Regional Research Institute Technical Documents

Regional Research Institute

3-12-2014

Generating MATLAB Matrices from IMPLAN CGE Files

Mason Pierce
mason.pierce@mail.wvu.edu

Regional Research Institute
Technical Document Series

Generating MATLAB Matrices from IMPLAN CGE Files

Mason Pierce

RRI TechDoc2013-05

Date Submitted: 3/12/2014

Key Words/Codes: IO, IMPLAN

Generating MATLAB Matrices from IMPLAN CGE Files

Abstract

This document describes the process used to create usable tables from the 27 individual GAMS files in .DAT format generated by IMPLAN software. 26 of these are social accounting matrix (SAM) data files and 1 is an employment-by-industry data file. By performing the appropriate command in MATLAB, the user can create the desired table, for example the use or make tables, which can then be used for various purposes.

IMPLAN 2 and IMPLAN 3 produce .DAT files of different formats. This document details the process of handling each of these file formats.

Introduction

IMPLAN social accounting reports give you the option to export the 26 individual detailed reports that together create the single GAMS file plus the employment-by-industry file. The 26 individual files are also the elements of the SAM. These files include interactions between industries, commodities, institutions, foreign trade, etc. However, the output format is not easy to work with. IMPLAN 3 entries take the form of:

```
Number.Number     Value
```

IMPLAN 2 data files take the form of:

```
Number          .          Number     Value
```

In both cases, the two numbers delimited by a decimal on the left identify the commodity/industry/institution/etc. interacting with another commodity/industry/institution/etc. The first number represents the rows of a final-form table, and the second number constitutes the columns. The interaction between them is the value.

For example, the 2x1.dat file, which is 1 of the 27 files IMPLAN will produce, is commodity by industry; it is the “domestic use of commodities by industries or payments to commodities.” The IMPLAN 3 file will contain entries similar to this:

```
3001.20     0.45602
```

This means that the value of industry 20’s use of commodity 1 is 0.45602.

In order to manipulate these types of entries into usable tables, code was written to read in the .DAT file from whichever workpath is designated, parse everything correctly, and put it into a workable matrix with 0's where there is no observation (for example, if industry 1 did not use any inputs or make any payments to commodity 2, there will not be an entry in the file and it will have a 0 in the corresponding spot in the produced matrix).

The code for each of the 27 files (the employment-by-industry file is handled in the same fashion as the SAM elements) was also augmented to make each matrix begin observations (usually) in the first row and first column. For example, commodity 3001 in the above example is actually commodity 1, just numbered to make it identifiable as a commodity instead of an industry. This also means that institutions in the produced matrices, for example, will have numbers from 1-18, as there are 18 total institutions, instead of the IMPLAN numbering scheme for institutions of 10001-14002. The user must just know which table they are commanding the program to make, thereby identifying what the rows and columns of the produced matrix will be; then, the numbers of rows and columns will come as no surprise.

This is easily augmentable to suit different preferences, however.

To generate the matrices, first shorten the titles of the files to be simply "1x2.dat", "5x4.dat", etc. if they are not already. IMPLAN 2 may have a prefix in the title of the file; delete it. Then, place all of the data files that you wish to use into your MATLAB workpath and follow the instructions below.

Supporting Algorithm(s)/Code

IMPLAN 3:

```
% PURPOSE: Generate usable matrices depicting input-output relationships
% between sectors using IMPLAN .dat outputs.
%--------------------------------------------------------------------------
% USAGE: The matrix generated begin with S; this is any one of 27 workable
% matrices generated from IMPLAN. The notation after the S denotes which
% matrix it is.
% outputs.
% INPUT:
% -> 'file.dat' corresponds to the file that is being turned into a
% workable matrix from the IMPLAN .dat output file.
% OUTPUT:
% -> S_x_ is the final usable matrix form of each of the 27 IMPLAN .dat
% files.
%--------------------------------------------------------------------------
% REFERENCES: None.
%--------------------------------------------------------------------------
% Written by: Mason Pierce, 3/12/2014
% Current email: mapierce@mix.wvu.edu
%--------------------------------------------------------------------------
% General form of parsing and forming a workable file from .dat files:
fid=fopen('file.dat');
M=textscan(fid,'%s %s');
fclose(fid);
```

```
A=M{1};
B=M{2};
E=str2double(B);
out=regexp(A,'\.','split');
out=str2double(reshape([out{:}],2,[])');
C=out(:,1);
D=out(:,2);
U=[C,D,E];
clear A
clear B
clear C
clear D
clear E
clear M
clear ans
clear fid
clear out

%--------------------------------------------------------------------------
% Taking this file output and turning it into a usable matrix.
% Note: The above code was general to all of the .dat files IMPLAN 3
% returns. The below code, executed after the above code, is specific
% to each individual .dat file. The below code is specific to the 4x5
% .dat file, but is included as a demonstration of the general form
% the codes take. Generally, this portion is used to subtract the
% arbitrary sectoring numbers in order to create a compact and
% maximally populated matrix. For example, the institutions (of which
% there are only ever 18) in this file are converted from an IMPLAN
% numbering scheme of 10001-14002 to a numbering scheme of 1-18.
% The 4x5 file is institutions by foreign trade; it is the foreign
% institutional export table. The 18 institutions have numbers
% subtracted from their IMPLAN numbering scheme as described above,
% and the foreign trade observations have 25000 subtracted from their
% numbers to make their number 1 instead of 25001. This makes all
% entries appear in a single column, the first one created in the
% matrix.
% This changes with each .dat file and is easily reconfigurable.
%--------------------------------------------------------------------------
A=U;
clear U
N=size(A,1);
for i=1:N
    if A(i,1)<=10999
        A(i,1)=A(i,1)-10000;
    elseif A(i,1)<=11999
        A(i,1)=A(i,1)-10991;
    elseif A(i,1)<=12999
        A(i,1)=A(i,1)-11988;
    elseif A(i,1)<=13999
        A(i,1)=A(i,1)-12985;
    elseif A(i,1)<=14999
        A(i,1)=A(i,1)-13984;
    else
        A(i,1)=A(i,1);
    end
end
for i=1:N
```

```
 A(i,2)=A(i,2)-25000;
end
B=A;
A=zeros(18,1);
for i=1:N
    A(B(i,1),B(i,2))=B(i,3);
end
S4x5=A;
clear A
clear B
clear N
clear i;


%--------------------------------------------------------------------------
```

IMPLAN 2:

```
% PURPOSE: Generate usable matrices depicting input-output relationships
% between sectors using IMPLAN .dat outputs.
%--------------------------------------------------------------------------
% USAGE: The matrix generated begin with S; this is any one of 27 workable
% matrices generated from IMPLAN. The notation after the S denotes which
% matrix it is.
% outputs.
% INPUT:
% -> 'file.dat' corresponds to the file that is being turned into a
% workable matrix from the IMPLAN .dat output file.
% OUTPUT:
% -> S_x_ is the final usable matrix form of each of the 27 IMPLAN .dat
% files.
%--------------------------------------------------------------------------
% REFERENCES: None.
%--------------------------------------------------------------------------
% Written by: Mason Pierce, 3/12/2014
% Current email: mapierce@mix.wvu.edu
%--------------------------------------------------------------------------
% General form of parsing and forming a workable file from .dat files:
fid=fopen('file.dat');
M=textscan(fid,'%s %s %s %s');
fclose(fid);
X=M{1};
Y=M{3};
Z=M{4};
X=str2double(X);
Y=str2double(Y);
Z=str2double(Z);
A=[X,Y,Z];
%--------------------------------------------------------------------------
% Taking this file output and turning it into a usable matrix.
% Note: The above code was general to all of the .dat files IMPLAN 2
% returns. The below code, executed after the above code, is specific
% to each individual .dat file. The below code is specific to the 3x1
% .dat file, but is included as a demonstration of the general form
% the codes take. Generally, this portion is used to subtract the
% arbitrary sectoring numbers in order to create a compact and
% maximally populated matrix. For example, the factors (of which
% there are only ever 4) in this file are converted from an IMPLAN
% numbering scheme of 5001-8001 to a numbering scheme of 1-4.
% The 3x1 file is factors by industry; it is the factor income/
% value-added elements. The 4 institutions have numbers
% subtracted from their IMPLAN numbering scheme as described above,
% but the industries are already numbered beginning with 1; their
% numbering scheme needs no manipulation.
% This changes with each .dat file and is easily reconfigurable.
%----------------------------------------------------------------------------
N=size(A,1);
for i=1:N
    if A(i,1)<=5999
        A(i,1)=A(i,1)-5000;
    elseif A(i,1)<=6999
        A(i,1)=A(i,1)-5999;
    elseif A(i,1)<=7999
        A(i,1)=A(i,1)-6998;
    else
        A(i,1)=A(i,1)-7997;
    end
end
B=A;
A=zeros(4,440);
for i=1:N
    A(B(i,1),B(i,2))=B(i,3);
end
S3x1=A;
clear A
clear M
clear X
clear Y
clear Z
clear ans
clear fid
clear B
clear N
clear i;
```