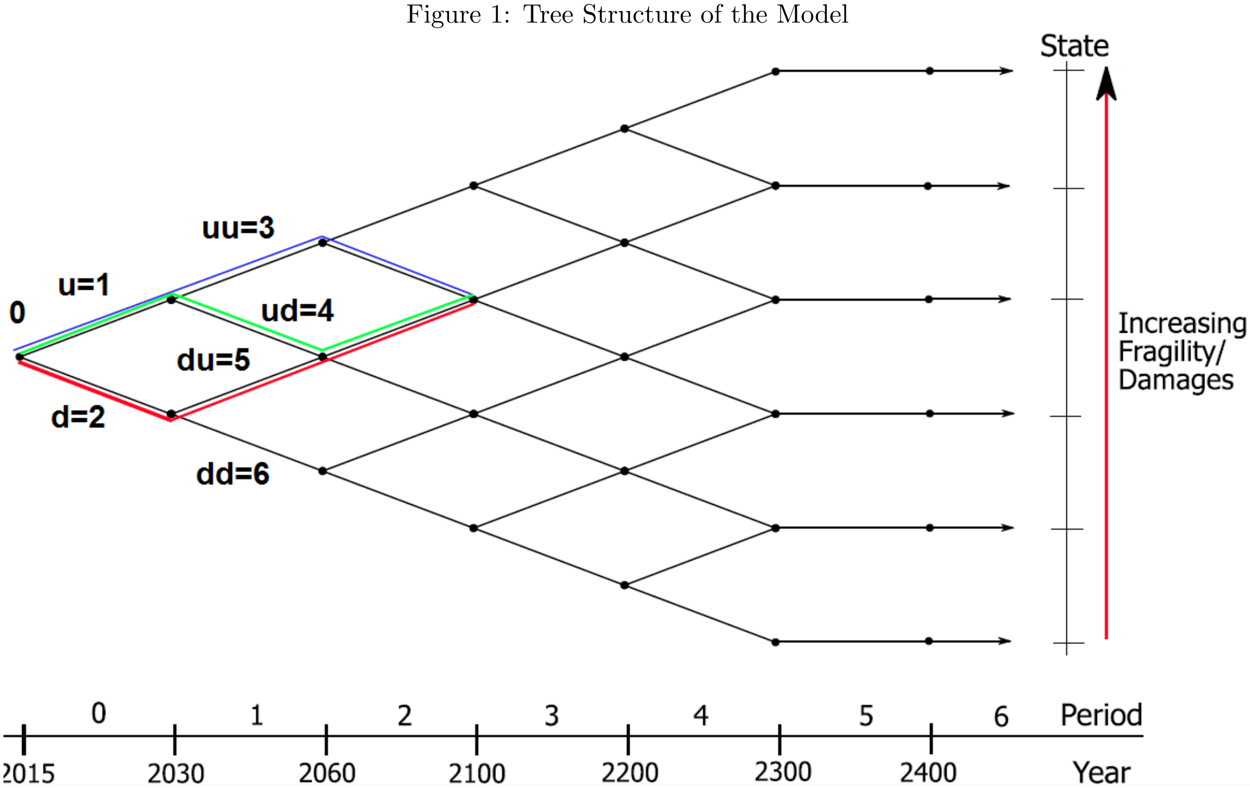
June 2017

1 Introduction

In order to apply Algorithmic Differentiation(AD) in our project, we try to put our essential problem into the (1.1) form. In our model, we have a representative agent who solves the optimization problem of trading off the (known) costs of climate mitigation against the uncertain future benefits associated with mitigation. She maximizes lifetime utility at each time and for each state of nature by choosing the optimal path of mitigation, $x_t^*(\theta_t)$, dependent on Earth's fragility θ_t.

To make the solution tractable, the EZ-Climate model employs a binomial tree for the resolution of uncertainty about climate change. Our baseline analysis uses a 7-period tree, beginning 2015. An initial mitigation decision is made in 2015, and subsequent mitigation decisions are made after information is revealed about climate fragility and the resulting damages in years 2030, 2060, 2100, 2200, and 2300 (See Figure 1). The final period, in which consumption simply grows at a constant rate, begins in 2400 and lasts forever.

Figure 1: Tree Structure of the Model



Basically, we look into fractional-mitigation levels $\mathbf{m}$ for 63 points, and we need to evaluate the scalar-valued function $f(\mathbf{m}) : \mathbb{R}^{63} \to \mathbb{R}$, which gives us the expected utility at the beginning of the very first period. Notice that $\mathbf{m}$ has entries x_t's that show up in the formulas in the rest of this article. As we learn from the paper, for an arbitrary period, its utility has to do with current consumption as well as expected utility of the periods to follow. The utility formulas are given below:

$$U_0 = \left[(1-\beta)c_0^\rho + \beta(E_0[U_1^\alpha])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \tag{1}$$

$$U_1 = \left[(1-\beta)c_1^\rho + \beta(E_1[U_2^\alpha])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \tag{2}$$

$$U_2 = \left[(1-\beta)c_2^\rho + \beta(E_2[U_3^\alpha])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \tag{3}$$

$$U_3 = \left[(1-\beta)c_3^\rho + \beta(E_3[U_4^\alpha])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \tag{4}$$

$$U_4 = \left[(1-\beta)c_4^\rho + \beta(E_4[U_5^\alpha])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \tag{5}$$

$$U_5 = \left[(1-\beta)c_5^\rho + \beta(E_5[U_6^\alpha])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \tag{6}$$

$$U_6 = \left[\frac{1-\beta}{1-\beta(1+g)^\rho}\right]^{\frac{1}{\rho}} c_6 \tag{7}$$

where $E_t[U_{t+1}^\alpha]$ is the certainty-equivalent of future lifetime utility, based on the agent's information at time t. Notice that we compute consumption c_t from mitigation levels $\mathbf{m}$, so the system can be represented as

$$U_0(\mathbf{m}) = \left[(1-\beta)c_0^\rho(\mathbf{m}) + \beta(E_0[U_1^\alpha(\mathbf{m})])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \tag{8}$$

$$U_1(\mathbf{m}) = \left[(1-\beta)c_1^\rho(\mathbf{m}) + \beta(E_1[U_2^\alpha(\mathbf{m})])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \tag{9}$$

$$U_2(\mathbf{m}) = \left[(1-\beta)c_2^\rho(\mathbf{m}) + \beta(E_2[U_3^\alpha(\mathbf{m})])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \tag{10}$$

$$U_3(\mathbf{m}) = \left[(1-\beta)c_3^\rho(\mathbf{m}) + \beta(E_3[U_4^\alpha(\mathbf{m})])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \tag{11}$$

$$U_4(\mathbf{m}) = \left[(1-\beta)c_4^\rho(\mathbf{m}) + \beta(E_4[U_5^\alpha(\mathbf{m})])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \tag{12}$$

$$U_5(\mathbf{m}) = \left[(1-\beta)c_5^\rho(\mathbf{m}) + \beta(E_5[U_6^\alpha(\mathbf{m})])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \tag{13}$$

$$U_6(\mathbf{m}) = \left[\frac{1-\beta}{1-\beta(1+g)^\rho}\right]^{\frac{1}{\rho}} c_6(\mathbf{m}) \tag{14}$$

In terms of the consumption model, in each period $t \in \{0, 1, 2, 3, 4, 5, 6\}$, the agent is endowed with a certain amount of the consumption good, $\bar{c_t}$. However, the agent is not able to consume the full endowed consumption for 2 reasons: climate change and climate policy. First, in period $t \in \{1, 2, 3, 4, 5, 6\}$, some of the endowed consumption may be lost due to climate change damages. Second, in periods $t \in \{0, 1, 2, 3, 4, 5\}$, the agent may elect to spend some of the endowed consumption to reduce his impact on the climate. The resulting consumption c_t, after damages and mitigation costs are taken into account, is given by:

$$c_0 = \bar{c_0} \cdot (1 - \kappa_0(x_0)) \tag{15}$$

$$c_1 = \bar{c_1} \cdot (1 - D_1(CRF_1, \theta_1) - \kappa_1(x_1)) \tag{16}$$

$$c_2 = \bar{c_2} \cdot (1 - D_2(CRF_2, \theta_2) - \kappa_2(x_2)) \tag{17}$$

$$c_3 = \bar{c_3} \cdot (1 - D_3(CRF_3, \theta_3) - \kappa_3(x_3)) \tag{18}$$

$$c_4 = \bar{c_4} \cdot (1 - D_4(CRF_4, \theta_4) - \kappa_4(x_4)) \tag{19}$$

$$c_5 = \bar{c_5} \cdot (1 - D_5(CRF_5, \theta_5) - \kappa_5(x_5)) \tag{20}$$

$$c_6 = \bar{c_6} \cdot (1 - D_6(CRF_6, \theta_6)) \tag{21}$$

1.1 Damage Function $D_t(CRF_t, \theta_t)$

Here, the climate change function $D_t(CRF_t, \theta_t)$ captures the fraction of the endowment consumption that is lost due to damages from climate change. This is handled by **damage.py**. D_t depends on 2 variables: CRF_t, which we define as the cumulative radiative forcing up to time t, which determines global average temperature, and θ_t, the Earth's fragility, a parameter that characterizes the uncertain relationship between the global average temperature and consumption damages.

The cumulative radiative forcing, CRF_t, in turn, depends on the level of mitigation in each period from 0 to t. The cumulative mitigation X_t is given by:

$$X_t = \frac{\sum_{s=0}^{t} g_s \cdot x_s}{\sum_{s=0}^{t} g_s} \tag{22}$$

where g_s is the flows of GHG emissions into the atmosphere in period s, for each period up to t, absent any mitigation. We manage the non-mitigation case as the base case and model it in **bau.py** for reference. Cumulative mitigation enters the determination of the rate of technology change, where we have constants θ_0 and θ_1. The level of mitigation at any time s is given by x_s.

To specify the damage function, we define it as a function of temperature changes, which, in turn, are a function of cumulative solar radiative forcing CRF_t, which, in our setting, are determined by the mitigation path up to that point in time. We then compare CRF_t to three baseline emissions paths, g_t, for which we have created associated damage simulations. The only way, then, to affect the level of damages is to change mitigation across time, x_t. The specification of damages has 2 components: a non-catastrophic component and an additional catastrophic component triggered by crossing a particular threshold. (The hazard rate associated with hitting that threshold increases with temperature.) If the threshold is crossed at any time, additional changes decreases consumption in all future periods.

We calculate the overall damage function $D_t(CRF_t, \theta_t)$ for the baseline emission paths, g_t, using Monte-Carlo simulation. We run a set of simulations for each of 3 constant mitigation levels X_t, which determine cumulative radiative forcing at each point in time. In each run of the simulation, we draw a set of random variables:

- global average temperature change.
- the parameter characterizing non-catastrophic damages as a function of temperature.
- an indicator variable that determines whether or not the atmosphere hits a tipping point at any particular time and state.
- the tipping point damage parameter.

The state variable θ_t indexes the distribution resulting from these sets of simulations, and interpolation across the three mitigation level gives us a continuous function D_t across cumulative radiative forcing levels CRF_t.

Therefore, based on the simulation results, we can determine D_t as a function of CRF_t and θ_t, where CRF_t is a function of the mitigation paths g_t of x_t.

1.1.1 CRF_t as a function of x_t

We have mitigation level x_t and the GHG emission under business-as-usual case, absent any mitigation, BAU_t for $t \in \{0, 1, 2, 3, 4, 5, 6\}$. We look into each period by unit called **subinterval**. For each subinterval we calculate **forcing** and **absorption**, with which we get cumulative forcings CRF_t and the ghg levels GHG_t. Their values in the very beginning are:

$$\begin{cases} CRF_0 & = 4.926 \\ GHG_0 & = 400 \\ sink_0 & = 35.396 \end{cases} \tag{23}$$

Here is how we calculate the variables for each subinterval:

- For each period t:

$$\begin{cases} \text{beginning emission} & g_{t,0} = (1 - x_t) \times BAU_t \\ \text{ending emission} & g_{t,t} = \begin{cases} (1 - x_t) \times BAU_{t+1}, & \text{if } t < 5 \\ (1 - x_t) \times BAU_t, & \text{else} \end{cases} \end{cases} \tag{24}$$

- For each subinterval i:

$$\begin{cases} \text{p\_co2}_i & = 0.71 \times \left[g_{t,0} + i \times \frac{g_{t,t} - g_{t,0}}{\text{number of subintervals in period } t} \right] \\ \text{p\_c}_i & = \frac{\text{p\_co2}_i}{3.67} \\ \text{add\_p\_ppm}_i & = \text{length of subinterval} \times \frac{\text{p\_c}_i}{2.13} \\ \text{lsc}_i & = 285.6268 + \text{cum\_sink}_i \times 0.88414 \\ \text{absorption} & = 0.5 \times 0.94835 \times |GHG_i - \text{lsc}_i|^{0.741547} \\ \text{cum\_sink}_i & = \text{cum\_sink}_{i-1} + \text{absorption}_i \\ GHG_i & = GHG_{i-1} + \text{add\_p\_ppm}_i - \text{absorption}_i \\ CRF_i & = CRF_{i-1} + 0.13183 \times |GHG_i - 315.3785|^{0.607773} \end{cases}$$

1.1.2 D_t as a function of CRF_t

In order to get the desired damage D_t, we run damage simulations for the 3 base scenarios and determine the damage coefficients in concern for each. Based on the CRF_i's and GHG_i's we get in the last section, we calculate the realized average mitigation $\bar{x}_t$ up to each period and compare with the constant mitigation levels in the base cases to interpolate the damage along the given path.

The 3 base scenarios correspond to maximum final GHG levels of 450, 650, and 1000 ppm. We calculate their cumulative radiative forcings CRF_t and constant mitigation levels, or average mitigation levels, $x_{450/650/1000} = \bar{x}_{450/650/1000}$, for each period. The way we infer the average mitigation from cumulative forcing up to a particular point is:

$$\bar{x}(CRF_t) = w_{450}(CRF_t) \times \bar{x}_{450} + w_{650}(CRF_t) \times \bar{x}_{650} \tag{25}$$

where the weights are given by:

- If $CRF_t > CRF_{650}$, $\begin{cases} w_{650} & = \frac{CRF_{1000} - CRF_t}{CRF_{1000} - CRF_{650}} \\ w_{450} & = 0 \end{cases}$

- If $CRF_{650} > CRF_t > CRF_{450}$, $\begin{cases} w_{650} & = \frac{CRF_t - CRF_{450}}{CRF_{650} - CRF_{450}} \\ w_{450} & = \frac{CRF_{650} - CRF_t}{CRF_{650} - CRF_{450}} \end{cases}$

- Otherwise, $\begin{cases} w_{650} & = 0 \\ w_{450} & = 1 + \frac{CRF_{450} - CRF_t}{CRF_{450}} \end{cases}$

We calculate an interpolated damage function using our 3 simulations where we have damage coefficients (for a given state and period) to find a smooth function that gives damages for any particular level of radiative forcing up to each point in time. For realized mitigation below that of 650 ppm, let the damage coefficients for state n at period t be $d_{t,n}^{650,1}$ and $d_{t,n}^{650,0}$. Similarly, for realized mitigation between those of 650 and 450 ppm, let the damage coefficients for state n at period t be $d_{t,n}^{450,2}$, $d_{t,n}^{450,1}$, and $d_{t,n}^{450,0}$. To find the function, we assume a linear interpolation of damages between the 650 and 1000 ppm scenarios, and a quadratic interpolation between 450 and 650 ppm. In addition, we impose a smooth pasting condition at 650 ppm, having the level and derivative of the interpolation between 650 ppm match the level and scope of the line above.

For consequent GHG level below 450 ppm, the realized mitigation is more than 100% and we assume climate damages exponentially decay toward 0. Mathematically, we let $S = \frac{d \cdot p}{l \cdot ln(0.5)}$, where d is the derivative of the quadratic damage interpolation function at 450 ppm, $p = 0.91667$ is the average mitigation in the

450 ppm simulation, and the level of damage is l. Radiative forcing at any point below 450 ppm then is y percent below that of the 450 ppm simulation, with $y = \frac{CRF_{450}-CRF}{CRF_{450}}$, where R is the radiative forcing in the 450 ppm simulation and r is the radiative forcing given the mitigation policy $\mathbf{m}$. Letting $\sigma = 60$, the extension of the damage function for $y > 0$ is defined as:

$$Damage(y) = l \cdot 0.5^{(y \cdot S)} e^{-\frac{(y \cdot p)^2}{\sigma}} \tag{26}$$

Therefore, given the cumulative forcing CRF_6 for node n up to period 6, we calculate the realized average mitigation $\bar{x}(CRF_6)$ and calculate its damage in the following way:

$$D_6(CRF_6, \theta_6) = \begin{cases} d_{6,n}^{650,1} \times \bar{x}(CRF_6) + d_{6,n}^{650,0}, & \text{for } \bar{x}(CRF_6) < \bar{x}_{650}; \\ d_{6,n}^{450,2} \times \bar{x}(CRF_6)^2 + d_{6,n}^{450,1} \times \bar{x}(CRF_6) + d_{6,n}^{450,0}, & \text{for } \bar{x}_{450} < \bar{x}(CRF_6) \leq \bar{x}_{650}; \\ 0.5^{\frac{\bar{x}(CRF_6)-\bar{x}_{450}+ln(d_{6,n}^{450})}{ln(0.5)}} e^{-\frac{(\bar{x}(CRF_6)-\bar{x}_{450})^2}{60}}, & \text{else} \end{cases} \tag{27}$$

The $d_{6,n}^{450}$ in the last case is the simulated damage for state n under the 450 ppm scenario.

The climate sensitivity- summarized by state of nature θ_t- is not known prior to the final period (t=6). Rather, what the representative agent knowns is the distribution of possible final states θ_6. We specify that the damage in period t, given a cumulative radiative forcing, CRF_t, up to time t, is the probabiity weighted average of the interpolated damage function over all final states of nature reachable from that node. In particular, teh damage function at time t, for the node indexed by θ_t is assumed to be:

$$D_t(CRF_t, \theta_t) = \sum_{\theta_6} \mathbb{P}(\theta_6|\theta_t) \cdot D_t(CRF_t, \theta_6) \tag{28}$$

In addition, we have a penalty function for the damages of the concentrations below pre-industrial levels:

$$f(GHG_t) = \left[1 + e^{0.05 \times (GHG_t - 200)}\right]^{-1} \tag{29}$$

Hence, the damage function at time t is:

$$D_t(CRF_t, \theta_t) = \sum_{\theta_6} \mathbb{P}(\theta_6|\theta_t) \cdot D_t(CRF_t, \theta_6) + \left[1 + e^{0.05 \times (GHG_t - 200)}\right]^{-1} \tag{30}$$

1.2 Cost Function $\kappa_t(x_t)$

Mitigation reduces the stock of GHGs in the atmosphere and leads to lower climate damages, and, hence, to higher future consumption. But mitigating GHG emissions is costly. Mitigating a fraction of x_t of emissions costs a fraction $\kappa_t(x_t)$ of the endowed consumption. In terms of the mitigation cost, we consider both the traditional cost regarding taxation and technological components. The latter one matters for x_t above the fractional-mitigation threshold, x^*, at which the backstop technology kicks in. The consequent upper and lower bound for marginal technology costs are $\tilde{\tau}$ and τ^* respectively. From these bounds we can calculate relevant parameters b and k. The general equation for marginal cost is given by:

$$\kappa_t(x_t) = [1 - \theta_0 - \theta_1 X_t] \begin{cases} \frac{g_0}{c_0} \cdot 92.08 \cdot x_t^{\alpha}, & x_t \leq x^* \\ \frac{g_0}{c_0} \cdot 92.08 \cdot x^{*\alpha} + \tilde{\tau} x_t - \tilde{\tau} x^* - \frac{b x_t (\frac{k}{x_t})^{\frac{1}{b}}}{b-1} + \frac{b x^* (\frac{k}{x^*})^{\frac{1}{b}}}{b-1}, & x_t > x^* \end{cases} \tag{31}$$

This part is from before and you can disregard it at this moment. We would like to start with the last period and reorganize our system into:

$$\begin{cases} \text{solve for } y_1: & F_1(\mathbf{m}) - M_1(\mathbf{m})y_1 = 0 \\ \text{solve for } y_2: & F_2(\mathbf{m}, y_1) - M_2(\mathbf{m}, y_1)y_2 = 0 \\ & \dots \\ \text{solve for } y_T: & F_T(\mathbf{m}, y_1, \dots, y_{T-1}) - M_T(\mathbf{m}, y_1, \dots, y_{T-1})y_T = 0 \\ \text{solve for } z: & \bar{f}(\mathbf{m}, y_1, \dots, y_T) - z = 0 \end{cases}$$

Essentially, we hope we can find appropriate maps that help us realize:

$$\left\{ \begin{array}{ccc} F_1(\mathbf{m}) = c_6(\mathbf{m}) & \mathrm{M}_1(\mathbf{m}) = \left[\frac{1-\beta}{1-\beta(1+g)^\rho}\right]^\rho & \mathrm{y}_1 = U_6(\mathbf{m}) \\ F_2(\mathbf{m}, y_1) = \left[(1-\beta)c_5^\rho(\mathbf{m}) + \beta(E_5[U_6^\alpha(\mathbf{m})])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} & \mathrm{M}_2(\mathbf{m}, y_1) = \mathbf{1} & \mathrm{y}_2 = U_5(\mathbf{m}, y_1) \\ F_3(\mathbf{m}, y_1, y_2) = \left[(1-\beta)c_4^\rho(\mathbf{m}) + \beta(E_4[U_5^\alpha(\mathbf{m})])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} & \mathrm{M}_3(\mathbf{m}, y_1, y_2) = \mathbf{1} & \mathrm{y}_3 = U_4(\mathbf{m}, y_1, y_2) \\ F_4(\mathbf{m}, y_1 \dots y_3) = \left[(1-\beta)c_3^\rho(\mathbf{m}) + \beta(E_3[U_4^\alpha(\mathbf{m})])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} & \mathrm{M}_4(\mathbf{m}, y_1 \dots y_3) = \mathbf{1} & \mathrm{y}_4 = U_3(\mathbf{m}, y_1 \dots y_3) \\ F_5(\mathbf{m}, y_1 \dots y_4) = \left[(1-\beta)c_2^\rho(\mathbf{m}) + \beta(E_2[U_3^\alpha(\mathbf{m})])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} & \mathrm{M}_5(\mathbf{m}, y_1 \dots y_4) = \mathbf{1} & \mathrm{y}_5 = U_2(\mathbf{m}, y_1 \dots y_4) \\ F_6(\mathbf{m}, y_1 \dots y_5) = \left[(1-\beta)c_1^\rho(\mathbf{m}) + \beta(E_1[U_2^\alpha(\mathbf{m})])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} & \mathrm{M}_6(\mathbf{m}, y_1 \dots y_5) = \mathbf{1} & \mathrm{y}_6 = U_1(\mathbf{m}, y_1 \dots y_5) \\ F_7(\mathbf{m}, y_1 \dots y_6) = \left[(1-\beta)c_0^\rho(\mathbf{m}) + \beta(E_0[U_1^\alpha(\mathbf{m})])^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} & \mathrm{M}_7(\mathbf{m}, y_1 \dots y_6) = \mathbf{1} & \mathrm{y}_7 = U_0(\mathbf{m}, y_1 \dots y_6) \\ \bar{f}(\mathbf{m}, y_1, \dots y_7) = y_7 & & \end{array} \right\}$$

Let $\mathbf{m}$ be a 1×32 vector. If we write the equations above into their matrix form, the first equation should look like:

$$\begin{cases} F_1(\mathbf{m}) & = c_6(\mathbf{m}) = \mathbf{A_1 m} + \mathbf{A_2} \\ M_1(\mathbf{m}) & = \left[\frac{1-\beta}{1-\beta(1+g)^\rho}\right]^\rho \\ y_1 & = U_6(\mathbf{m}) = \mathbf{A'_1 m} + \mathbf{A'_2}, \text{ where } \mathbf{A'_i} = \left[\frac{1-\beta}{1-\beta(1+g)^\rho}\right]^\rho \mathbf{A_i} \end{cases}$$

For F_2, we have $c_5(\mathbf{m}) = \mathbf{B_{5,1} m_5^2} + \mathbf{B_{5,2} m_5} + \mathbf{B_{5,3}}$ and $F_2(\mathbf{m_5}, y_1) = [(1-\beta)c_5^\rho(\mathbf{m}) + \beta y_1^\rho]^{\frac{1}{\rho}}$, so y_2 should also be of dimension 1×32. Here the $\mathbf{m_5}$ is simply $\mathbf{m}$.
In what to follow, we look into F_i in $F_3, F_4 \dots$ up to F_7 and let $j = 7 - i$. We have decreasing size of y_i's as they represent utilities of earlier stages where we have fewer decision nodes. y_i has the same dimension as $\mathbf{m_j}$ and we need to get $\mathbf{m_j}$ from $\mathbf{m}$. In general, we need the following components to get F_i:

- Get $\mathbf{m_j}$ from $\mathbf{m}$ via $\mathbf{m_j} = \mathbf{\Omega_j m}$.
- Consumption $c_j(\mathbf{m}) = \mathbf{B_{j,1} m_j^2} + \mathbf{B_{j,2} m_j} + \mathbf{B_{j,3}}$.
- The expectation of future utility in the equation should be given by

$$E_j[U_{j+1}^\alpha(\mathbf{m})] = \frac{1}{2} \times (y_{i-1,1}^\alpha + y_{i-1,2}^\alpha)$$

  where $y_{i-1,1}, y_{i-1,2}$ have the even- and odd-indexed entries of y_i respectively. Hence, y_i should have half the number of y_{i-1}'s entries and is of dimension $2^{6-i} \times 1$.