Trading The 4th Order Least Squares Curve On IBM 5min Bars

What is Least Squares?

First let us imagine a set of closing prices on a graph with time as the horizontal axis and price as the vertical axis. Let us further suppose that we have only fifty closing price dots at fifty time intervals. How can we draw a curve through those fifty prices where the sum of all the squared vertical distances between the prices at each time interval and the curve that is being fit to the data is minimized? This can be solved by a mathematical technique called least squares curve fitting. However before we can implement this technique we have to define the formula for the curve. Here we will use a familiar formula for a curve called a polynomial. The polynomial has the form $a_0+a_1*t+a_2*t^2+a_3*t^3+\ldots.+a_n*t^n$. Where t is the price bar number. That is for the first price bar t equals 1. For the second price bar t is assign the number two and so on. When the curve formula is described in this fashion and when the net squared distance (also called the sum of the squared errors) is minimized, a unique set of coefficients $a_0,a_1,a_2,\ldots,a_n$ for the polynomial are determined.

Most traders are familiar with the simplest form of the polynomial, namely the first order polynomial also called the linear regression line of the form a_0+a_1*t . If you are fitting the first order polynomial to N prices then the coefficients can be solved for quite easily and are given by

$$a_0 = [2(2N+1)/N(N-1)]\sum_{t=1}^{N} p(t) + [6/(N(N-1)]\sum_{t=1}^{N} t * p(t)$$

$$a_1 = [12/N(N^2-1)]\sum_{t=1}^{N} t * p(t) - [6/N(N-1)]\sum_{t=1}^{N} p(t)$$

Where p(t) is the price a point t and N is the number of prices we are using to calculate the coefficients.

Using the a_0 and a_1 coefficients found above we can forecast the next bars estimated price using the formula:

$P_{\text{next day}} = a_0 + a_1*(N+1)$

Unfortunately the coefficients of the higher order polynomials do not easily reduce to simple formulas such as the first order polynomial. Matrix algebra is usually used to solve for the higher order polynomial coefficients. What is interesting though, is that using the least squares technique on a polynomial curve produces a specially weighted moving average of past prices.

Data Discussion

For this article we will use 5 minute bar prices of IBM from 7/5/2001 to 8/3/2001 and 7/11/2001 to 8/11/2001 to develop the system parameters and 5 minute bar prices from 8/6/2001 to

8/10/2001 and 8/13/2001 to 8/17/2001 for out-of-sample performance for the parameters chosen in the in-sample test segments.

The Least Squares Polynomial System.
First order polynomial curves do not respond fast enough to intraday changes in price. To develop a curve that responds faster to price changes we need to use higher polynomial orders, such as second, third and forth. During optimization we can let the computer determine what is the best polynomial order for intraday trading of our stock. For this system we will use a polynomial curves of the form $a_0+a_1*t+a_2*t^2$, $a_0+a_1*t+a_2*t^2+a_3*t^3$ and $a_0+a_1*t+a_2*t^2+a_3*t^{3+}a_4*t^4$.

If we use a sliding window of N data points, for each window we can solve for the coefficients of the polynomial using the least squares technique. The coefficients will then yield a best estimate of price at the end point N, given by the equation

$$P_{est}(N) = a_0+a_1*N+a_2*N^2+a_3*N^{3+}a_4*N^4$$

Where here we have used a fourth order polynomial here for demonstration purposes.

Since we are interested in turning points, it is best to use the next bar's forecasted price to create a leading indicator given by the equation

$$P_f(N=1) = a_0+a_1*(N+1)+a_2*(N+1)^2+a_3*(N+1)^{3+}a_4*(N+1)^4$$

Before proceeding it is instructive to observe the differences between the next bar's forecasted price using different polynomial orders. Figure 5a presents a plot of IBM 5minute bars on 8/14/01. Superimposed on this chart are the least squares polynomials forecasted next bar price of order 1, 2 and 4. The black line is the 1st order polynomial, the red line is the 2nd order polynomial and the blue line is the 4th order polynomial. Notice how the 4th order line responds much faster to changes in price many time leading the peak or bottoms in price. The 1st order line responds to slow to changes in price to be useful in intraday trading. The 2nd order line is smoother than the 4th order line but also lags the price series by 1 to 3 bars. While the 4th order line dues lead the price series, it tend to zigzag two much. This problem is solved by demanding that the 4th order line move a minimum percentage amount from a previous high or low before taking action.

We are now ready to describe the least squares curve system. We use the next day forecast price because changes in the trend are more quickly reflected in the forecast price than in the end point price.

The 4th Order Least Squares Polynomial System Defined
The least squares forecast is constructed by solving for the least squares coefficients at each 5min Bar using the last **N** bars of IBM 5minute close prices. $\mathbf{P_f}$ is then constructed from the equation above and plotted on the price chart. In general what we will be doing is following the plotted curve of $\mathbf{P_f}$. When the curve increases by a percentage amount *pctup* from the previous prior

low of the curve we will go long. When the curve falls by the percentage amount *pctdn* from the previous prior high of the curve we will go short.

Buy Rule:

- **IF P_f** has moved up by more than the percentage amount of *pctup* from the lowest low recorded in **P_f** while short then buy at the market.

Sell Rule:

- **IF P_f** has moved down by more than the percentage amount *pctdn* from the highest high recorded in **P_f** while long then sell at the market.

Exit Rule

- Close the position 5 minutes before the NYSE close (no trades will be carried overnight).

Walk Forward Optimization

Walk forward optimization will be used here because of the changing nature of intraday markets for stocks. Intraday price dynamics are constantly changing due to current economic surprises, events and trader sentiment. Also the time of year changes the nature of intraday markets, such as the seasons, holidays, vacation time, etc. As such, optimizations on intraday data performed 3 months ago may no longer be representative of today's intraday price dynamics. In walk forward testing enough data is needed to model most of the price dynamics that will be encountered in the out-of-sample segment, but not so much data that when the price dynamics start to change they are swamped by the weight of distant past data price dynamics that no longer are valid. We found that 4 weeks of 5 minute bar test data for IBM to be appropriate.

The walk forward procedure will be applied as follows. A period of 4 weeks from the start of the IBM 5 minute bar data, July 5th, 2001 through August 3rd, 2001, is chosen and system parameter values are found through optimization on this intraday data segment. The parameter values found are then applied to the out-of-sample 5 minute intraday bar data following the test segment which in this case is August 6th, 2001 to August 10th, 2001. This process is repeated by moving the test data window forward one week to July 11th , 2001 through August 10th, 2001, and again finding the parameters values through optimization on this new data test window. The parameter values found are then applied to the next out-of-sample 5 minute intraday bar data following this new test window data which in this case is August 13th, 2001 to August 17th, 2001.

Finding The System Parameters Using Walk Forward Optimization

There are four system parameters to find. *n,* the polynomial order. *N,* the number of prices used to find the least squares coefficients. *pctup,* the percent amount the forecasted price curve has to increase from a previous low to generate a buy signal. *Pctdn,* the percent amount the forecasted price curve has to drop from a previous high to generate a sell signal.. The best parameters will be defined as those values that give the best Net Profits and best Total winning Bars to Total Losing Bars ratio with the minimum drawdown and minimum largest losing trades. In addition, the results should be stable, e.g. the profits, wins, and drawdowns should not change by much as the parameters move by a small amount away from their optimum values. Also in choosing the "best" parameters, we considered only those parameters sets whose maximum consecutive losses

were 4 or less. The maximum consecutive losses are constrained to 4 or less because in real time trading it is tough to follow a system that has more than 4 losses in a row. Optimization is defined as the search for the parameter values that give the best results as defined above. It should be noted that in this stage of system development, the only thing indicated by the optimum values that are found in the test portion is that the data has been *curve fitted* as best it can with this system. Without further testing on out-of-sample data there is no way to tell if the system will work in the future.

Results

Figure 1 presents a table of the test window selected optimum parameters for the IBM 5min data series.

Start Date	End Date	n	N	pctup	pctdn
07/05/01	08/03/01	4	40	0.7	0.5
07/11/01	08/10/01	4	20	1.05	0.9

Figure 1 Optimum Parameter Values For Test Data Segments

Figures 2a and 2b presents the performance summary of the test windows using the optimum parameters for the test windows shown in Figure 1.

Figure 3 presents the combined performance summary of the two out-of-sample data segment from 8/6/01 to 8/10/01 and 8/13/01 to 8/17/01. This performance represents what would have happened in *real time* if one used the parameters found in the test sections. Slippage, and commissions are not included.

Figure 4 presents a specialized percentage trade by trade summary from 8/6/01 to 8/17//01. Note that the trades from 8/6/01 to 8/17/01 are the out-of-sample trades generated from the optimized parameters from the two test sections of 7/5/01 to 8/3/01 and 7/11/01 to 8/10/01.
The in sample trades were generated by the curve fit and are not of interest here.

Figures 5a presents a 5 minute bar chart of IBM with the 4th Order polynomial superimposed and all the buy and sell signals from the trade by trade summary of Figure 3 indicated on the charts. Also included at the bottom of the charts are the bar by bar profit or loss of each trade. The lower plot tracks the runup and drawdown of each trade.

Discussion of System Performance

As can be observed from the test sample Performance summary in Figures 2a, 2b and the out-of-sample performance summary of Figure 3, the out-of-sample performance was in line with the test sample performance. The out-of-sample section average bars in winners and losers, the drawdowns, and the profits factors were very close to the test in-sample sections This inline performance in the out-of-sample section indicates that 4 weeks of test data was enough to capture the intraday price dynamics of IBM.

Observing the out-of-sample performance summary of Figure 3, we can see that the system did equally well on the short trades and long trades. The average trade (win & loss) was $145 in the

test section and $165 in the out-of-sample section indicating stability in the parameter selection. There were no really big winners or big losers indicating steady returns.

In observing the chart we can see that the system did very well in catching every major intraday trend of IBM 5 minute bars. As can be seen, many times the 4th order polynomial curve overshoots the price curve because the curve is plotting the next bar's estimated value based upon the previous 20 bars. However, as we can observe from the charts, while this overshoot is disconcerting, it is not a problem for there is very little lag between the curve and the price bars at the price turning points where many times the curve leads the price bar turning points. This no lag feature of the 4th order polynomial system is a valuable attribute. Overall the 4th order polynomial system did a good job in minimizing the losses due to the inevitable whipsaws that will occur in any trading system and maximizing the profits from the major intraday trend moves of IBM.

In order to use this system in real time trading, at least ten to twenty more test and out-of-sample windows would have to be examined to make sure that the above results above were not due to pure chance.

Info on Dennis Meyers

Dennis Meyers (info@MeyersAnalytics.com.) has a doctorate in applied mathematics in engineering. He is a private trader, and president of Meyers Analytics (www.meyersanalytics.com). His firm specializes in consulting for financial institutions and developing publicly available analytical software for traders.

Figure 2a Test Window Performance Summary for IBM Least Squares Polynomial System
IBM-5 min bars 07/05/2001 - 08/03/2001

Performance Statistics based upon Buying and Selling 1000 shares of IBM

Performance Summary: All Trades

Total Net Profit	$16,230.00	Open position P/L	$0.00
Gross Profit	$35,030.00	Gross Loss	($18,800.00)
Total # of trades	114	Percent profitable	53.51%
Number winning trades	61	Number losing trades	53
Largest winning trade	$2,120.00	Largest losing trade	($1,360.00)
Average winning trade	$574.26	Average losing trade	($354.72)
Ratio avg win/avg loss	1.62	Avg trade (win & loss)	$142.37
Max consec. Winners	11	Max consec. losers	4
Avg # bars in winners	17	Avg # bars in losers	10
Max intraday drawdown	($2,560.00)		
Profit Factor	1.86	Max # contracts held	1

Performance Summary: Long Trades

Total Net Profit	$6,120.00	Open position P/L	$0.00
Gross Profit	$16,880.00	Gross Loss	($10,760.00)
Total # of trades	56	Percent profitable	51.79%
Number winning trades	29	Number losing trades	27
Largest winning trade	$2,110.00	Largest losing trade	($1,360.00)
Average winning trade	$582.07	Average losing trade	($398.52)
Ratio avg win/avg loss	1.46	Avg trade (win & loss)	$109.29
Max consec. Winners	5	Max consec. losers	4
Avg # bars in winners	14	Avg # bars in losers	8
Max intraday drawdown	($2,850.00)		
Profit Factor	1.57	Max # contracts held	1

Performance Summary: Short Trades

Total Net Profit	$10,110.00	Open position P/L	$0.00
Gross Profit	$18,150.00	Gross Loss	($8,040.00)
Total # of trades	58	Percent profitable	55.17%
Number winning trades	32	Number losing trades	26
Largest winning trade	$2,120.00	Largest losing trade	($890.00)
Average winning trade	$567.19	Average losing trade	($309.23)
Ratio avg win/avg loss	1.83	Avg trade (win & loss)	$174.31
Max consec. Winners	6	Max consec. losers	3
Avg # bars in winners	20	Avg # bars in losers	12
Max intraday drawdown	($1,630.00)		
Profit Factor	2.26	Max # contracts held	1

Figure 2b Test Window Performance Summary for IBM Least Squares Polynomial System
IBM-5 min bars 07/11/2001 - 08/10/2001

Performance Statistics based upon Buying and Selling 1000 shares of IBM

Performance Summary: All Trades

Total Net Profit	$15,910.00	Open position P/L	$0.00
Gross Profit	$34,330.00	Gross Loss	($18,420.00)
Total # of trades	109	Percent profitable	54.13%
Number winning trades	59	Number losing trades	50
Largest winning trade	$2,070.00	Largest losing trade	($1,090.00)
Average winning trade	$581.86	Average losing trade	($368.40)
Ratio avg win/avg loss	1.58	Avg trade (win & loss)	$145.96
Max consec. Winners	8	Max consec. losers	4
Avg # bars in winners	18	Avg # bars in losers	11
Max intraday drawdown	($3,030.00)		
Profit Factor	1.86	Max # contracts held	1

Performance Summary: Long Trades

Total Net Profit	$9,080.00	Open position P/L	$0.00
Gross Profit	$17,240.00	Gross Loss	($8,160.00)
Total # of trades	53	Percent profitable	56.60%
Number winning trades	30	Number losing trades	23
Largest winning trade	$2,070.00	Largest losing trade	($970.00)
Average winning trade	$574.67	Average losing trade	($354.78)
Ratio avg win/avg loss	1.62	Avg trade (win & loss)	$171.32
Max consec. Winners	4	Max consec. losers	4
Avg # bars in winners	14	Avg # bars in losers	7
Max intraday drawdown	($1,800.00)		
Profit Factor	2.11	Max # contracts held	1

Performance Summary: Short Trades

Total Net Profit	$6,830.00	Open position P/L	$0.00
Gross Profit	$17,090.00	Gross Loss	($10,260.00)
Total # of trades	56	Percent profitable	51.79%
Number winning trades	29	Number losing trades	27
Largest winning trade	$1,480.00	Largest losing trade	($1,090.00)
Average winning trade	$589.31	Average losing trade	($380.00)
Ratio avg win/avg loss	1.55	Avg trade (win & loss)	$121.96
Max consec. Winners	7	Max consec. losers	5
Avg # bars in winners	22	Avg # bars in losers	15
Max intraday drawdown	($2,020.00)		
Profit Factor	1.67	Max # contracts held	1

Figure 3 Combined Walk Forward Out-Of-Sample Performance Summary for Least Squares Polynomial System IBM-5 min bars 08/06/2001 - 08/17/2001

Performance Statistics based upon Buying and Selling 1000 shares of IBM

Performance Summary: All Trades

Total Net Profit	$6,950.00	Open position P/L	$0.00
Gross Profit	$12,200.00	Gross Loss	($5,250.00)
Total # of trades	42	Percent profitable	59.52%
Number winning trades	25	Number losing trades	17
Largest winning trade	$1,470.00	Largest losing trade	($680.00)
Average winning trade	$488.00	Average losing trade	($308.82)
Ratio avg win/avg loss	1.58	Avg trade (win & loss)	$165.48
Max consec. Winners	5	Max consec. losers	5
Avg # bars in winners	22	Avg # bars in losers	12
Max intraday drawdown	($2,140.00)		
Profit Factor	2.32	Max # contracts held	1

Performance Summary: Long Trades

Total Net Profit	$3,250.00	Open position P/L	$0.00
Gross Profit	$5,070.00	Gross Loss	($1,820.00)
Total # of trades	20	Percent profitable	55.00%
Number winning trades	11	Number losing trades	9
Largest winning trade	$1,470.00	Largest losing trade	($450.00)
Average winning trade	$460.91	Average losing trade	($202.22)
Ratio avg win/avg loss	2.28	Avg trade (win & loss)	$162.50
Max consec. Winners	6	Max consec. losers	3
Avg # bars in winners	14	Avg # bars in losers	10
Max intraday drawdown	($910.00)		
Profit Factor	2.79	Max # contracts held	1

Performance Summary: Short Trades

Total Net Profit	$3,700.00	Open position P/L	$0.00
Gross Profit	$7,130.00	Gross Loss	($3,430.00)
Total # of trades	22	Percent profitable	63.64%
Number winning trades	14	Number losing trades	8
Largest winning trade	$1,130.00	Largest losing trade	($680.00)
Average winning trade	$509.29	Average losing trade	($428.75)
Ratio avg win/avg loss	1.19	Avg trade (win & loss)	$168.18
Max consec. Winners	6	Max consec. losers	5
Avg # bars in winners	28	Avg # bars in losers	15
Max intraday drawdown	($2,150.00)		
Profit Factor	2.08	Max # contracts held	1

FIGURE 4 Out-Of-Sample Trade By Trade Summary
IBM 5min Noise Channel 2 System Trade Size = 1000 Shares 03/26/2001 to 04/06/2001

Entry Date	Entry Time		Entry Price	Exit Date	Exit Time	Exit Price	Bars InTrd	Trade $P&L	Trade %P&L	Trade Max$Pft	Time	Trade Max$DD	Time
8/6/01	940	Sell	107.16	8/6/01	1040	106.94	12	220	0.21%	460	1000	(80)	945
8/6/01	1040	Buy	106.94	8/6/01	1135	106.60	11	(340)	-0.32%	0	1040	(340)	1135
8/6/01	1135	Sell	106.60	8/6/01	1525	106.26	46	340	0.32%	770	1440	(30)	1140
8/6/01	1525	Buy	106.26	8/6/01	1555	106.47	6	210	0.20%	210	1555	(100)	1525
8/7/01	945	Sell	105.80	8/7/01	1035	106.32	10	(520)	-0.49%	0	945	(520)	1025
8/7/01	1035	Buy	106.32	8/7/01	1120	107.10	9	780	0.73%	1050	1045	0	1035
8/7/01	1120	Sell	107.10	8/7/01	1545	105.97	53	1130	1.06%	1800	1515	(110)	1125
8/7/01	1545	Buy	105.97	8/7/01	1555	106.22	2	250	0.24%	250	1555	0	1545
8/8/01	940	Sell	104.99	8/8/01	1035	105.64	11	(650)	-0.62%	0	940	(660)	1030
8/8/01	1035	Buy	105.64	8/8/01	1135	105.50	12	(140)	-0.13%	400	1100	(160)	1050
8/8/01	1135	Sell	105.50	8/8/01	1230	105.48	11	20	0.02%	100	1210	(60)	1225
8/8/01	1230	Buy	105.48	8/8/01	1315	105.19	9	(290)	-0.27%	0	1230	(420)	1305
8/8/01	1315	Sell	105.19	8/8/01	1555	104.20	32	990	0.94%	1060	1545	0	1315
8/9/01	940	Sell	103.39	8/9/01	1035	103.06	11	330	0.32%	640	1015	0	940
8/9/01	1035	Buy	103.06	8/9/01	1140	103.21	13	150	0.15%	550	1110	(210)	1035
8/9/01	1140	Sell	103.21	8/9/01	1300	103.37	16	(160)	-0.16%	240	1235	(290)	1140
8/9/01	1300	Buy	103.37	8/9/01	1355	103.92	11	550	0.53%	1120	1335	0	1300
8/9/01	1355	Sell	103.92	8/9/01	1500	103.76	13	160	0.15%	370	1430	0	1355
8/9/01	1500	Buy	103.76	8/9/01	1555	104.04	11	280	0.27%	450	1525	0	1500
8/10/01	940	Sell	104.00	8/10/01	1100	103.04	16	960	0.92%	1140	1025	0	940
8/10/01	1100	Buy	103.04	8/10/01	1210	103.55	14	510	0.49%	760	1150	(80)	1110
8/10/01	1210	Sell	103.55	8/10/01	1300	104.06	10	(510)	-0.49%	100	1215	(510)	1300
8/10/01	1300	Buy	104.06	8/10/01	1345	104.11	9	50	0.05%	400	1315	(60)	1330
8/10/01	1345	Sell	104.11	8/10/01	1455	104.24	14	(130)	-0.12%	200	1420	(170)	1435
8/10/01	1455	Buy	104.24	8/10/01	1550	104.66	11	420	0.40%	670	1520	0	1455
8/10/01	1550	Sell	104.66	8/10/01	1555	104.84	1	(180)	-0.17%	0	1550	(180)	1555
8/13/01	940	Buy	105.08	8/13/01	1005	105.07	5	(10)	-0.01%	320	945	(20)	940
8/13/01	1005	Sell	105.07	8/13/01	1215	105.75	26	(680)	-0.65%	320	1035	(680)	1215
8/13/01	1215	Buy	105.75	8/13/01	1235	105.30	4	(450)	-0.43%	0	1215	(570)	1225
8/13/01	1235	Sell	105.30	8/13/01	1500	105.90	29	(600)	-0.57%	350	1305	(650)	1455
8/13/01	1500	Buy	105.90	8/13/01	1535	106.30	7	400	0.38%	660	1520	0	1500
8/13/01	1535	Sell	106.30	8/13/01	1555	105.80	4	500	0.47%	550	1545	0	1535
8/14/01	940	Buy	106.62	8/14/01	1005	106.48	5	(140)	-0.13%	0	940	(310)	940
8/14/01	1005	Sell	106.48	8/14/01	1115	106.11	14	370	0.35%	910	1100	(200)	1025
8/14/01	1115	Buy	106.11	8/14/01	1425	106.01	38	(100)	-0.09%	570	1240	(170)	1130
8/14/01	1425	Sell	106.01	8/14/01	1555	105.83	18	180	0.17%	360	1540	(270)	1445
8/15/01	940	Buy	106.37	8/15/01	955	106.10	3	(270)	-0.25%	130	940	(320)	950
8/15/01	955	Sell	106.10	8/15/01	1555	105.10	72	1000	0.94%	1000	1520	(100)	955
8/16/01	940	Sell	104.45	8/16/01	1025	104.14	9	310	0.30%	690	1010	(480)	940
8/16/01	1025	Buy	104.14	8/16/01	1555	105.61	66	1470	1.41%	1800	1540	(220)	1035
8/17/01	940	Sell	105.10	8/17/01	1555	104.56	75	540	0.51%	1180	1530	0	940
								Total 6950	Average 0.16%	Average 526		Average (194)	

Figure 5a IBM-5 min bars 08/14/2001 - 08/17/2001 Least Squares Polynomial System