

Trading mit Regressionskurven

1. Was soll Regression sein und wozu kann man das brauchen?

Da die diversen Methoden, mit denen Bots traden können, durch das "Rauschen" der Kurswerte behindert werden, versucht man mit mathematischen Methoden, das Rauschen wegzurechnen, um die grundlegende Kursbewegung klarer zu erkennen.

Dazu gibt es viele Verfahren.

Eines der bekanntesten ist das der gleitenden Mittelwerte (SMA und EMA), die wohl jeder Cryptonutzer kennen wird.

Gleitende Mittelwerte haben neben ihren Vorzügen jedoch auch Nachteile; sie hinken dem Kursgeschehen nach (SMA stärker als EMA) und sie erlauben als "passive" Methoden nur schwache Annahmen über die Zukunft.

Regression ist ein anderes mathematisches Verfahren. Hier nimmt an, dass die Kurse einer Geraden, einer Parabel oder einer anderen festgelegten mathematischen Funktion folgen und kann dann nachfolgend ermitteln, in wie weit diese Annahme stimmt. ;-)

Passen die Kurse gut zur Annahme, dann vermutet man, dass sie sich zumindest in näherer Zukunft noch an diese mathematische Funktion halten werden, und kann damit abschätzen, ob der Kauf oder Verkauf eines Coins aussichtsreich erscheint.

Die (inzwischen nicht mehr mitgelieferte) Methode LinReg arbeitete mit Geraden, während QuadReg auf Parabeln aufsetzt.

2. Regressionsgeraden

Hier geht man davon aus, dass die Kurse um eine gerade Linie herum streuen. Welche Linie das ist, und wie genau oder ungenau die Kurse dieser theoretischen Linie folgen, kann man mit linearer Regression errechnen.

Zuerst wird aus den Kurswerten die bestmögliche Gerade ausgerechnet, die den Kursen "am wenigsten schlecht" entspricht, d.h. diejenige Gerade, bei der die Fehler (Abstände der Kurswerte von der Gerade) am kleinsten ist.

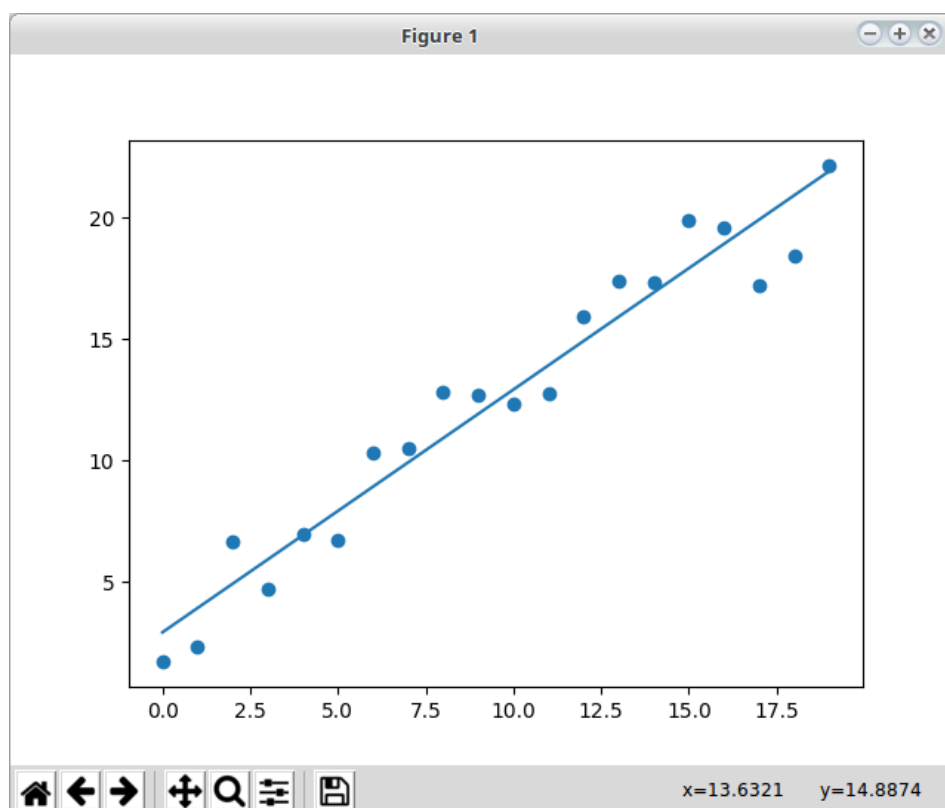
Über die Summe der ganzen Fehler kann man dann einen Güte-Koeffizienten ("R2") ausrechnen, um die Qualität der Geraden mathematisch beurteilen zu können.

Dabei bedeutet $R^2 = 1.0$, dass die Punkte perfekt auf einer Gerade liegen; kleinere Werte geben an, dass die Punkte um die Gerade streuen. (Die Ausdrucksweise ist hier ein bisschen flapsig, da sich die Gerade ja nach den Punkten richtet, und nicht die Punkte nach der Gerade.)

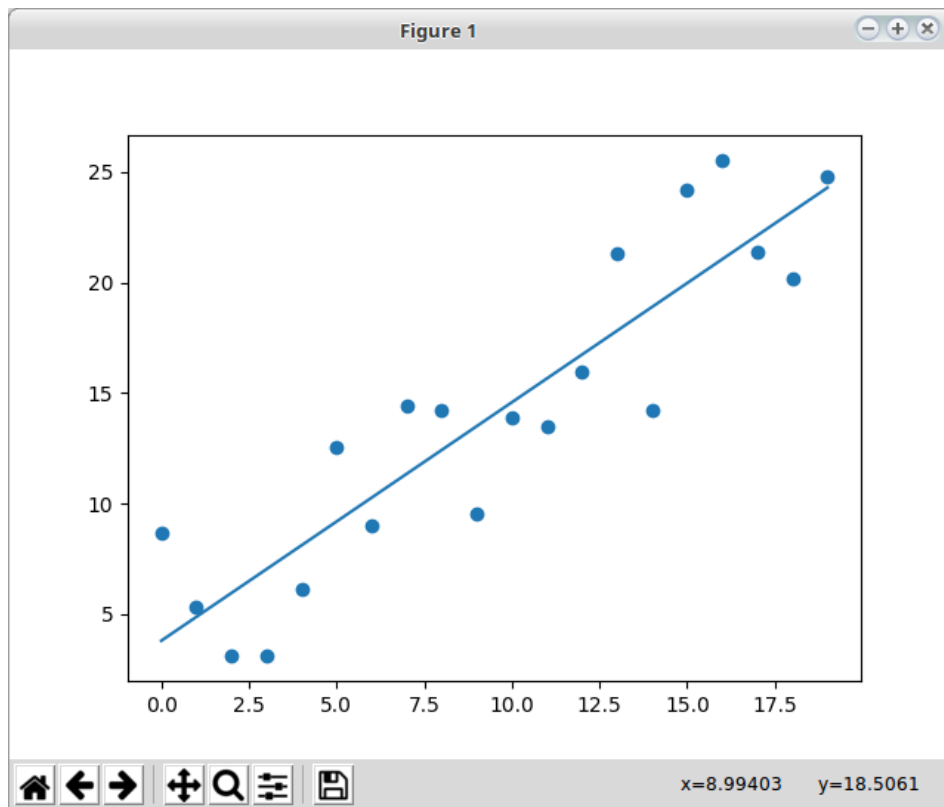
Die Qualität nimmt schneller ab, als man nach der Zahl vermuten würde. Ein R^2 um 0.60 herum ist schon kaum mehr brauchbar.

Mehrere Beispiele:

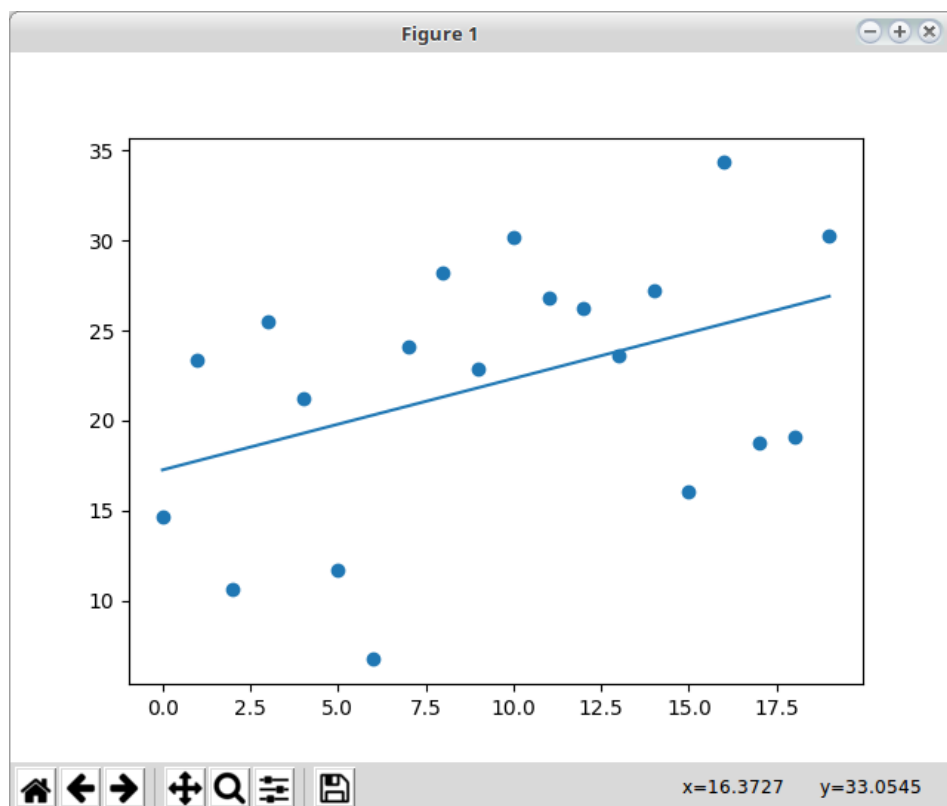
Die blauen Punkte sollen die Schlusskurse eines Coins darstellen. Die blaue Gerade ist die mit der Regressionsmethode ermittelte beste Näherung.



Hier ist die Näherung sehr gut, R^2 ist 0.994.



Hier streuen die Punkte schon viel mehr, man kann den Trend der Kurswerte in Richtung der Geraden aber gut erkennen. R^2 ist noch 0.811.



Hier haben die Punkte eigentlich nichts mehr mit der Gerade zu tun, die Gerade ist zwar die beste, noch mögliche Näherung, hat bei dieser Streuung aber keine Aussagekraft mehr. R^2 ist dementsprechend nur noch 0.172.

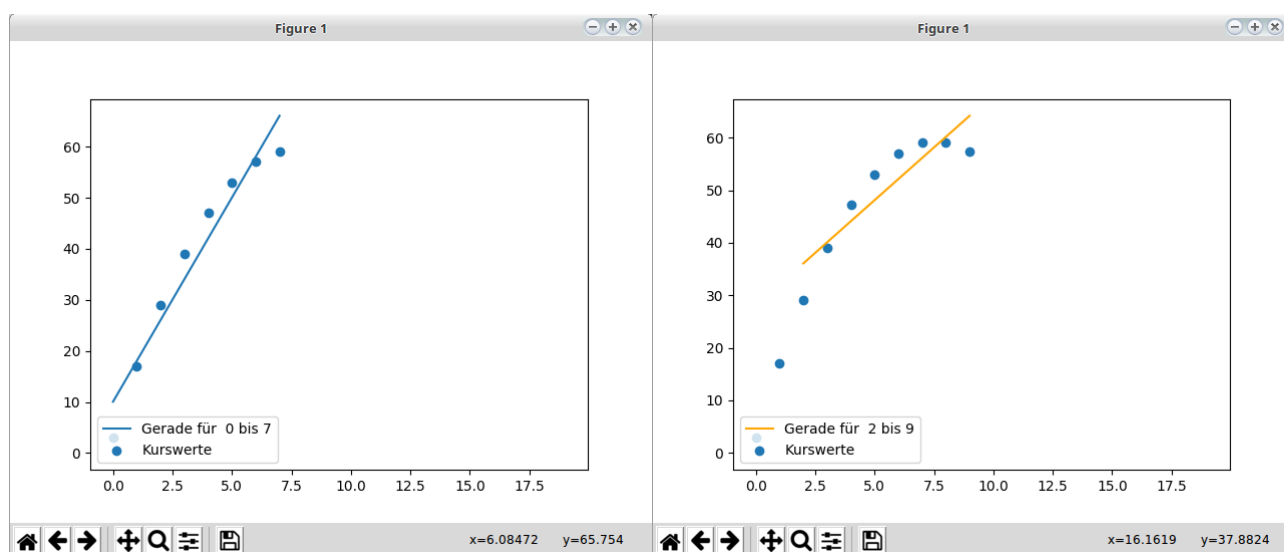
Nun ist es sehr selten, dass Kurse stur entlang von Geraden verlaufen. Normalerweise gibt es Hochs und Tiefs, Wellen, längere und kürzere Steigungen oder Abfälle.

Da Geraden keine Kurven können, behilft man sich damit, dass man nur die letzten soundsoviel Kurswerte für die Berechnung benutzt, und immer, wenn von der Exchange eine neue Candle kommt, wieder eine neue Gerade berechnet.

Damit man sich das besser vorstellen kann, hier ein Beispiel.

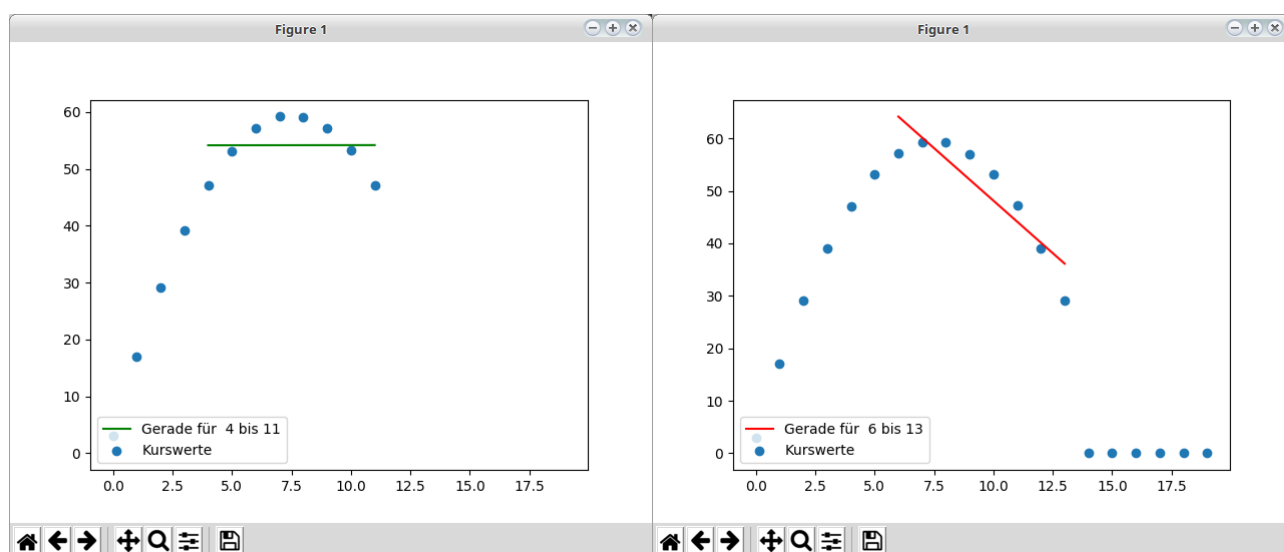
Wir bekommen sukzessive von der Börse die Kursdaten und berechnen immer über die letzten 8 Candles die Regressionsgerade.

(Der Übersichtlichkeit halber hier nicht für jede neue Candle, sondern nur für jede zweite. Und stört euch nicht an, dass die erste Candle mit Candle 0 bezeichnet wird - Informatiker fangen immer mit der Null das Zählen an.)



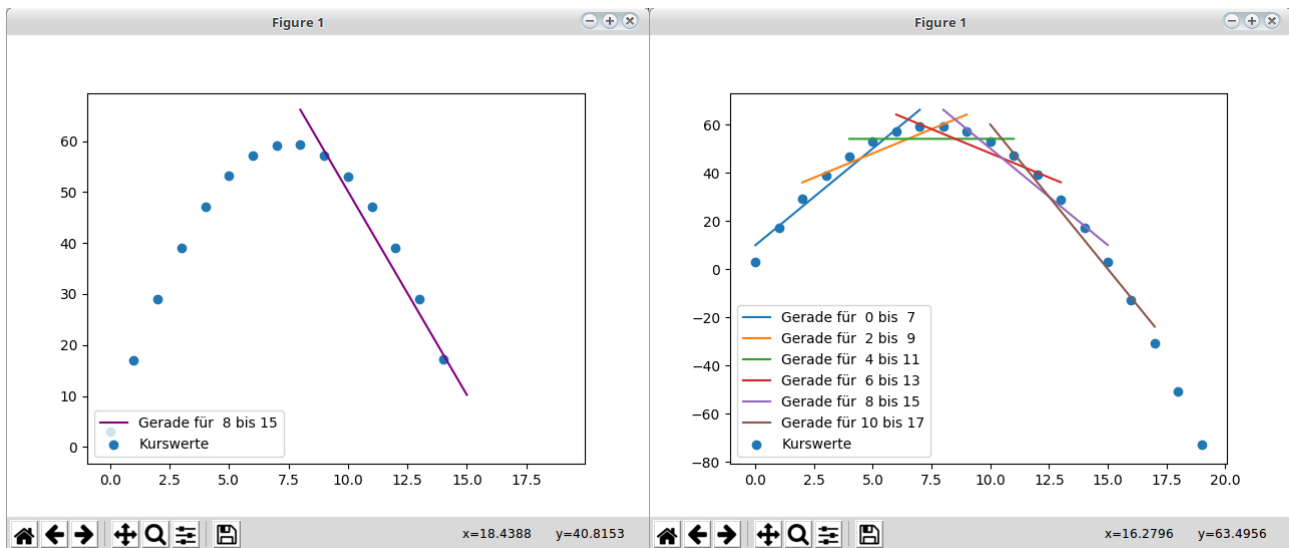
Candles von 0 bis 7 bekannt

Candles von 0 bis 9 bekannt



Candles von 0 bis 11 bekannt

Candles von 0 bis 13 bekannt



Candles von 0 bis 15 bekannt

Hier alle Candles und Geraden in einem Bild

Man sieht, wie die Geraden die Kursform "entlangwandern". Man sieht ebenfalls, dass erst die rote Gerade (die über die Candles 6 bis 13 berechnet wird) deutlich anzeigt, dass der Kurs zu fallen begonnen hat.

Die Näherung von Kursen durch Regressionsgeraden hinkt also immer ein Stück hinterher. Beim Traden heißt das, dass ein Teil der erreichten Gewinne erst wieder abgegeben wird, bis eine fallende Gerade den Verkauf signalisiert.

Das klingt im ersten Moment ziemlich traurig, ist beim Traden aber völlig normal. Nahezu alle Hilfsmittel zum Bestimmen des Kursverhaltens hinken mehr oder weniger hinterher.

Das liegt in der Natur der Sache - Kurse verlaufen nicht so ideal wie die Kurven in den Bildern hier, sondern schwanken zusätzlich zum großen Trend noch rauf und runter. Wer bei jedem kleinen Kursabfall sofort verkauft, sichert zwar den momentanen Wert der Coins, fällt aber dafür oft auf kleine Unterbrechungen oder Korrekturen eines eigentlich beständigen Aufwärtstrends herein:



In diesem Beispiel eines langen Aufwärtstrends finden sich mehrere Stellen, an denen man in Versuchung sein konnte, zu verkaufen. Wer es nicht tat und erst nach der großen roten Kerze am 05. September 2018 verkauft hat, konnte den ganzen Trend mitnehmen.

Insofern ist das ganze ein Balance-Akt - man muss die richtige Mitte zwischen frühem Verkauf und spätem Verkauf finden, um auf Dauer ein gutes Gesamtergebnis zu erzielen.

Dazu benutzt die Regressionsmethode einen Trick.

Man berechnet nicht jeweils eine Gerade über genau festgelegte Zahl xx an Candles, sondern man "probiert" viele verschiedene Werte für dieses xx aus und entscheidet sich letztlich für die Gerade, deren R^2 am besten ist, die also dem Kurs am besten entspricht.

Das kann eine Gerade über mehrere Wochen sein, wenn wie im August letzten Jahres der BTC einen relativ gleichmäßigen Aufstieg an den Tag legt. Da wäre eine Gerade vom 20. August bis Anfang September sehr genau gelegen und hätte sicherlich eine Güte von R^2 nahe 1.0 erzielt.

Am 8. September in obiger Grafik wäre dann vermutlich eine kurze Gerade vom 6. September bis zum 8. September "besser" gewesen als eine lange vom 20. August bis zum 8. September. Denn durch den groben Abfall am 05./06.09. wurde die bislang hohe Qualität der langen Geraden zerstört.

Um diese Mechanik zu verstehen, wären viele, viele Beispiele angebracht, was über den Rahmen dieser Kurzeinführung weit hinaus ginge.

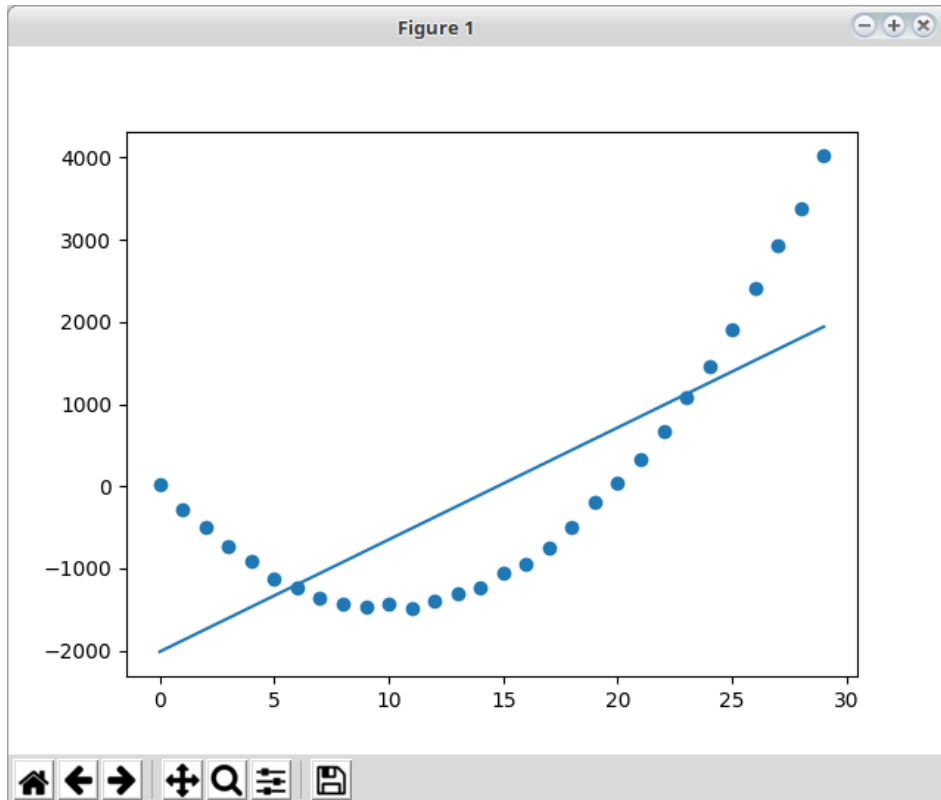
Interessant ist das Fazit:

Durch diesen Trick mit der variablen Länge ist sozusagen ein "dynamisches Trailing" in das Regressionsgeraden-Trading eingebaut.

Je länger der Trend läuft - und desto größer damit die Gewinne - desto größer können Korrekturen oder kleinere Einbrüche sein. Das "Trailing" wächst mit den Gewinnen.

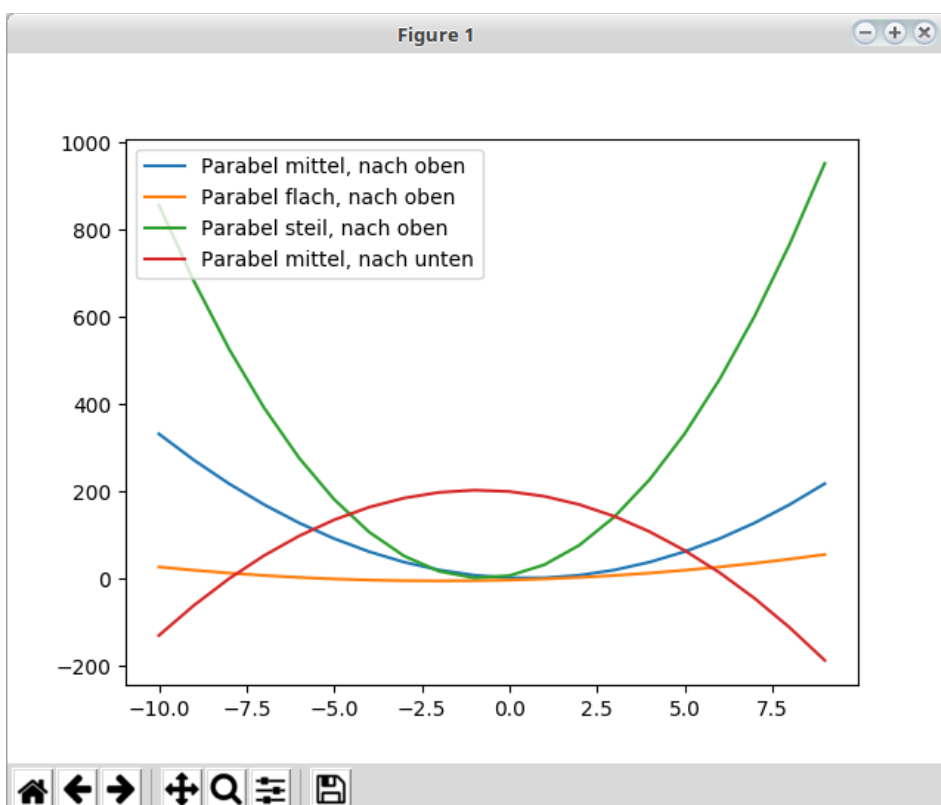
3. Regressionsparabel

Oft folgen Kurse durchaus einer erkennbaren Form, lassen sich mit einer Geraden aber schlecht beschreiben. Kein Wunder - eine Gerade kann halt nur gerade. ;-)

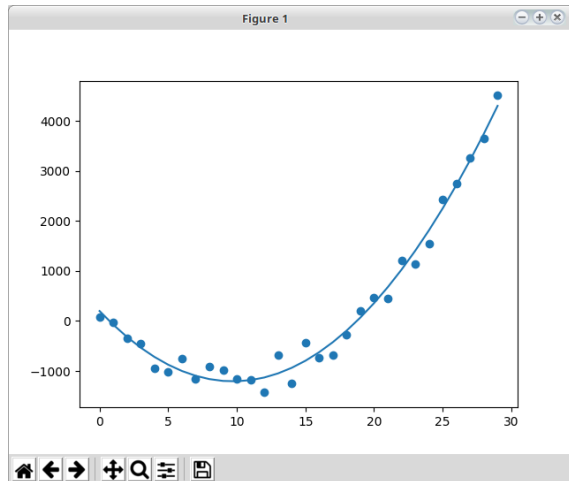


Hier ist die Gerade ($R^2 = 0.592$) nicht mehr zu gebrauchen. Dafür kommt jetzt die Parabel ins Spiel.

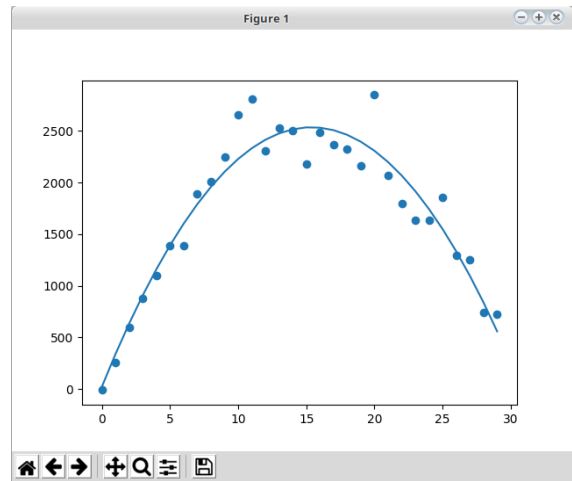
Parabeln gibt es in vielen Varianten:



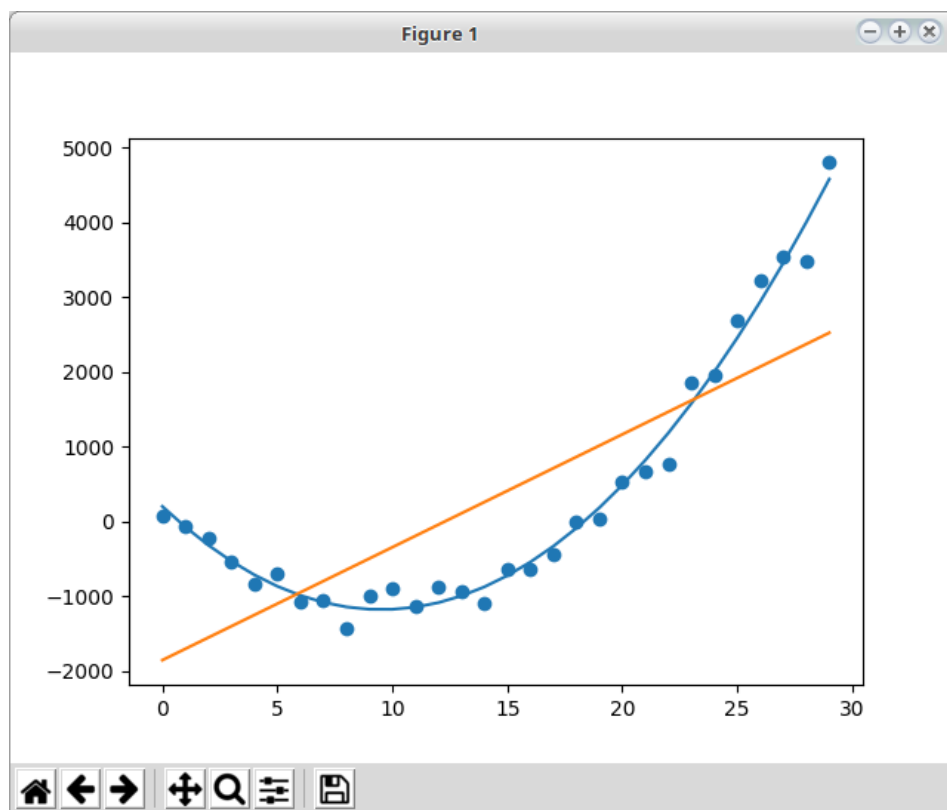
Die wichtigsten sind aber die zwei Hauptvarianten:



Die, die wie große 'U' oder 'V' aussehen -
"nach oben geöffnet".



Die, die wie 'U' oder 'V' auf dem Kopf stehend
aussehen - "nach unten geöffnet".



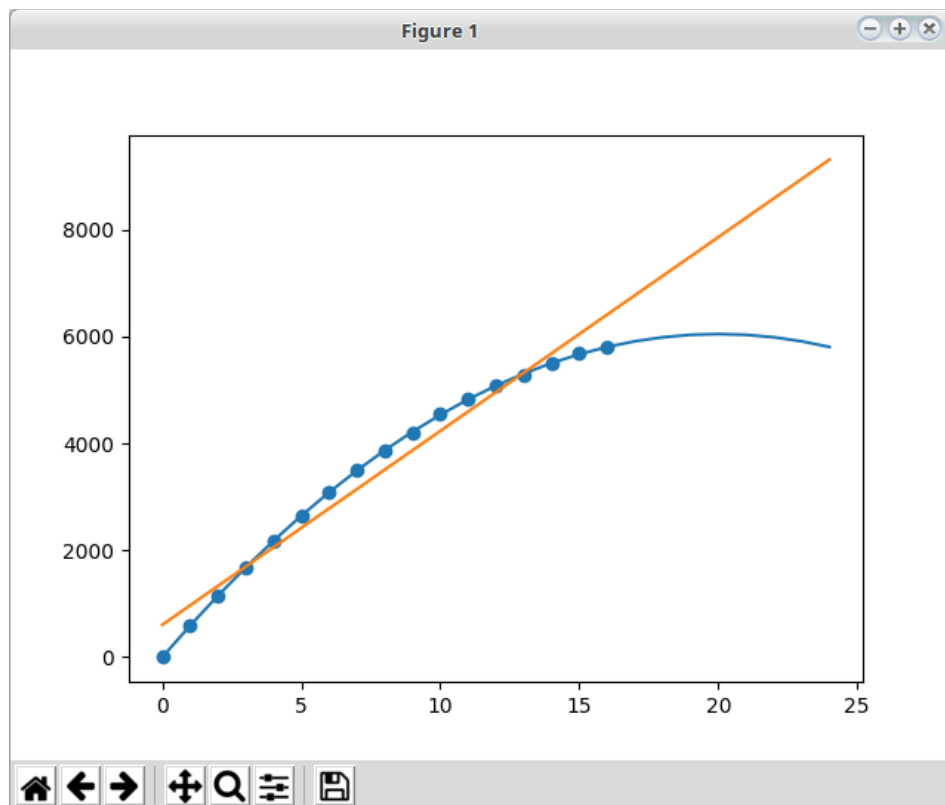
Eine Parabel kann gebogene Verläufe sehr viel besser annähern. Hier ist $R^2 = 0.986$, während die Gerade nur auf $R^2 = 0.614$ kommt.

Weil Parabel vielseitiger sind - siehe das vorangegangene Bild mit den vier bunten Parabeln - passen Parabeln an verschiedene Kurssituationen und lassen Veränderungen für die Zukunft vermuten, während Geraden "nur" gerade sind und darauf spekulieren, dass sich der gleiche Trend unverändert in der Zukunft fortsetzen könnte.

Bei allen Grafiken gehen wir davon aus, dass alle Kurswerte, die wir im Moment kennen, durch blaue Punkte dargestellt werden. Aus diesen Werten lassen wir die am besten dazu passende Gerade bzw. Parabel errechnen und ins gleiche Diagramm einzeichnen.

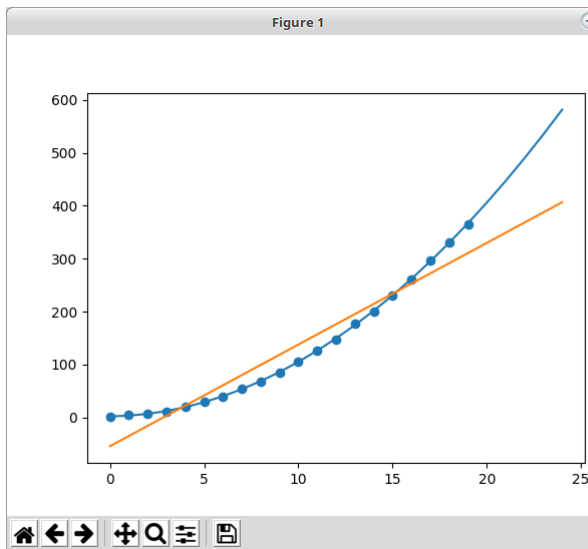
Da wir ja an der Einschätzung, wie sich der Kurs weiterentwickeln könnte, interessiert sind, lassen wir zusätzlich die Geraden und Parabeln noch ein bisschen in die Zukunft weiterlaufen.

Parabeln erlauben differenziertere Prognosen als Geraden:

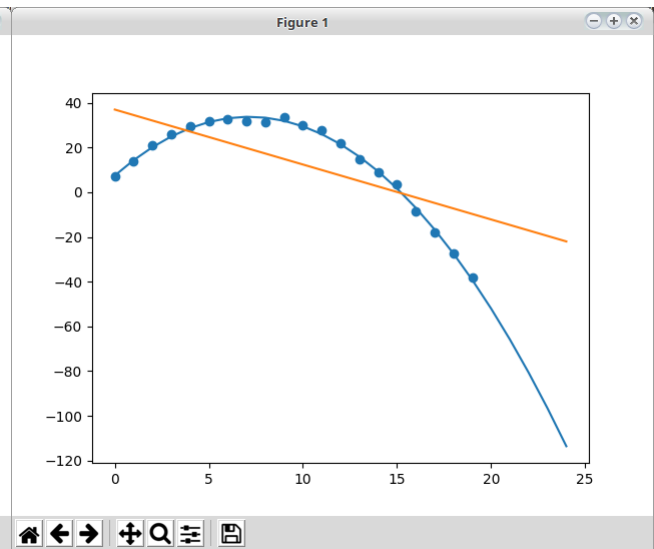


Während die Gerade den Kurstrend unverdrossen unverändert in die Zukunft fortschreibt, "erkennt" die Parabel, dass sich der Anstieg immer weiter abflacht, und der Kurs bald ein Top bilden könnte.

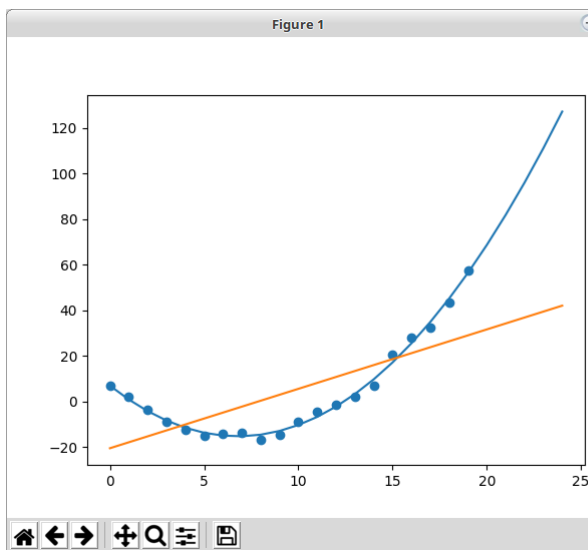
Wie man am letzten Bild der vorhergehenden Seite schon erkennen kann, können Parabeln eine Vielzahl von Kursformationen besser annähern:



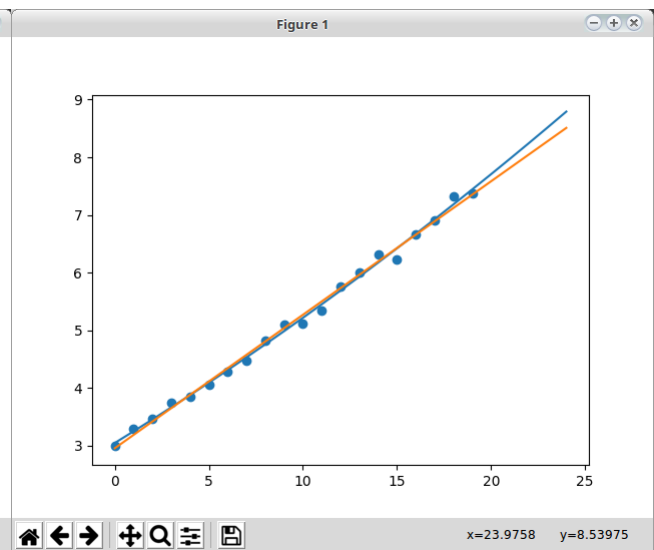
Anstiege



Erreichen eines Tops mit nachfolgendem Abfall

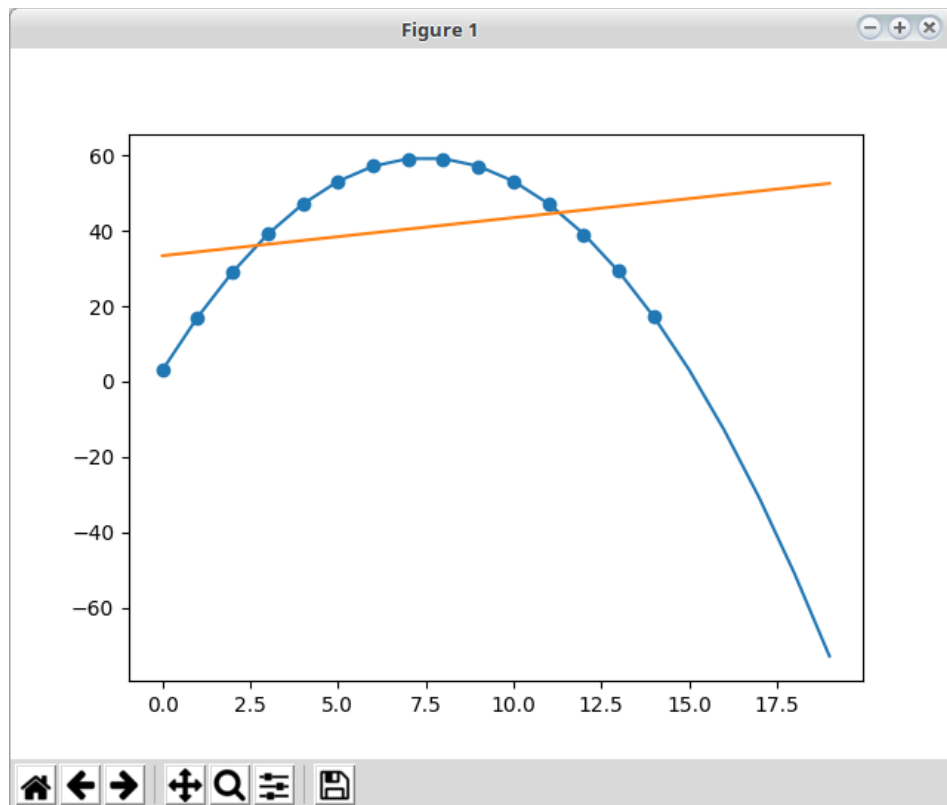


Bodenbildung und Wiederanstieg



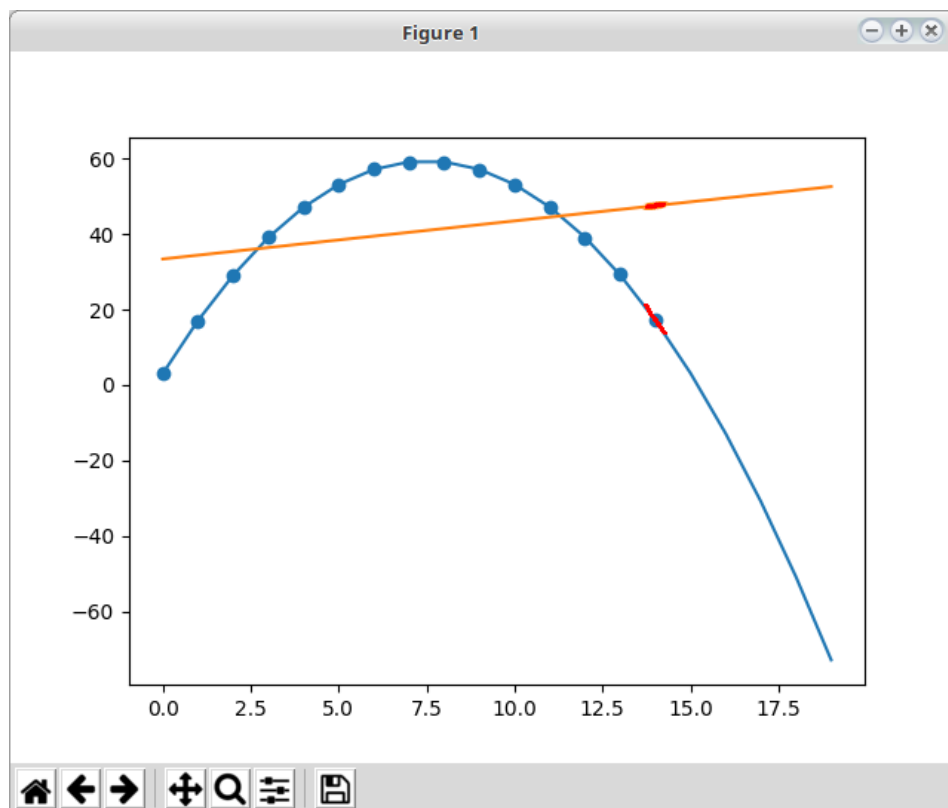
Selbst Geraden können sie annähern, weil mathematisch die Gerade nur ein Sonderfall der Parabel ist.

Ein weiterer großer Vorteil der Parabeln ist die wesentlich höhere "Reaktionsgeschwindigkeit".



Während die Gerade bei diesen 15 Kurswerten noch davon ausgeht, dass wir uns in einem Anstieg befinden und zum Kauf rät, zeigt die Parabel schon bei weniger als 15 Punkten längst einen immer negativer werdenden Abwärtstrend und rät zum Verkauf.

Da wir die beiden Funktionen nach rechts weiterberechnen lassen, können wir sehen, wie sich Gerade und Parabel die weitere Entwicklung der Kurse "vorstellen".



Da der Computer nach der Bewegung hinter dem Rauschen sucht, nimmt man für die vermutete aktuelle Steigung des Kursverlaufes nicht einfach den Unterschied der letzten Candle zur aktuellen Candle, sondern die Steigung der Regressionsfunktion.

Der vermutete momentane Kurstrend, also die aktuelle Steigung der Kurswerte, ist dann die Steigung der Geraden oder der Parabel am rechten Kurswert, hier in Rot eingezeichnet.

Anhand dieser errechneten Steigung (dargestellt durch das kleine rote Strichlein an der Gerade bzw Parabel) entscheidet der Computer, ob er kaufen oder verkaufen soll.

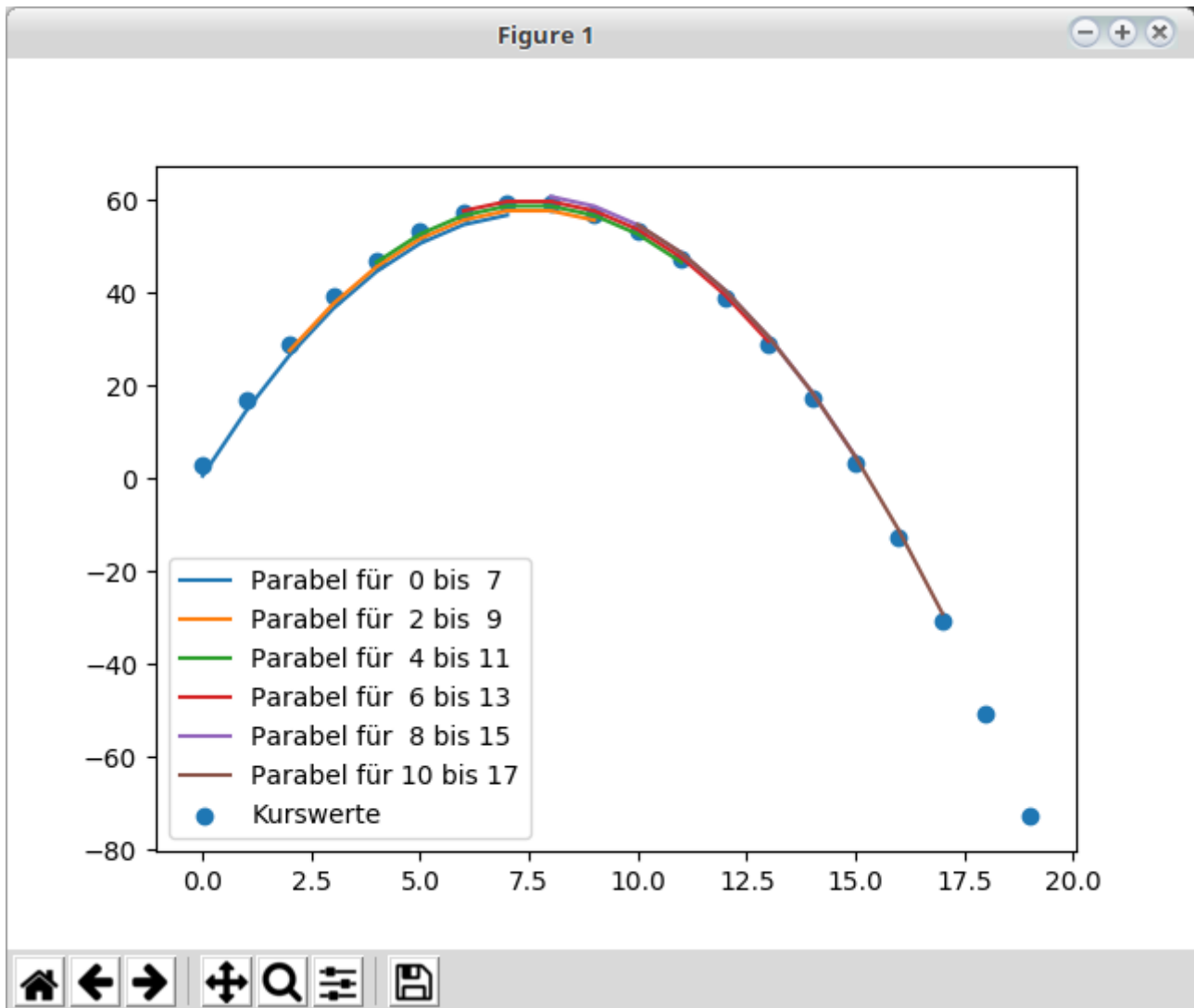
Hat dieses kleine rote Strichlein eine Tendenz von links unten nach rechts oben (wie bei der Geraden), bedeutet das für den Bot, dass der Kurs durchschnittlich steigt. Fällt das Strichlein von links oben nach rechts unten (wie bei der Parabel), wird ein durchschnittliches Fallen des Kursen angenommen.

Die Gerade über die Kurswerte 0 bis 14 sagt den Kurs völlig falsch voraus. Ihrer Meinung nach steigt der Kurs weiterhin gleichmäßig.

Die Parabel zeigt am letzten Kurswert für $x = 14$ ein starkes Abfallen - der kleine rote Strich an der Parabel zeigt in Richtung sinkende Kurse.

Wie bei den Geraden berechnet man eine Parabel nicht für alle Kurswerte der Vergangenheit, sondern beschränkt sich auf die letzten xx Werte. Wir haben gesehen, wie sich die Gerade mit jeder weiteren Candle verändert und immer weiter den Kurs entlangwandert.

Das Bild, auf dem alle Kurswerte und alle Geraden gleichzeitig dargestellt waren, jetzt in der Parabelversion:



Hier kommen alle Parabeln zum gleichen Ergebnis, egal, für welche Candles sie berechnet wurden, und liegen deshalb aufeinander. (Sie sind hier leicht verschoben gezeichnet, damit man sie besser erkennen kann, lägen aber in Wirklichkeit genau aufeinander.)

Die Grafik zeigt, dass bereits die orange Parabel, die auf den Candles 2 bis 9 basiert, an ihrer letzten Candle 9 einen ersten Kurs-Abfall zeigt. Und spätestens die grüne Parabel zeigt bei ihrer letzten Candle 11 deutlich an, dass verkauft werden sollte.

Bei den Geraden mussten wir auf die Gerade zur Candle 13 warten. Die Parabelmethode hat in diesem Beispiel also einen Vorsprung von 2 Candles.

Wo viel Licht ist, muss natürlich auch Schatten sein.

Im Fall der Parabeln stellt ihre Anpassbarkeit an vielerlei Kurssituationen auch ein Problem dar.

Da Parabeln so vielgestaltig sein können, kann man in viel mehr Fällen als bei Geraden gut passende Parabeln errechnen.

Damit kommt es öfters als bei Geraden vor, dass eine kurze Parabel über wenig Candles eine höhere Genauigkeit aufweist und aufgrund ihres besseren R^2 die längere Parabel über den bisherigen Trend überstimmen kann. (Wir erinnern uns - die Kurve mit dem besten R^2 -Wert "gewinnt".)

Das Aussitzen längeren Trends, das sich bei LinReg so positiv bemerkbar macht, wird also vermutlich ein bisschen schwieriger. (Im Gegenzug kompensieren das allerdings die Vorteile der Parabeln beim Finden eines Anfangs und des schnelleren Abbrechens bei Marktdrehungen.)

Ein zweites Problem ist die Doppelnatur als nach oben geöffnete oder nach unten geöffnete Parabel.

Wenn wir von der besten Parabel über unsere aktuellen Kursdaten die momentane Steigung an der neuesten Candle berechnen, müssen wir zusätzlich noch ermitteln, welchen der beiden Parabeltypen wir vor uns haben.

Eine positive Steigung alleine genügt nicht. Auch die linke Hälfte einer kopfstehenden Parabel hat positive Trendwerte, die aber mehr oder weniger bald auf das Top zulaufen, nach dem die Kurse möglicherweise wieder abwärts gehen.

Theoretisch könnte man auch auf diesem linken Ast der kopfstehenden Parabel mit Erfolg kaufen, wenn man noch weit genug vom Top entfernt ist. Ich habe mit verschiedenen Varianten experimentiert, habe in den Tests aber nie einen merklichen Vorteil damit bekommen.

Deswegen wird in QuadReg nur auf dem rechten, aufstrebenden Ast einer nach oben geöffneten Parabel gekauft.