

Work	Status	Due Date	Time Planned
<ul style="list-style-type: none"> <input type="checkbox"/> Getting the Time Series Data <i>Getting the data involves contacting the appropriate people and also possibly preprocessing.</i> 			
<ul style="list-style-type: none"> <input type="checkbox"/> TAQ 			
<ul style="list-style-type: none"> <input type="checkbox"/> Producing time series from TAQ: <i>We are debugging.</i> 		5/26/09	2h
<ul style="list-style-type: none"> <input type="checkbox"/> Series we will produce: 			
<ul style="list-style-type: none"> <input type="checkbox"/> Daily VWAP for actively traded stocks from 1993–2008 <i>The S&P 500 changes from time to time. I will use it as a guide. It has been noted that VWAP (which is the average we are currently producing) is more accurate than closing prices.</i> 		5/27/09	
<ul style="list-style-type: none"> <input type="checkbox"/> Hourly VWAP from 1993–2008 		5/28/09	
<ul style="list-style-type: none"> <input type="checkbox"/> Ten-minute VWAP for selected months <i>anything beyond this point and we need Kalman filtering.</i> 		5/29/09	
<ul style="list-style-type: none"> <input type="checkbox"/> High-frequency series with Kalman filtering 		6/5/09	
<ul style="list-style-type: none"> <input type="checkbox"/> Future enhancements: 			
<ul style="list-style-type: none"> <input type="checkbox"/> Incorporate dividend data from TAQ <i>I noticed recently that TAQ has dividend data available. We can use it to obtain more accurate results.</i> 			
<ul style="list-style-type: none"> <input type="checkbox"/> Incorporate Kalman filtering <i>We need to do this in order to work with any time series with missing data.</i> 		5/29/09	3h
<ul style="list-style-type: none"> <input type="checkbox"/> This will be tested with known data: 			
<ul style="list-style-type: none"> <input type="checkbox"/> 1. spring systems (2nd order linear differential equations) 			
<ul style="list-style-type: none"> <input type="checkbox"/> 2. actual market data with random points removed 			
<ul style="list-style-type: none"> <input type="checkbox"/> Google Trends 			
<ul style="list-style-type: none"> <input type="checkbox"/> I will contact them! We would like an API for access to their data. 		5/27/09	
<ul style="list-style-type: none"> <input type="checkbox"/> These people are: Eyal Cabri; Amos Fiat and his students, including Dan Feldman, from Google Israel. www.isragood.com—google-israel-improves-google-trends.html www.israel21c.org—en.jsp blog.israeltech.net—78 blog.google.org—tracking-flu-trends.html 			
<ul style="list-style-type: none"> <input type="checkbox"/> We still need to decide what data we're interested in. 			
<ul style="list-style-type: none"> <input type="checkbox"/> Economic data is definitely interesting 			
<ul style="list-style-type: none"> <input type="checkbox"/> We don't have enough data points for Google News Trends, but regular Google Trends is fine 			
<ul style="list-style-type: none"> <input type="checkbox"/> I have the list of data we discussed around here somewhere 			
<ul style="list-style-type: none"> <input type="checkbox"/> Twitter data 			
<ul style="list-style-type: none"> <input type="checkbox"/> NowPublic 		5/26/09	
<ul style="list-style-type: none"> <input type="checkbox"/> I am still uncertain as to what kernel we'll use or whether this will work, but I'll reply tonight. 			
<ul style="list-style-type: none"> <input type="checkbox"/> Other Important Economic Data <i>Google Trends actually makes these available, but it's better to go to the source.</i> 			
<ul style="list-style-type: none"> <input type="checkbox"/> the Bureau of Labor Statistics at www.bls.gov <i>They produce key economic data: the employment numbers.</i> 		6/5/09	
<ul style="list-style-type: none"> <input type="checkbox"/> Interest rates: We can use wrds.wharton.upenn.edu—bmyield <i>Interest rate data is key.</i> 		6/5/09	
<ul style="list-style-type: none"> <input type="checkbox"/> Census data at www.census.gov—estimates.html <i>I don't know whether we need this.</i> 			
<ul style="list-style-type: none"> <input type="checkbox"/> Producing Kernels from Time Series <i>The code appears to be working well.</i> 			
<ul style="list-style-type: none"> <input type="checkbox"/> Run a variety of experiments 			
<ul style="list-style-type: none"> <input type="checkbox"/> Start with: Daily VWAP, Hourly VWAP, ten-minute VWAP, 			
<ul style="list-style-type: none"> <input type="checkbox"/> Possibly introduce new kernels <i>I'll read through the relevant bits of "Learning with Kernels" by Scholkopf and Smola.</i> 			
<ul style="list-style-type: none"> <input type="checkbox"/> Analysis of Time Series 			
<ul style="list-style-type: none"> <input type="checkbox"/> We need to run a variety of experiments. 			
<ul style="list-style-type: none"> <input type="checkbox"/> The code needs to be improved substantially to automate running experiments. <i>It's currently extremely minimal.</i> 		5/28/09	3h
<ul style="list-style-type: none"> <input type="checkbox"/> For example, automatically interpret the results: Show the labels together with the coefficients. 			
<ul style="list-style-type: none"> <input type="checkbox"/> Organize the results of multiple experiments 			
<ul style="list-style-type: none"> <input type="checkbox"/> We need to adjust the step size to achieve faster convergence. <i>We can consider Barzilai-Borwein and the other methods used by Hale-Yin-Zhang. We can achieve results before doing this, but once we're scripting automated experiments we need to do this.</i> 		6/5/09	1d
<ul style="list-style-type: none"> <input type="checkbox"/> Miscellaneous 			
<ul style="list-style-type: none"> <input type="checkbox"/> We need to create conventions for output data files and directories, and also for config files. <i>We could use file extensions: .raw, .series, .kernel, .analysis, .conf</i> 			
<ul style="list-style-type: none"> <input type="checkbox"/> Programming 			
<ul style="list-style-type: none"> <input type="checkbox"/> Documentation 			
<ul style="list-style-type: none"> <input type="checkbox"/> I will update the comments to produce a javadoc API documentation. <i>Not having this available seems to be slowing me down in coding.</i> 		5/28/09	1h
<ul style="list-style-type: none"> <input type="checkbox"/> Refactoring <i>I think the refactoring I currently want to do is minor.</i> 			
<ul style="list-style-type: none"> <input type="checkbox"/> Research 			
<ul style="list-style-type: none"> <input type="checkbox"/> Start writing about what we already have <i>I will go through several drafts.</i> 			
<ul style="list-style-type: none"> <input type="checkbox"/> Sliding Window Kernel Computation 			
<ul style="list-style-type: none"> <input type="checkbox"/> Start writing about this and determining how to structure the code. 			
<ul style="list-style-type: none"> <input type="checkbox"/> We'll start with the kernels we're using. 			
<ul style="list-style-type: none"> <input type="checkbox"/> Need to do a literature search. 			
<ul style="list-style-type: none"> <input type="checkbox"/> Improve the Optimization 			
<ul style="list-style-type: none"> <input type="checkbox"/> Fixed Point Continuation 			
<ul style="list-style-type: none"> <input type="checkbox"/> Different Starting Points 			
<ul style="list-style-type: none"> <input type="checkbox"/> Improve the Analysis of the FPC algorithm's local minima 			
<ul style="list-style-type: none"> <input type="checkbox"/> Experiment with different starting points, both sparse and not sparse 			
<ul style="list-style-type: none"> <input type="checkbox"/> Adjust the step sizes 			
<ul style="list-style-type: none"> <input type="checkbox"/> Barzilai-Borwein 			
<ul style="list-style-type: none"> <input type="checkbox"/> From FPC code: method of van den Berg, E. and M. Friedlander. In pursuit of a root. Preprint, 2007 			
<ul style="list-style-type: none"> <input type="checkbox"/> Iteratively Reweighted Least Squares 			
<ul style="list-style-type: none"> <input type="checkbox"/> Talk with Sinan Gunturk 			
<ul style="list-style-type: none"> <input type="checkbox"/> References: Daubechies et al, Candes et al, Needell. 			
<ul style="list-style-type: none"> <input type="checkbox"/> Other optimization ideas 			

Work	Status	Due Date	Time Planned
<ul style="list-style-type: none"> • <input type="checkbox"/> min $Ax ^2 + u x ^2$ with $x ^2 - 1 = 0$ <i>this is homogeneous in degree. Perhaps that makes it easier to analyze. The results should be roughly the same, as if we changed the scaling factor.</i> 			
<ul style="list-style-type: none"> • <input type="checkbox"/> x _1 < c: LASSO 			
<ul style="list-style-type: none"> • <input type="checkbox"/> Ridge regression <i>This is regression regularized by the L2 norm of an affine transformation.</i> 			
<ul style="list-style-type: none"> • <input type="checkbox"/> Basis pursuit 			
<ul style="list-style-type: none"> • <input type="checkbox"/> Misc Questions 			
<ul style="list-style-type: none"> • <input type="checkbox"/> Cholesky Decomposition: Is it useful? 			
<ul style="list-style-type: none"> • <input type="checkbox"/> LLL + truncation 			
<ul style="list-style-type: none"> • <input type="checkbox"/> Experiment 			