These are some simple notes about how to use automatic differentiation to solve an equilibrium model with market power or optimal policy.

The key problem is that naive autodiff is only able to compute partial derivatives, whereas a granular agent's FOC involves the total derivative, to their action, of an objective function that is implicitly defined in terms of atomistic agents' FOCs and market clearing conditions.

1 Static Case

In the following, almost everything will be implicitly vector-valued. Thus, e.g., G will represent a vector of market clearing conditions, A a vector of household actions,[1]. $\frac{\partial G}{\partial A}$ and $\frac{dG}{dA}$ are the Jacobian of partial and total derivatives, respectively, of G with respect to the components of A.

1.1 Decomposition of Monopolist's or Policymaker's FOC

I outline a case where a granular agent makes a decision that is "quantity-like" in the sense that it enters directly into the equilibrium (i.e. market-clearing) conditions, and only indirectly enters into the household's budget constraint. This arises, for instance, in Cournot competition. This will look slightly different in the case where the granular agent's decison is "price-like", as in Bertrand competition, entering directly into the budget constraint and only indirectly into the

1.1.1 Environment

Suppose you have a model with atomistic agent types $i \in I$, granular agents (monopolists or policymakers) $\ell \in L$, and equilibrium variables (typically prices) p_k indexed by $k \in K$. [2] Prices are defined by a vector of market clearing conditions,

$$0 = G\left(A, P, Q\right)$$

where $A \equiv \left\{a_i\right\}_i$ are the actions of atomistic agents, $P \equiv \left\{p_k\right\}_k$ are equilibrium variables, and $Q \equiv \left\{q_\ell\right\}_\ell$ are the actions of the granular agents.

Agents i differ by beginning-of-period state x_i and choose actions a_i maximize utility subject to a budget constraint H,

$$\begin{aligned} a_i &= \arg\max_a u\left(x_i, a \mid \Gamma, P\right) \\ 0 &= H_i \equiv H\left(x_i, a \mid \Gamma, P\right). \end{aligned}$$

Granular agents ℓ choose action q_ℓ to maximize an objective Π_ℓ,

[1] If each household type takes more than one action, then A will be a block vector, with one block per household type, and within each block, one entry per choice variable.

[2] Thanks to Lukas Mann for talking this through with me. The format of this first part is cribbed from his job market paper, though I believe it is fairly standard.

$$
\begin{aligned}
\Pi_\ell &= \Pi_\ell\left(\Gamma, A, P, Q\right) \\
Q_\ell &= \max_{Q_\ell} H_\ell\left(\Gamma, A, P, \{Q_1, \ldots, Q_\ell, \ldots, Q_L\}\right).
\end{aligned}
$$

1.1.2 Granular Agent's Problem

The granular agent ℓ internalizes the response of atomistic agents and equilibrium variables to their action q_ℓ.[3]

That is, their FOC is,

$$\frac{d\Pi_\ell}{dQ_\ell} = 0$$

where the response of A and P are internalized, so that the FOC becomes,

$$0 = d\Pi_\ell = \frac{\partial \Pi_\ell}{\partial A} dA + \frac{\partial \Pi_\ell}{\partial P} dP + \frac{\partial \Pi_\ell}{\partial q_\ell} dQ_\ell \tag{1}$$

1.2 Total Differentiation

To get this in terms of partial derivatives, first totally differentiate the market clearing conditions,

$$
\begin{aligned}
0 &= G\left(A, P, Q\right) \\
&= \frac{\partial G}{\partial A} dA + \frac{\partial G}{\partial P} dP + \frac{\partial G}{\partial Q_\ell} dQ_\ell.
\end{aligned} \tag{2}
$$

Next, take the atomistic agent's FOC,[4]

$$
\begin{aligned}
0 &= \frac{\partial u'}{\partial a_i} da_i \\
0 &= \frac{\partial H_i}{\partial a_i} da_i + \frac{\partial H_i}{\partial P} dP + \frac{\partial H_i}{\partial Q_\ell} dQ_\ell.
\end{aligned}
$$

Stacking $\frac{\partial H_i}{\partial a_i}$ and $\frac{\partial u'}{\partial a_i}$ into a single invertible matrix $\frac{\partial \widetilde{H}_i}{\partial a_i}$ so that, setting things up so that $\frac{\partial u}{\partial P} = 0$,

$$
\begin{aligned}
0 &= \frac{\partial \widetilde{H}_i}{\partial a_i} da_i + \frac{\partial \widetilde{H}_i}{\partial P} dP + \frac{\partial \widetilde{H}_i}{\partial Q_\ell} dQ_\ell \\
da_i &= -\left(\frac{\partial \widetilde{H}_i}{\partial a_i}\right)^{-1} \left(\frac{\partial \widetilde{H}_i}{\partial P} dP + \frac{\partial \widetilde{H}_i}{\partial Q_\ell} dQ_\ell\right).
\end{aligned}
$$

[3] However, they do not internalize the response of other granular agents, as this leads to suffering. I mutter the words, "Nash equilibrium," I wave my hands...

[4] I denote the FOC by u', so that $\frac{\partial u'}{\partial a_i}$ is the *second* partial derivative.

Stacking these gives,[5]

$$dA = -\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1}\left(\frac{\partial \widetilde{H}}{\partial P}dP + \frac{\partial \widetilde{H}}{\partial Q_\ell}dQ_\ell\right).$$

That is, the total effect of the changes in P and Q_ℓ on the household Lagrangian FOC must be exactly countered by the change in A.

Plugging into (2),

$$\begin{aligned}
0 &= -\frac{\partial G}{\partial A}\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1}\frac{\partial \widetilde{H}}{\partial P}dP - \frac{\partial G}{\partial A}\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1}\frac{\partial \widetilde{H}}{\partial Q_\ell}dQ_\ell + \frac{\partial G}{\partial P}dP + \frac{\partial G}{\partial Q_\ell}dQ_\ell \\
&= \left(\frac{\partial G}{\partial A}\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1}\frac{\partial \widetilde{H}}{\partial P} - \frac{\partial G}{\partial P}\right)dP + \left(\frac{\partial G}{\partial A}\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1}\frac{\partial \widetilde{H}}{\partial Q_\ell} - \frac{\partial G}{\partial Q_\ell}\right)dQ_\ell \\
dP &= -\left(\frac{\partial G}{\partial A}\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1}\frac{\partial \widetilde{H}}{\partial P} - \frac{\partial G}{\partial P}\right)^{-1}\left(\frac{\partial G}{\partial A}\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1}\frac{\partial \widetilde{H}}{\partial Q_\ell} - \frac{\partial G}{\partial Q_\ell}\right)dQ_\ell.
\end{aligned} \tag{3}$$

That is, in GE, the GE effect of a change in Q_ℓ on the equilibrium conditions G must be zero. Thus, the sum of the direct effect, $\frac{\partial G}{\partial X}dX$, and indirect effect, $-\frac{\partial G}{\partial A}\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1}\frac{\partial \widetilde{H}}{\partial X}dX$, for $X = Q_\ell$ must be exactly counteracted by the sum of effects for $X = P$.

Finally, plugging back into 1,

$$\begin{aligned}
0 &= \frac{\partial \Pi_\ell}{\partial A}dA + \frac{\partial \Pi_\ell}{\partial P}dP + \frac{\partial \Pi_\ell}{\partial Q_\ell}dQ_\ell \\
\text{Where}\quad dA &= -\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1}\left(\frac{\partial \widetilde{H}}{\partial P}dP + \frac{\partial \widetilde{H}}{\partial Q_\ell}dQ_\ell\right) \\
dP &= -\left(\frac{\partial G}{\partial A}\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1}\frac{\partial \widetilde{H}}{\partial P} - \frac{\partial G}{\partial P}\right)^{-1}\left(\frac{\partial G}{\partial A}\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1}\frac{\partial \widetilde{H}}{\partial Q_\ell} - \frac{\partial G}{\partial Q_\ell}\right)dQ_\ell
\end{aligned} \tag{4}$$

Equation (4) is the granular agent's FOC.

1.2.1 Interpretation

This looks complicated. This is largely because it is so general, encompassing any well-behaved static model consisting of atomistic agents, granular agents,

[5]It might seem like we are blowing up the size of $\frac{\partial \widetilde{H}}{\partial A}$ by making it one big matrix, when we know it is block-diagonal. However, Julia allows us to specify that a matrix is block-diagonal, then acts on it as efficiently as it would on the individual matrices.

equilibrium variables, and market-clearing conditions. The above expression nests many special cases quite beautifully, and it may be instructive to see some such special cases.

Bertrand Competition Suppose that prices do not affect the market clearing conditions directly and granular firms choose some prices. (Prices chosen by firms are in Q.) Then,

$$\frac{\partial G}{\partial P} = \frac{\partial G}{\partial Q_\ell} = 0,$$

so that Equation (3) becomes,[6]

$$\begin{aligned}
0 &= \frac{\partial \Pi_\ell}{\partial A} dA + \frac{\partial \Pi_\ell}{\partial P} dP + \frac{\partial \Pi_\ell}{\partial Q_\ell} dQ_\ell \\
\text{Where} \quad dA &= -\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1} \left(\frac{\partial \widetilde{H}}{\partial P} dP + \frac{\partial \widetilde{H}}{\partial Q_\ell} dQ_\ell\right) \\
dP &= -\left(\frac{\partial G}{\partial A}\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1} \frac{\partial \widetilde{H}}{\partial P}\right)^{-1} \left(\frac{\partial G}{\partial A}\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1} \frac{\partial \widetilde{H}}{\partial Q_\ell}\right) dQ_\ell
\end{aligned}$$

Cournot Competition Suppose that firms choose quantities, which enter into market clearing conditions directly but not into the household budget constraint. Also assume that prices P do not enter into the market clearing conditions directly. Then,

$$\frac{\partial G}{\partial P} = \frac{\partial \widetilde{H}}{\partial Q_\ell} = 0,$$

so that Equation (3) becomes,

$$\begin{aligned}
0 &= \frac{\partial \Pi_\ell}{\partial A} dA + \frac{\partial \Pi_\ell}{\partial P} dP + \frac{\partial \Pi_\ell}{\partial Q_\ell} dQ_\ell \\
\text{Where} \quad dA &= -\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1} \left(\frac{\partial \widetilde{H}}{\partial P} dP + \frac{\partial \widetilde{H}}{\partial Q_\ell} dQ_\ell\right) \\
dP &= -\frac{\partial G}{\partial A}\left(\frac{\partial \widetilde{H}}{\partial A}\right)^{-1} \frac{\partial \widetilde{H}}{\partial P}^{-1} \frac{\partial G}{\partial Q_\ell} dQ_\ell.
\end{aligned}$$

1.3 Automatic Differentiation

Now that we have the granular agent's FOC in terms of automatically computable partial derivatives, we can compute it automatically. Thus, the full procedure is:

[6] The last line below looks suspiciously like the regression projection $(X'X)^{-1}X'y$. If you can see the deep connection, please let me know!

1. Categorize the equations in the model into the categories G, H, Π, u.
2. Use automatic differentiation routines to compute the partial derivative Jacobians.
   (a) If using Julia, remember to type your matrices appropriately in order to take advantage of sparsity.
3. Plug the Jacobians into (4) to obtain the FOC.

We can even use gradient-based solvers such as (L-)BFGS by automatically differentiating the FOC! Since it is already in terms of auto-diff-able partial derivatives, higher order derivatives are easy.