Since the matrix A^tA is symmetric positive definite, we obtain that the Hessian $D^2 f(x_0)$ is symmetric positive definite, and therefore x_0 is a minimum point for the function $f(x)$. We conclude that the point x_0 is a global minimum point for $f(x)$, since x_0 is the only critical point of $f(x)$.

Thus, the solution to the least squares problem (8.1) is given by (8.15), i.e.,

$$x \;=\; (A^tA)^{-1}A^ty. \tag{8.16}$$

Note that the numerical value of x from (8.16) is computed by solving the linear system $(A^tA)x = A^ty$ using the Cholesky solver from Table 6.2, since A^tA is a symmetric positive definite matrix; see the pseudocode from Table 8.1 for details.

Table 8.1: Least squares implementation

```
Function Call:
x = least_squares(A,y)

Input:
A = m × n matrix; m > n
y = column vector of size m

Output:
x = solution to min ||y − Ax||

x = linear_solve_cholesky(A^t A, A^t y);
```

8.1.1 Least squares for implied volatility computation

Consider a European call or put option[1] on an underlying asset whose price is assumed to follow a lognormal model. The implied volatility of the option is the unique value of the volatility parameter σ from the lognormal model that makes the Black–Scholes value of the option equal to the market price of the option.

More precisely, if C_m and P_m are the market prices of a European call option and of a European put option, respectively, with strike K and maturity T on an underlying asset with spot price S paying dividends continuously at the rate q, and assuming that interest rates are constant and equal to r, the implied volatility σ_{imp} corresponding to the price C_m is, by definition, the solution $\sigma = \sigma_{imp}$ to

$$C_{BS}(S, K, T, \sigma, r, q) \;=\; C_m; \tag{8.17}$$

the implied volatility σ_{imp} corresponding to price P_m is the solution $\sigma = \sigma_{imp}$ to

$$P_{BS}(S, K, T, \sigma, r, q) \;=\; P_m. \tag{8.18}$$

Here, $C_{BS}(S, K, T, \sigma, r, q)$ and $P_{BS}(S, K, T, \sigma, r, q)$ are the Black–Scholes values of a call option and of a put option given by (10.77–10.80), i.e.,

$$C_{BS}(S, K, T, \sigma, r, q) \;=\; Se^{-qT}N(d_1) - Ke^{-rT}N(d_2); \tag{8.19}$$

$$P_{BS}(S, K, T, \sigma, r, q) \;=\; Ke^{-rT}N(-d_2) - Se^{-qT}N(-d_1), \tag{8.20}$$

[1]See Section 10.3 for a brief overview of European options.

respectively, where $N(z)$ is the cumulative distribution of the standard normal variable, i.e.,

$$N(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{x^2}{2}} \, dx,$$

and

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r - q + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}; \quad d_2 = d_1 - \sigma\sqrt{T}. \tag{8.21}$$

For any plain vanilla European option, the option price C_m or P_m, the maturity T, the strike K, and the spot price S of the underlying asset are known. However, a continuous dividend yield q for the underlying asset is very rarely quoted in the markets and the interest rate r could be chosen from several different discount curves. As seen below,[2] the least squares method and Put-Call parity can be used to overcome these issues and find the implied volatility using Newton's method.

As an example of how implied volatilities are computed in practice, consider the snapshot from Table 8.2 of the mid prices[3] on March 9, 2012, of the S&P 500 options (ticker symbol SPX)[4] maturing on December 22, 2012. These options are European options and therefore we can use the Black–Scholes framework. Although not needed for this method, the corresponding spot price of the index was 1,370.

Table 8.2: Dec 2012 SPX option prices on 3/9/2012

Call Strike	Price	Put Strike	Price
C1175	225.40	P1175	46.60
C1200	205.55	P1200	51.55
C1225	186.20	P1225	57.15
C1250	167.50	P1250	63.30
C1275	149.15	P1275	70.15
C1300	131.70	P1300	77.70
C1325	115.25	P1325	86.20
C1350	99.55	P1350	95.30
C1375	84.90	P1375	105.30
C1400	71.10	P1400	116.55
C1425	58.70	P1425	129.00
C1450	47.25	P1450	143.20
C1500	29.25	P1500	173.95
C1550	15.80	P1550	210.80
C1575	11.10	P1575	230.90
C1600	7.90	P1600	252.40

Recall that the Put-Call parity states that taking a long position in a European call option and a short position in a European put option with the same strike K and maturity T is equivalent to taking a long position in a forward contract with

[2] This method was implemented in 2010 in Bloomberg terminals, providing a tenfold improvement in the accuracy of implied volatility calculations over the prior method.

[3] The mid price of an option is the average of the bid price and ask price of the option.

[4] More information on SPX options can be found on the Chicago Board Options Exchange (CBOE); see http://www.cboe.com/products/indexopts/spx_spec.aspx

delivery price K and maturity T, and therefore the following relationship between the values C and P of the call and put options must hold for no–arbitrage:

$$C - P \;=\; Se^{-qT} - Ke^{-rT}. \tag{8.22}$$

Let $F = Se^{(r-q)T}$ be the forward price of the asset at time T. Then, $Se^{-qT} = Fe^{-rT}$ and the Put–Call parity (8.22) can be written as

$$C - P \;=\; Fe^{-rT} - Ke^{-rT}. \tag{8.23}$$

Denote by $disc = e^{-rT}$ the discount factor, and let $PVF = Fe^{-rT}$ be the present value of the forward price. Then, (8.23) is the same as

$$C - P \;=\; PVF - K \cdot disc. \tag{8.24}$$

The data from Table 8.2 provides call and put options values for 16 different strikes. From (8.24), it follows that the values of PVF and $disc$ can be obtained by solving a least square problem $y \approx Ax$, see (8.2), with $x = \begin{pmatrix} PVF \\ disc \end{pmatrix}$ and with the following 16×2 matrix A and the following 16×1 column vector y corresponding to $C - P$ for each strike:

$$A \;=\; \begin{pmatrix} 1 & -1175 \\ 1 & -1200 \\ 1 & -1225 \\ 1 & -1250 \\ 1 & -1275 \\ 1 & -1300 \\ 1 & -1325 \\ 1 & -1350 \\ 1 & -1375 \\ 1 & -1400 \\ 1 & -1425 \\ 1 & -1450 \\ 1 & -1500 \\ 1 & -1550 \\ 1 & -1575 \\ 1 & -1600 \end{pmatrix}; \quad y \;=\; \begin{pmatrix} 178.80 \\ 154.00 \\ 129.05 \\ 104.20 \\ 79.00 \\ 54.00 \\ 29.05 \\ 4.25 \\ -20.40 \\ -45.45 \\ -70.30 \\ -95.95 \\ -144.70 \\ -195.00 \\ -219.80 \\ -244.50 \end{pmatrix}.$$

The solution $x = (A^tA)^{-1}A^ty$ to this least squares problem, see (8.16), is computed as $x = \text{least\_squares}(A,y)$ by using the routine from Table 8.1. We find that

$$x \;=\; \begin{pmatrix} PVF \\ disc \end{pmatrix} \;=\; \begin{pmatrix} 1349.54 \\ 0.9964 \end{pmatrix}. \tag{8.25}$$

In order to use the values $PVF = 1349.54$ and $disc = 0.9964$ obtained above to compute implied volatilities, we first show that the Black–Scholes formulas (8.19–8.21) can be written in terms of PVF and $disc$ without any dependence on r, q, or the spot price S; see (8.34–8.36).

Recall that $PVF = Fe^{-rT}$. Since $F = Se^{(r-q)T}$, it follows that

$$\begin{aligned} PVF &= Fe^{-rT} \;=\; Se^{(r-q)T} \cdot e^{-rT} \\ &= Se^{-qT}. \end{aligned} \tag{8.26}$$

Also, recall that

$$disc \;=\; e^{-rT}. \tag{8.27}$$

Then, using (8.26) and (8.27), the Black–Scholes formulas (8.19) and (8.20) can be written as

$$\begin{aligned} C_{BS} &= PVF \cdot N(d_1) - K \cdot disc \cdot N(d_2); & (8.28)\\ P_{BS} &= K \cdot disc \cdot N(-d_2) - PVF \cdot N(-d_1). & (8.29) \end{aligned}$$

Moreover,

$$\begin{aligned} \ln\left(\frac{S}{K}\right) + (r-q)T &= \ln\left(\frac{S}{K}\right) + \ln\left(e^{(r-q)T}\right) \;=\; \ln\left(\frac{S}{K}\cdot e^{(r-q)T}\right) & (8.30)\\ &= \ln\left(\frac{Se^{-qT}}{K}\cdot e^{rT}\right) \;=\; \ln\left(\frac{Se^{-qT}}{Ke^{-rT}}\right)\\ &= \ln\left(\frac{PVF}{K\cdot disc}\right), & (8.31) \end{aligned}$$

where for (8.30) we used the facts that $\ln(e^x) = x$ for any x and $\ln(a)+\ln(b) = \ln(ab)$ for any $a, b > 0$, and for (8.31) we used (8.26) and (8.27), i.e., $Se^{-qT} = PVF$ and $e^{-rT} = disc$.

Then, from (8.21) and (8.31), we obtain that

$$\begin{aligned} d_1 &= \frac{\ln\left(\frac{S}{K}\right) + (r-q)T}{\sigma\sqrt{T}} + \frac{\frac{\sigma^2}{2}T}{\sigma\sqrt{T}} \;=\; \frac{\ln\left(\frac{PVF}{K\cdot disc}\right)}{\sigma\sqrt{T}} + \frac{\sigma\sqrt{T}}{2}; & (8.32)\\ d_2 &= d_1 - \sigma\sqrt{T} \;=\; \frac{\ln\left(\frac{PVF}{K\cdot disc}\right)}{\sigma\sqrt{T}} - \frac{\sigma\sqrt{T}}{2}. & (8.33) \end{aligned}$$

From (8.28), (8.29), (8.32), and (8.33), we conclude that the Black–Scholes option values can be written as functions of PVF, $disc$, K, T, and σ as follows:

$$\begin{aligned} C_{BS}(PVF, disc, K, T, \sigma) &= PVF \cdot N(d_1) - K \cdot disc \cdot N(d_2); & (8.34)\\ P_{BS}(PVF, disc, K, T, \sigma) &= K \cdot disc \cdot N(-d_2) - PVF \cdot N(-d_1), & (8.35) \end{aligned}$$

where

$$d_1 \;=\; \frac{\ln\left(\frac{PVF}{K\cdot disc}\right)}{\sigma\sqrt{T}} + \frac{\sigma\sqrt{T}}{2}; \quad d_2 \;=\; \frac{\ln\left(\frac{PVF}{K\cdot disc}\right)}{\sigma\sqrt{T}} - \frac{\sigma\sqrt{T}}{2}. \tag{8.36}$$

Since K and T are known and PVF and $disc$ have been computed using least squares, see (8.25), we can use Newton's method to solve either

$$C_{BS}(PVF, disc, T, \sigma) \;=\; C_m \tag{8.37}$$

or

$$P_{BS}(PVF, disc, T, \sigma) \;=\; P_m \tag{8.38}$$

for $\sigma = \sigma_{imp}$ for every option from Table 8.2.

For call options, we look at (8.37) as a function of only one variable, σ. Then, finding the implied volatility for a call option requires solving the nonlinear problem

$$f_C(x) \;=\; 0, \tag{8.39}$$

using Newton's method, where $x = \sigma$ and

$$f_C(x) \;=\; PVF \cdot N(d_1(x)) - K \cdot disc \cdot N(d_2(x)) \;-\; C_m, \tag{8.40}$$

with $d_1(x)$ and $d_2(x)$ given by (8.36), i.e.,

$$d_1(x) \;=\; \frac{\ln\left(\frac{PVF}{K \cdot disc}\right)}{x\sqrt{T}} + \frac{x\sqrt{T}}{2}; \quad d_2(x) \;=\; \frac{\ln\left(\frac{PVF}{K \cdot disc}\right)}{x\sqrt{T}} - \frac{x\sqrt{T}}{2}.$$

The value of x thus computed is the implied volatility σ_{imp}.

Note that differentiating the function $f_C(x)$ with respect to x is the same as computing the vega of the call option, which is equal to

$$\text{vega}_C \;=\; Se^{-qT}\sqrt{\frac{T}{2\pi}}\,\exp\left(-\frac{(d_1(x))^2}{2}\right); \tag{8.41}$$

see (10.85). Then, since $Se^{-qT} = PVF$, we obtain that

$$f_C'(x) \;=\; PVF\sqrt{\frac{T}{2\pi}}\,\exp\left(-\frac{(d_1(x))^2}{2}\right). \tag{8.42}$$

The Newton's method recursion for solving (8.39) is

$$x_{k+1} \;=\; x_k - \frac{f_C(x_k)}{f_C'(x_k)}, \tag{8.43}$$

where the functions $f_C(x)$ and $f_C'(x)$ are given by (8.40) and (8.42), respectively.

Similarly, for put options we look at (8.38) as a function of only one variable, σ. Then, finding the implied volatility for a put option requires solving the nonlinear problem

$$f_P(x) \;=\; 0, \tag{8.44}$$

where $x = \sigma$ and

$$f_P(x) \;=\; K \cdot disc \cdot N(-d_2(x)) - PVF \cdot N(-d_1(x)) \;-\; P_m, \tag{8.45}$$

with $d_1(x)$ and $d_2(x)$ given by (8.36), i.e.,

$$d_1(x) \;=\; \frac{\ln\left(\frac{PVF}{K \cdot disc}\right)}{x\sqrt{T}} + \frac{x\sqrt{T}}{2}; \quad d_2(x) \;=\; \frac{\ln\left(\frac{PVF}{K \cdot disc}\right)}{x\sqrt{T}} - \frac{x\sqrt{T}}{2}.$$

Differentiating the function $f_P(x)$ with respect to x is the same as computing the vega of the put option, which is equal to

$$\text{vega}_P \;=\; Se^{-qT}\sqrt{\frac{T}{2\pi}}\,\exp\left(-\frac{(d_1(x))^2}{2}\right), \tag{8.46}$$

respectively; see (10.86) and (10.85).

Then, since $Se^{-qT} = PVF$, we obtain that

$$f_P'(x) \;=\; PVF\sqrt{\frac{T}{2\pi}}\,\exp\left(-\frac{(d_1(x))^2}{2}\right). \tag{8.47}$$

Note that $f'_P(x) = f'_C(x)$, since $\text{vega}_P = \text{vega}_C$.

The Newton's method recursion for solving (8.44) is

$$x_{k+1} \;=\; x_k - \frac{f_P(x_k)}{f'_P(x_k)}, \tag{8.48}$$

where the functions $f_P(x)$ and $f'_P(x)$ are given by (8.45) and (8.47), respectively.

A good initial guess for Newton's method is 25% volatility, i.e., $x_0 = 0.25$, and the algorithm is stopped when two consecutive approximations in Newton's method are within 10^{-6} of each other; see the pseudocode from Table 8.3 for finding the implied volatility for both call and put options, i.e., for solving either (8.43) or (8.48).

Table 8.3: Pseudocode for computing implied volatility

```
Input:
V_m = option price
      // V_m = C_m for call implied vol; V_m = P_m for put implied vol
K = strike price of the option
T = maturity of the option
PVF = present value of the forward price of the underlying asset
disc = discount factor corresponding to time T
tol = tolerance for Newton's method convergence
f_BS(x) = Black–Scholes option value; x = volatility
      // f_BS(x) = f_C(x) for calls; f_BS(x) = f_P(x) for puts
      // f'_BS(x) = f'_C(x) = f'_P(x)

Output:
x_new = implied volatility

x_0 = 0.25;            // initial guess: 25% volatility
x_new = x_0; x_old = x_0 − 1; tol = 10^(−6)
while (|x_new − x_old| > tol)
   x_old = x_new
   x_new = x_old − (f_BS(x_old) − V_m) / f'_BS(x_old)
end
```

For the options from Table 8.2, note that the options maturity is $T = \frac{199}{252}$, i.e., the ratio of 199, the number of trading days between March 9, 2012 and December 22, 2012, and 252, the total number of trading days in a year. Using the values $PVF = 1349.54$ and disc $= 0.9964$ computed using least squares and the Newton's method from Table 8.3, we obtain the implied volatilies from Table 8.4.

A consequence of the Put–Call parity is that the theoretical values of implied volatilities of calls and puts with the same strike are equal. Note that, indeed, the implied volatilities from Table 8.4 corresponding to calls and puts with the same strike are nearly identical.

Table 8.4: Implied volatilies for SPX options

Strike	Implied Vol Call	Implied Vol Put	Strike	Implied Vol Call	Implied Vol Put
1175	25.73%	25.72%	1375	19.69%	19.66%
1200	24.96%	24.92%	1400	18.94%	18.94%
1225	24.19%	24.16%	1425	18.26%	18.25%
1250	23.44%	23.40%	1450	17.53%	17.68%
1275	22.63%	22.65%	1500	16.34%	16.24%
1300	21.86%	21.91%	1550	15.05%	15.08%
1325	21.15%	21.20%	1575	14.48%	14.47%
1350	20.41%	20.43%	1600	14.13%	14.02%

8.2 Linear regression: ordinary least squares for time series data

Linear regression for time series data (also called ordinary least squares for time series data) requires finding the best approximation of the time series data of a random variable by a linear combination of the time series data of other random variables and a constant vector.

Let Y and $X_1, X_2, \ldots, X_n$ be random variables given by time series data at N data points t_i, $i = 1 : N$. Assume that the sample covariance matrix corresponding to the time series data for $X_1, X_2, \ldots, X_n$ is nonsingular, or, equivalently, that the column vectors of the time series data for $X_1, X_2, \ldots, X_n$ and the $N \times 1$ column vector with all entries equal to 1 are linearly independent; cf. Theorem 7.2.

We look for the best linear approximation of the time series data for Y by a linear combination of the time series data for $X_1, X_2, \ldots, X_n$ plus a constant vector, i.e., we look for constants $a, b_1, b_2, \ldots, b_n$ such that

$$\begin{aligned}
\begin{pmatrix} Y(t_1) \\ Y(t_2) \\ \vdots \\ Y(t_N) \end{pmatrix} &\approx \begin{pmatrix} a + \sum_{k=1}^n b_k X_k(t_1) \\ a + \sum_{k=1}^n b_k X_k(t_2) \\ \vdots \\ a + \sum_{k=1}^n b_k X_k(t_N) \end{pmatrix} \\
\Longleftrightarrow \quad \begin{pmatrix} Y(t_1) \\ Y(t_2) \\ \vdots \\ Y(t_N) \end{pmatrix} &\approx \begin{pmatrix} a \\ a \\ \vdots \\ a \end{pmatrix} + \sum_{k=1}^n b_k \begin{pmatrix} X_k(t_1) \\ X_k(t_2) \\ \vdots \\ X_k(t_N) \end{pmatrix} \\
\Longleftrightarrow \quad \begin{pmatrix} Y(t_1) \\ Y(t_2) \\ \vdots \\ Y(t_N) \end{pmatrix} &\approx \begin{pmatrix} 1 & X_1(t_1) & \ldots & X_n(t_1) \\ 1 & X_1(t_2) & \ldots & X_n(t_2) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_1(t_N) & \ldots & X_n(t_N) \end{pmatrix} \begin{pmatrix} a \\ b_1 \\ \vdots \\ b_n \end{pmatrix}. \quad (8.49)
\end{aligned}$$

Denote by T_Y and T_{X_k} the column vectors of the time series data for the random variables Y and X_k, for $k = 1 : n$, respectively, i.e.,

$$T_Y = (Y(t_i))_{i=1:N}; \quad (8.50)$$