Visualizing United State's Traffic Flow

Tharald S. Fongaard & Jacob Perricone

Advisor: Professor Alain Kornhauser

Submitted in partial fulfillment

of the requirements for the degree of

Bachelor of Science in Engineering

Department of Operations Research and Financial Engineering

Princeton University

12 May 2015

We hereby declare that we are the sole authors of this thesis.

We authorize Princeton University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

Jacob Perricone

Tharald S. Fongaard

We further authorize Princeton University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Jacob Perricone

Tharald S. Fongaard

Abstract

By analyzing trip generation in the United States, we will display key statistics in an informative graphic user interface. Specifically, given a node network of the entire nation and trip generation files, we will provide a means to determine the number of vehicles, people, and the average vehicle occupancy on each link at each time throughout the day.

Acknowledgements

Here are the acknowledgments. I'm sure you know what to do here. . . thanks, and thanks!

To The Hereditary Kingdom of Norway

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This is the introduction of my sample thesis. Let me just state that this is a very, very limited introduction to LaTeX, and I can not do justice to it through the next few pages.

I recommend reading this .pdf file along with the .tex file open on the side so that you can compare the pure ASCII text and the final result after compilation.

1.1 A Different Word Processing System

I highly recommend the booklet "The Not So Short Introduction to LaTeX 2_ε" which is free and very well-written. It is available on:

`http://tobi.oetiker.ch/lshort/lshort.pdf`.

I refer you to the introduction there.

Chapter 2

Development of the Department of Operations Research and Financial Engineering

The first step of the visualization required that the United States be broken up into pixels so that trips could be easily identified and tracked. Our analysis considers the 48 contiguous states, which together span around 3.8 million square miles. Each square pixel spans an area of .25 miles, implying that around 12 million pixels are needed to uniquely map the entire country. The center of the National grid is at 37°N Latitude, -97.50°W Longitude, placing the grid's centroid close to the geographic center of the continental US along the time meridian.

2.1 Pixel Structure and Algorithm Method

The primary motivation for creating a pixelated grid is to establish a simple method to track, store and retrieve trips within the pixel. Thus each pixel object must contain a unique id, an (i,j) coordinate, and the latitude and longitude of its four corners and its centroid.

2.1.1 Geometric Analysis

In order to algorithmically create a pixelated grid, a methodology of mapping latitude and longitude to (i,j) coordinates was established. Given that along a Meridian each degree of latitude traverses approximately 69.174 miles and that along a Parallel each degree of longitude traverses approximately $69.174 \cdot \cos(Latitude)$ one can map any latitude and longitude coordinate to an (i,j) pixel. To determine the set of (lon, lat) coordinates within a pixel, the angular height and width of each pixel as calculated. The Y-height of a pixel is 0.00722814 degrees Latitude and the X-width of a pixel is 0.00944344 degrees Latitude. Each (i,j) Pixel thus includes all points within:

$$\mathit{YHeight} = 0.00722814 \mid \mathit{CenterLat} = 97.5 \mid \mathit{CenterLon} = 37.0 \tag{2.1}$$

$$\frac{CenterLat + YHeight \cdot i}{\cos(CenterLon + YHeight \cdot j)} \leq \text{Longitude} < \frac{CenterLat + YHeight \cdot (i+1)}{\cos(CenterLon + YHeight \cdot (j+1))} \tag{2.2}$$

$$CenterLat + YHeight \cdot j \leq \text{Latitude} < CenterLat + YHeight \cdot (j+1) \tag{2.3}$$

The inverse of the equation above was used to map (lon, lat) coordinates to pixel coordinates and was imperative in constraining our algorithm to the US.

$$\mathit{xPixel} = i = floor(138.348 \cdot .sllongitude + CenterLat \cdot \cos(latitude)) \tag{2.4}$$

$$\mathit{yPixel} = j = floor(138.348 \cdot .sllatitude - CenterLon) \tag{2.5}$$

Where floor casts the decimal result as an integer, rounded down, and 138.348 is a correcting constant.

2.1.2 Methodoloy

The southwest and northeast corner of the US provided the initial and final (lon, lat) coordinates of our grid. The initial point (southwest corner) has a longitude of -120.5°and a latitude of -63.5°. The final point (northeast corner) has a longitude of -63.5°and a latitude of 48.9°. Using equation 2.4 and 2.5, the initial and final (i,j) coordinates corresponding to the corners of the nation were found and used as starting and breaking conditions during pixel construction. Then, running through a double for loop indexed by i and j, each (i,j) pair was sent to a Pixel class for creation and then added to a hashtable. The Pixel class takes an (i,j) coordinate in its constructor and using the equations in 2.1.1 determines the longitude and latitude of its four corners and centroid. Furthermore, in order to prevent collisions within the hashtable, each pixel creates a unique id by using its (i,j) coordinates as the input for the Cantor pairing function, which uniquely encodes two natural numbers into a single natural number. Since the Cantor pairing function only holds for positive numbers, i and j are offset by the absolute value of the initial (i,j) point.

$$f_{ij} = (1/2)\cdot(i+|(InitialX+InitialY)|+j)\cdot(i+|(InitialX+InitialY)|+j+1)+j+|(InitialY)| \tag{Cantor Function}$$

A sketch of the initialization loop is shown below.

```
for i in range(iStart, iEnd):
    for j in range(jStart, jEnd):
            Pixel = createPixel(i,j)
            Pixels.add(Pixel.key(), Pixel)
```

2.1.3 Additional Structure: The Use of Subsections

We are in a subsection now (two levels down from a Chapter). When we refer to *X.Y.Z*, we mean Chapter X, Section Y, and subsection Z. We declare a Chapter by the command `\chapter{`*title of Chapter*`}`, a Section by `\section{`*title of Section*`}`, and so on.

The nice thing about LaTeX is that it takes care of the chapter, section, and subsection numbering automatically. If I were to add another subsection before this one the subsection number would change (increment by one). This section is 2.1.1 and I referred to it using the command `\ref{`*label of this section*`}`. I inserted a label right after the `\subsection` declaration by typing `\label{`*label of this section*`}`.

A subsubsection

Just for fun! Notice that no number is alloted for such a low level environment but it sometimes useful.

2.1.4 Another Subsection

And so on. . . .

2.2 Mathematical Symbols

Let $X = \{X_n, n \in \mathbb{N}\}$ be a Markov chain with state space $\mathcal{D}$. Throughout this thesis, we use the notation

$$p_{ij} := \mathbb{P}\{X_{n+1} = j \mid X_n = i\}, \quad i, j \in \mathcal{D} \tag{2.6}$$

for the transition probabilities of the Markov chain X. Furthermore, we denote by P the transition matrix, $P = [p_{ij}]_{i,j\in\mathcal{D}}$.

When we wrote (2.1) we implicitly assumed that the Markov chain X is time-homogeneous.

Let us also define Y,

$$Y = (Y_n)_{n=0,1,2,\ldots}$$

to be another process. Notice that the second equation does not take a number on the right—this is the use of `\begin{equation*}` environment.

Notice that the all the math characters, X, $\mathcal{D}$, and others such as α, β, γ are part of the text in LaTeX. On the contrary, Word includes such characters as foreign objects (usually images), which increases the size of the document file, sometimes makes them disappear, but most importantly are not as aesthetically pleasing as the resulting characters here.

2.3 Citing and Bibliography

When working with large documents you need an easy way to cite your references without having to go back to your list all the time to remember the names of the authors and the year of publication. Even more importantly, you need to have all your references listed in the end of the document in alphabetical order. Of course, they all need to be syntactically the same so that alone makes the manual entry of references a big pain. Thankfully, LaTeX takes care of that in a very easy and elegant way, using BibTeX.

I cite here a few books, papers, and technical reports, and please go to page 15 to see the resulting bibliography.

According to the books by Çinlar (1975), Bielecki and Rutkowski (2002), and Musiela and Rutkowski (1997) and the articles by Duffie and Gârleanu (2001), Blanco et al. (2005), and Cotton et al. (2004) we conclude absolutely nothing. However, in his report, Aas (2004) claims that otherwise. All these citations were entered by

`\cite{`*citation label*`}`.

Notice the different citation style that follows: it is parenthetical, and observe that only one pair of parentheses is required (see Theorem 5.2 Antonov et al., 2005, pg. 32). This citation is entered by typing `\cite[see Theorem 5.2][pg. 32]{AMM05}` in the `.tex` file. (Here, the citation label corresponding to Antonov et al. (2005) is obviously `AMM05`.)

The citations are included in the file `refs.bib` under the folder `Bibliography`. You can modify it and make your own references. I highly recommend using *JabRef* for managing your bibliography entries, because it makes it a piece of cake to do a lot of *dirty* work. *JabRef* is free and it works as a Java Application.

Also notice that LaTeX, by default, includes in the Bibliography section only the references you actually cited throughout the text. If you want a source to appear in the Bibliography section without actually citing it anywhere in your text use the command `\nocite{`*citation label*`}`. For example here I type `\nocite{B95}` and you see no citation appear—however look at the fourth entry of the Bibliography. That cited book does not appear anywhere in this thesis, other than the Bibliography.

2.4 Referencing Figures and Tables

The very informative Figure 3.1 is on page 8. Both of these numbers were automatically generated—which is great when you add a new figure before the one you just inserted, because the numbering changes automatically for you. Use `\ref{fig:dens}` for the figure number and `\pageref{fig:dens}` for the page number where the figure is located. Here, `fig:dens` was the label of the figure (see actual `.tex` file for more information). Remember that LaTeX does not work like Word—the figures and tables are **not** always placed exactly where you want them, so avoid writing "according to the figure below...," and prefer writing "according to Figure [*figure number*]...," instead. The same things go unchanged for tables. Notice that when I talk about figures and tables in general, I do not need to capitalize them, however if I talk specifically about Figure 3.1 and Table 3.1, I'd better respect them and capitalize the 'f' and the 't.'

Since we're at it, notice that the quotes ‘, ’, “, ” are not inserted like in Word. For ‘ you need to use the `‘` key that is located above the `Tab` button. For ’ you just press the `’` key, exactly to the left of the `Enter` key. For double quotes just double the appropriate single quotes without leaving any space.

Chapter 3

Analysis of Problem

Time for some analysis, and probably graphs and tables. Thankfully LaTeX (and LaTeX 2$_\varepsilon$) provides very nice environments for both.

3.1 Preliminary Analysis

Assume that the random variable D given $\tilde{p} = p$ is binomially distributed with parameters 50 and p (probability of success, in this case default). We also take the cumulative distribution function of $\tilde{p}$ to be

$$F(\theta) = \mathbb{P}\{\tilde{p} \leq \theta\} = \Phi\left(\frac{1}{\rho}\left(\sqrt{1-\rho^2}\,\Phi^{-1}(\theta) - \Phi^{-1}(\bar{p})\right)\right)$$

where Φ is the cumulative standard normal distribution function, ρ is the correlation coefficient between the idiosyncratic and market factors and $\bar{p}$ is the mean default probability ($\bar{p} = \mathbb{E}\tilde{p}$). To calculate the density of $\tilde{p}$ let

$$\begin{aligned} h(\theta, \rho, \bar{p}) &:= \frac{1}{\rho}\left(\sqrt{1-\rho^2}\,\Phi^{-1}(\theta) - \Phi^{-1}(\bar{p})\right), \\ \varphi(\theta) &:= \frac{d}{d\theta}\Phi(\theta) = \frac{1}{\sqrt{2\pi}} e^{-\theta^2/2}, \end{aligned}$$

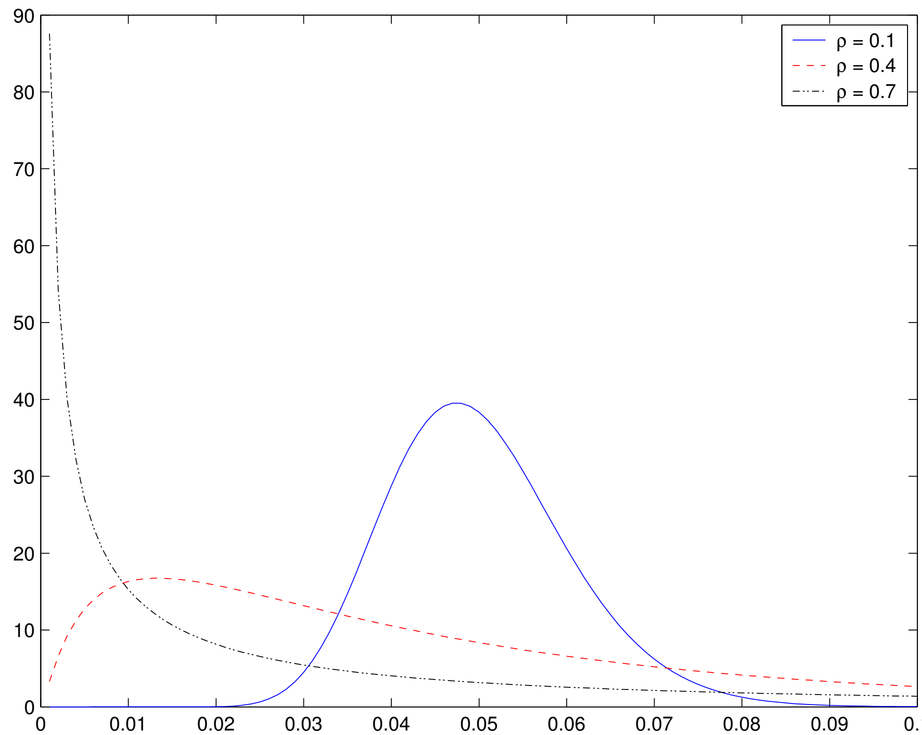

Figure 3.1: The density function f as given in (3.1) for three different ρ's and $\bar{p} = 0.05$. (Plotted on $[0, 0.1]$ for convenience.)

and notice that since Φ is a bijection we have

$$\Phi \circ \Phi^{-1}(\theta) = \Phi^{-1} \circ \Phi(\theta) = \theta,$$

for every $\theta \in \mathbb{R}$. Then, we have for the density of $\tilde{p}$,

$$\begin{aligned} f(\theta, \rho, \bar{p}) &= \frac{d}{d\theta} F(\theta) = \Phi'(h(\theta, \rho, \bar{p})) \frac{\partial}{\partial \theta} h(\theta, \rho, \bar{p}) \\ &= \varphi(h(\theta, \rho, \bar{p})) \frac{\sqrt{1-\rho^2}}{\rho} \frac{d}{d\theta} \Phi^{-1}(\theta) \\ &= \varphi(h(\theta, \rho, \bar{p})) \frac{\sqrt{1-\rho^2}}{\rho} \frac{1}{\varphi(\Phi^{-1}(\theta))}, \end{aligned} \tag{3.1}$$

for $\theta \in (0,1)$ and zero otherwise, since

$$\begin{aligned}
\frac{d}{d\theta}\left(\Phi(\Phi^{-1}(\theta))\right) &= \Phi'(\Phi^{-1}(\theta))\frac{d}{d\theta}\Phi^{-1}(\theta) \Leftrightarrow \\
\frac{d}{d\theta}\theta &= \varphi(\Phi^{-1}(\theta))\frac{d}{d\theta}\Phi^{-1}(\theta) \Leftrightarrow \\
\frac{d}{d\theta}\Phi^{-1}(\theta) &= \frac{1}{\varphi(\Phi^{-1}(\theta))}.
\end{aligned}$$

The density of $\tilde{p}$ is shown in Figure 3.1 for three different values of ρ and $\bar{p} = 0.05$. The effect of the correlation is to put more mass towards higher default probabilities as the correlation increases, thus resulting in larger number of defaults as the correlation increases.

Table 3.1: Values of the CDO Tranches.

		ρ		
		0.1	0.4	0.7
Equity	$C_E^\star(T)$	2.5712	2.8957	3.5642
Junior	$C_J^\star(T)$	9.9289	9.6137	9.2120
Senior	$C_S^\star(T)$	35.0000	34.9905	34.7239
Sum	$C_E^\star(T) + C_J^\star(T) + C_S^\star(T)$	47.5000	47.5000	47.5001

Using the default values for the parameters as above we get the values for the tranches in Table 3.1.

Let's see one more result on Table 3.2. According to Frazier (2006)...

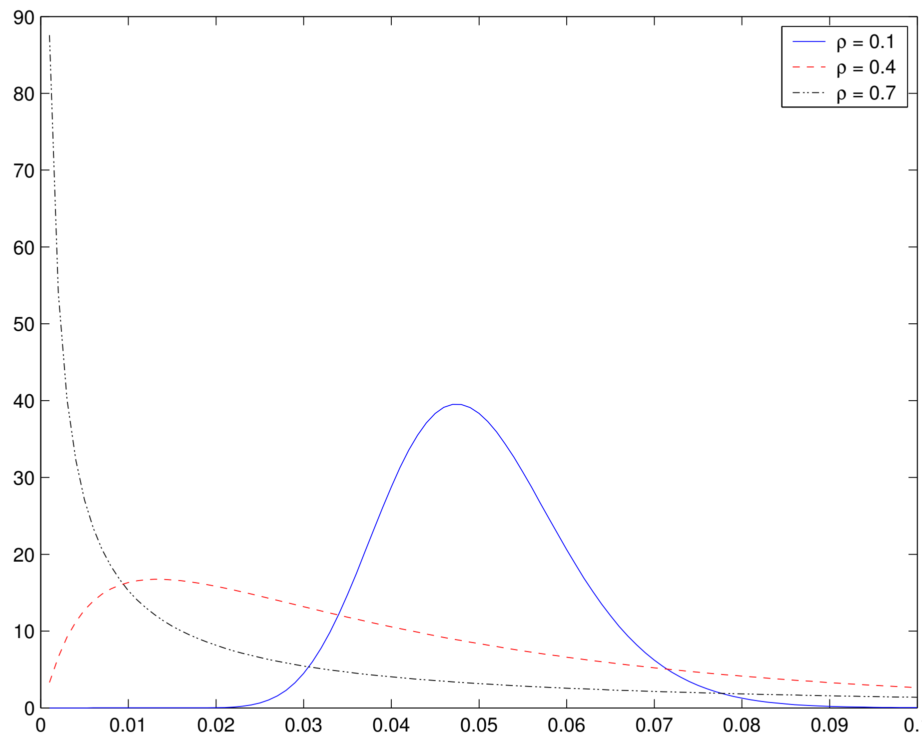

Figure 3.2: The density function f as given in (3.1) for three different ρ's and $\bar{p} = 0.05$. (Plotted on $[0, 0.1]$ for convenience.)

Table 3.2: These are the data.

DJ CDX Tranches:	0-3%	3-7%	7-10%	10-15%	15-30%	RMSE
Market mid spread	40.00%	312.5	122.5	42.5	12.5	
Bid/Ask Spread	2.00%	15.0	7	7	3	
Stochastic vol intensities	41.80%	308.9	116.2	44.9	2	1.68
Jump-diffusion intensities	46.90%	340.2	119.7	61.9	14.3	2.17
Pure-diffusion intensities	49.30%	442.9	94.9	16.8	0.4	5.34
Gaussian copula	46.80%	474.4	131.8	36.9	2.9	5.3
RFL Gaussian copula	48.60%	334.9	125.5	66.5	9.2	2.59
Double-t copula	45.10%	367	114.9	54.9	20	2.44

Source: Some source that I quote in the form Author (YYYY).

Appendix A

Code

```
% Question 1
warning('off'); % To protect our nerves
p = 0.05;
K_E = 5;
K_J = 10;
K_S = 35;

tmp1 = [];
for i = 0:(K_E-1)
     tmp1(i+1) = (K_E - i) * binopdf(i,50,p);
end
C_E = sum(tmp1)

tmp2 = [];
for i = (K_E+1):50
    tmp2(i-K_E) = (i-K_E) * binopdf(i,50,p);
end
```

```
tmp3 = [];
for i = (K_E + K_J +1):50
    tmp3(i- K_E - K_J) = (i - K_E - K_J) * binopdf(i,50,p);
end
C_J = K_J - sum(tmp2) +sum(tmp3)

C_S = K_S - sum(tmp3)

% Sanity check
C_E + C_J + C_S
```

Bibliography

Aas, K. (2004). Modelling the dependence structure of financial assets: A survey of four copulas. Technical report, Norwegian Computing Center. SAMBA/22/04.

Antonov, A., Mechkov, S., and Misirpashaev, T. (2005). Analytical techniques for synthetic cdos and credit default risk measures. Work in progress, NumeriX.

Bielecki, T. R. and Rutkowski, M. (2002). *Credit Risk: Modeling, Valuation and Hedging*. Springer.

Billingsley, P. (1995). *Probability and Measure*. Wiley, 3rd edition.

Blanco, R., Brennan, S., and Marsh, I. W. (2005). An empirical analysis of the dynamic relation between investment-grade bonds and credit default swaps. *Journal of Finance*, 60(5):2255–2281.

Çinlar, E. (1975). *Introduction to Stochastic Processes*. Prentice-Hall.

Cotton, P., Fouque, J.-P., Papanicolaou, G., and Sircar, R. (2004). Stochastic volatility corrections for interest rate derivatives. *Mathematical Finance*, 14(2):173–200.

Duffie, D. and Gârleanu, N. (2001). Risk and valuation of collateralized debt obligations. *Financial Analysts Journal*, 57(1):41–59.

Frazier, P. (2006). A nice title. *Journal of Everything*, 12(1):34–76.

Musiela, M. and Rutkowski, M. (1997). *Martingale Methods in Financial Modelling.* Springer.