Optimization in several variables I, first and second order conditions

This note covers the basics of (unconstrained) optimization. In particular, we will see

- *why critical points are critical, and*
- *why the second derivative test works.*

1. Local extreme values and critical points.

1.1 Local extreme values.

The value $z_0 = f(x_0, y_0)$ is called a ***local minimum value*** of the function if there is a number $\delta > 0$ such that

$$f(x,y) \geq f(x_0,y_0) \tag{1}$$

for all points (x, y) satisfying $|x - x_0| < \delta$ and $|y - y_0| < \delta$. Likewise, $z_1 = f(x_1, y_1)$ is a ***local maximum value*** if

$$f(x,y) \leq f(x_1,y_1) \tag{2}$$

for all points (x, y) satisfying $|x - x_1| < \delta$ and $|y - y_1| < \delta$, for some number $\delta > 0$.† The local maximum and minimum values of a function are also called ***local extreme values***, when the distinction between maximum and minimum is not important.

More generally, $\tilde{w} = f(\tilde{x}_1, \ldots, \tilde{x}_n)$ is a local maximum (or minimum) value if and only if there is some number $\delta > 0$ such that

$$f(x_1,\ldots,x_n) \leq f(\tilde{x}_1,\ldots,\tilde{x}_n) \qquad (\text{or } f(x_1,\ldots,x_n) \geq f(\tilde{x}_1,\ldots,\tilde{x}_n),\ \text{respectively})$$

for all points $(x_1, \ldots, x_n)$ satisfying $|x_j - \tilde{x}_j| < \delta$, for $j = 1, 2, \ldots, n$.

1.2 First order conditions.

By studying the *first order Taylor approximation* of the function $z = f(x, y)$, we can identify the points (x_0, y_0) where local extreme values *might* occur. Recall that the first order approximation can be written

$$f(x,y) - f(x_0,y_0) \approx f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0), \tag{3}$$

where the approximation is accurate for all points (x, y) *sufficiently close* to (x_0, y_0), i.e., for all points satisfying $|x - x_0| < \delta$ and $|y - y_0| < \delta$ for some positive number δ.‡

Suppose that $f_x(x_0, y_0) > 0$, then by choosing $y = y_0$ and $x_1 > x_0$, the approximation (3) tells us that

$$f(x_1,y_0) - f(x_0,y_0) \approx f_x(x_0,y_0)(x_1-x_0) > 0,$$

† The condition '$|x - x_1| < \delta$ and $|y - y_1| < \delta$, for some number $\delta > 0$' is a more precise way of saying 'sufficiently close to (x_1, y_1)'.

‡ For our purposes in this note, the precise value of δ is not important. The important thing is that δ exists.

as long as x_1 is sufficiently close to x_0. This means that $f(x_0, y_0)$ cannot possibly be a local maximum value, since there are points arbitrarily close to (x_0, y_0) where the function takes *larger* values.

Still assuming that $f_x(x_0, y_0) > 0$ and staying with $y = y_0$, but choosing $x_1 < x_0$, we see that

$$f(x_1, y_0) - f(x_0, y_0) \approx f_x(x_0, y_0)(x_1 - x_0) < 0,$$

as long as x_1 is sufficiently close to x_0. This implies that $f(x_0, y_0)$ cannot possibly be a local minimum value, since there are points arbitrarily close to (x_0, y_0) where the function takes *smaller* values.

The preceding two paragraphs show that if $f_x(x_0, y_0) > 0$, then $f(x_0, y_0)$ is not a local extreme value. An analogous line of reasoning shows that if $f_x(x_0, y_0) < 0$, then $f(x_0, y_0)$ is not a local extreme value. Furthermore, there is nothing special about the variable x — if $f_y(x_0, y_0) \neq 0$, then the same reasoning shows that $f(x_0, y_0)$ is not a local extreme value. All of this implies the following basic fact:

*if $f(x_0, y_0)$ **is** a local extreme value, then $f_x(x_0, y_0) = 0$ and $f_y(x_0, y_0) = 0$.*

These arguments work just as well for a function of 17 variables as a function of two variables, so we can draw the following general conclusion.

Fact 1.

If $f(\tilde{x}_1, \ldots, \tilde{x}_n)$ is a local extreme value of the function $f(x_1, \ldots, x_n)$, then

$$f_{x_j}(\tilde{x}_1, \ldots, \tilde{x}_n) = 0 \quad \textit{for each of the } n \textit{ variables } x_1, \ldots, x_n.$$

This fact leads to the the following definition.

Definition 1.

*A point $(\tilde{x}_1, \ldots, \tilde{x}_n)$ is called a **critical point** (or a **stationary point**) of the function $f(x_1, \ldots, x_n)$ if the system of n equations*

$$\left.\begin{array}{rcl} f_{x_1}(\tilde{x}_1, \ldots, \tilde{x}_n) & = & 0 \\ f_{x_2}(\tilde{x}_1, \ldots, \tilde{x}_n) & = & 0 \\ & \vdots & \\ f_{x_n}(\tilde{x}_1, \ldots, \tilde{x}_n) & = & 0 \end{array}\right\} \tag{4}$$

*are all satisfied. The value of the function at a critical point, $f(\tilde{x}_1, \ldots, \tilde{x}_n)$, is called a **critical value** (or a **stationary value**).*

*The equations in (4) are called the **first order conditions** for $f(\tilde{x}_1, \ldots, \tilde{x}_n)$ to be a local extreme value.*

The first order conditions must be satisfied if $f(\tilde{x}_1, \ldots, \tilde{x}_n)$ is a local extreme value, but they do *not* guarantee that the critical value is a local extreme value.[§]

[§]Mathematicians call these *necessary* conditions, but they are not *sufficient* conditions.

Example 1. The function $f(x,y) = x^2 - y^2$ has exactly one critical point, $(0,0)$, since the two first order conditions for this function

$$\begin{array}{rcrcl} f_x &=& 2x &=& 0 \\ f_y &=& -2y &=& 0 \end{array}$$

have the unique solution $x = 0$ and $y = 0$, and the critical value is $f(0,0) = 0$.

On the other hand, $f(x,0) = x^2 > 0$ for any $x \neq 0$, however small $|x|$ is. This means that $f(0,0) = 0$ is *not* a local maximum, since the function assumes larger values at points that are arbitrarily close to $(0,0)$.

Likewise, $f(0,y) = -y^2 < 0$ for any $y \neq 0$, however small $|y|$ is, which means that the function takes smaller values at points arbitrarily close to $(0,0)$ as well. This means that $f(0,0)$ is not a local minimum either.

Therefore, the critical value $f(0,0)$ is not a local extreme value in this example. ■

Example 2. The first order conditions for the function $g(x,y) = x^2 + y^2 - 2x + 6y + 10$ are

$$\begin{array}{rcrcl} g_x &=& 2x-2 &=& 0 \\ g_y &=& 2y+6 &=& 0. \end{array}$$

The first equation has the solution $x_0 = 1$ and the second equation has the solution $y_0 = -3$. So $(x_0, y_0) = (1, -3)$ is the only critical point for this function, and the critical value is $g(1,-3) = 0$.

A little bit of algebraic manipulation (completing the square separately in x and y) shows that

$$g(x,y) = x^2 + y^2 - 2x + 6y + 10 = (x-1)^2 + (y+3)^2.$$

This means that $g(x,y) \geq 0$ for all (x,y), and that $g(x,y) = 0$ *only* at the critical point $(1,-3)$. In other words, if $(x,y) \neq (1,-3)$, then

$$g(x,y) > 0 = g(1,-3),$$

so $g(1,-3)$ is a local minimum value, and in fact that it is the *global* minimum value, since it is the smallest value of the function over all. ■

The main goal of the remainder of this note is to explain the *second derivative test* which is described (for functions of two variables) in Section 17.6 of the textbook. The second derivative test provides a way of deciding whether a critical value is a local maximum value, a local minimum value or neither.

2. *A necessary digression: quadratic forms.*

The two examples in the previous section illustrate the fact that analyzing the critical values of quadratic functions is relatively easy, requiring only basic algebra (i.e., completing the square) and the fact that $u^2 > 0$ for all $u \neq 0$.

To understand the second derivative test, we need only analyze the behavior of quadratic functions that have no constant term and no linear terms. In the case of two variables, they look like this

$$Q(u,v) = au^2 + buv + cv^2,$$

and in the case of three variables, they have the form

$$Q(u,v,w) = au^2 + bv^2 + cw^2 + duv + euw + fvw.$$

This type of function is called a ***quadratic form***.

As it turns out, what we need to know about quadratic forms is when such a form produces only positive values, only negative values, or when it produces both positive and negative values. This behavior is completely determined by the *coefficients* of the quadratic form, as I will illustrate for forms in two and three variables. Moreover, we don't need calculus to do this analysis.

2.1 ***The case of two variables.***

Consider the form $Q(u,v) = au^2 + buv + cv^2$. If $a \neq 0$, then a short sequence of algebraic manipulations shows that

$$Q(u,v) = a\left(u + \frac{b}{2a}v\right)^2 + \left(\frac{4ac-b^2}{4a}\right)v^2 = a\left[\left(u + \frac{b}{2a}v\right)^2 + D\left(\frac{v}{2a}\right)^2\right]. \tag{5}$$

The expression $D = 4ac - b^2$ that appears in this expression is called the ***discriminant*** of the form $Q(u,v)$.

There are three possibilities to consider: $D > 0$, $D < 0$ and $D = 0$.

(a) If $D > 0$, then

$$\left[\left(u + \frac{b}{2a}v\right)^2 + D\left(\frac{v}{2a}\right)^2\right] > 0$$

for all $(u,v) \neq (0,0)$. This means that the nature of $Q(u,v)$ depends on a in this case.

i. If $a > 0$, then $Q(u,v) > 0$ for all $(u,v) \neq (0,0)$.

ii. If $a < 0$, then $Q(u,v) < 0$ for all $(u,v) \neq (0,0)$.

(b) If $D < 0$, then the expression $S(u,v) = \left[\left(u + \frac{b}{2a}v\right)^2 + D\left(\frac{v}{2a}\right)^2\right]$ takes both positive and negative values. For example, if $v = 0$ and $u \neq 0$, then $S(u,v) = u^2 > 0$. On the other hand, if $v \neq 0$ and $u = -bv/2a$, then $S(u,v) = D(v/2a)^2 < 0$. Since $Q(u,v) = a \cdot S(u,v)$, it follows that in this case, $Q(u,v)$ also takes both positive and negative values.

(c) If $D = 0$, then $Q(u,v) = a\,(u + bv/2a)^2$. In this case, if $a > 0$, then $Q(u,v) \geq 0$ for all (u,v), and if $a < 0$, then $Q(u,v) \leq 0$ for all (u,v). At first glance, this case seems to be like case (a), but there is a big difference.

If $D = 0$, *then* $Q(u,v) = 0$ *for all points* (u,v) *satisfying* $u = -bv/2a$.

In cases (a) and (b), $Q(u,v) = 0$ only if $(u,v) = (0,0)$. This difference makes the case $D = 0$ unhelpful in the second derivative test.

What happens if $a = 0$? First of all, the quadratic form simplifies to

$$Q(u,v) = buv + cv^2,$$

and second the discriminant simplifies to $D = -b^2$. If $b = 0$ as well, then $D = 0$ and $Q(u,v) = cv^2$, and this reduces to a special case of (c), above, with the same conclusions.

If $b \neq 0$, then $D < 0$ and once again $Q(u,v)$ takes both positive and negative values. I will leave this for you to check.

The key points of the analysis above are summarized in the following fact.

Fact 2.
If $Q(u, v) = au^2 + buv + cv^2$ *and* $D = 4ac - b^2 \neq 0$*, then*

1. *If* $D > 0$ *and* $a > 0$*, then* $Q(u, v) > 0$ *for all* $(u, v) \neq (0, 0)$*.*

2. *If* $D > 0$ *and* $a < 0$*, then* $Q(u, v) < 0$ *for all* $(u, v) \neq (0, 0)$*.*

3. *If* $D < 0$*, then* $Q(u, v)$ *produces both positive and negative values.*

2.2 *The case of three variables.*

As you can imagine, with three variables, there are more cases to consider in the algebraic analysis. I'll spare you the details, and merely summarize the main points in Fact 3.

Fact 3.
Let $Q(u, v, w) = au^2 + bv^2 + cw^2 + duv + euw + fvw$*, and set*

$$D_2 = 4ab - d^2 \quad \text{and} \quad D_3 = 4abc + def - be^2 - af^2 - cd^2.$$

1. *If* $a > 0$*,* $D_2 > 0$ *and* $D_3 > 0$*, then* $Q(u, v, w) > 0$ *for all* $(u, v, w) \neq (0, 0, 0)$*.*

2. *If* $a < 0$*,* $D_2 > 0$ *and* $D_3 < 0$*, then* $Q(u, v, w) < 0$ *for all* $(u, v, w) \neq (0, 0, 0)$*.*

3. *If* $D_2 < 0$*, or if* $D_2 > 0$ *but* $a \cdot D_3 < 0$*, then* $Q(u, v, w)$ *yields both positive and negative values.*

I won't discuss the case of quadratic forms with more than three variables here, beyond saying that Facts 2 and 3 generalize directly to give similar criteria.

3. *The second order conditions.*

We have already seen that for $f(\tilde{x}_1, \ldots, \tilde{x}_n)$ to be a local extreme value, it is necessary that the point $(\tilde{x}_1, \ldots, \tilde{x}_n)$ be a *critical point* of the function. I.e., the point must satisfy the *first order conditions* in (4). On the other hand, we also saw that the first order conditions, by themselves, are not enough to *guarantee* that the critical value $f(\tilde{x}_1, \ldots, \tilde{x}_n)$ is a local extreme value.

In the case that $f(x_1, \ldots, x_n)$ is a quadratic function, it is relatively easy to analyze the nature of the critical value by completing the square.¶ To extend this same analysis to general (nonquadratic) functions, we will look at the second order Taylor approximation, centered at the critical point $(\tilde{x}_1, \ldots, \tilde{x}_n)$.

To keep the explanation relatively simple, I will focus on the case of two variables.‖ The problem that we want to solve is the following:

> *Suppose that* (x^*, y^*) *is a critical point of the function* $f(x, y)$*. How can we tell if* $f(x^*, y^*)$ *is bigger than or smaller than all the nearby values of* $f(x, y)$*?*

¶Though what I mean by 'easy' may not jibe with what you mean by 'easy', and in any case, easy or not, the analysis is admittedly somewhat tedious.

‖The explanation for functions of three or more variables is essentially the same, but the expressions get longer and longer, and possibly more confusing.

In other words, we want a test that will tell us if $f(x,y) < f(x^*,y^*)$ for all points (x,y) that are sufficiently close to (x^*,y^*), if $f(x,y) > f(x^*,y^*)$ for all points (x,y) that are sufficiently close to (x^*,y^*), or if neither is true. To produce such a test, we will use the quadratic Taylor polynomial for $f(x,y)$, centered at the critical point (x^*,y^*).

If (x^*,y^*) is a critical point of the function $f(x,y)$, then

$$f_x(x^*,y^*) = f_y(x^*,y^*) = 0$$

and the quadratic Taylor polynomial for f, centered at (x^*,y^*) simplifies to

$$T_2(x,y) = f(x^*,y^*) + \frac{f_{xx}(x^*,y^*)}{2}(x-x^*)^2 + f_{xy}(x^*,y^*)(x-x^*)(y-y^*) + \frac{f_{yy}(x^*,y^*)}{2}(y-y^*)^2$$

because the linear terms $f_x(x^*,y^*)(x-x^*)$ and $f_y(x^*,y^*)(y-y^*)$ both vanish!

Now, if $(x,y) \approx (x^*,y^*)$, then $f(x,y) \approx T_2(x,y)$, so

$$f(x,y) - f(x^*,y^*) \approx T_2(x,y) - f(x^*,y^*),$$

and this means that

$$f(x,y) - f(x^*,y^*) \approx \frac{f_{xx}(x^*,y^*)}{2}(x-x^*)^2 + f_{xy}(x^*,y^*)(x-x^*)(y-y^*) + \frac{f_{yy}(x^*,y^*)}{2}(y-y^*)^2. \quad (6)$$

The expression on the righthand side of Equation (6) is a *quadratic form*, $Q_f(u,v)$, in the two variables $u = x - x^*$ and $v = y - y^*$, with coefficients

$$a = \frac{f_{xx}(x^*,y^*)}{2}, \quad b = f_{xy}(x^*,y^*) \text{ and } c = \frac{f_{yy}(x^*,y^*)}{2},$$

that is

$$Q_f(u,v) = au^2 + buv + cv^2 = \frac{f_{xx}(x^*,y^*)}{2}(x-x^*)^2 + f_{xy}(x^*,y^*)(x-x^*)(y-y^*) + \frac{f_{yy}(x^*,y^*)}{2}(y-y^*)^2.$$

Furthermore, it follows from (6) that we have the following three cases.

Q1. If $Q_f(u,v)) > 0$ for all points $(u,v) \neq (0,0)$, then

$$f(x,y) > f(x^*,y^*)$$

for all points (x,y) that are *sufficiently close* to (x^*,y^*)), in which case $f^* = f(x^*,y^*)$ is a *relative minimum value.*

Q2. If $Q_f(u,v)) < 0$ for all points $(u,v) \neq (0,0)$, then

$$f(x,y) < f(x^*,y^*)$$

for all points (x,y) that are *sufficiently close* to (x^*,y^*)), in which case $f^* = f(x^*,y^*)$ is a *relative maximum value.*

Q3. If $Q_f(u,v)$ takes both positive and negative values, then there will be points (x,y) close to (x^*,y^*) such that $f(x,y) > f(x^*,y^*)$ and there will also be points (x,y) close to (x^*,y^*) such that $f(x,y) < f(x^*,y^*)$. In this case, $f^* = f(x^*,y^*)$ is neither a minimum nor a maximum value.

Since the coefficients of Q_f are given by the values of the second order derivatives of the function $f(x, y)$ at the critical point (x^*, y^*), these second derivative values determine which of the three cases, above, occur (if any). To complete the analysis, we use Fact 2 and for this we need the discriminant of $Q_f(u, v)$ (which depends on the critical point (x^*, y^*)):

$$\begin{aligned} D_f(x^*, y^*) &= 4ac - b^2 = 4\left(\frac{f_{xx}(x^*, y^*)}{2}\right)\left(\frac{f_{yy}(x^*, y^*)}{2}\right) - f_{xy}(x^*, y^*)^2 \\ &= f_{xx}(x^*, y^*) f_{yy}(x^*, y^*) - f_{xy}(x^*, y^*)^2 \end{aligned}$$

Fact 4.

Suppose that (x^*, y^*) satisfies the first order conditions $f_x(x^*, y^*) = 0$ and $f_y(x^*, y^*) = 0$, then:

1. *If $D_f(x^*, y^*) > 0$ and $f_{xx}(x^*, y^*) > 0$, then $f^* = f(x^*, y^*)$ is a local* ***minimum*** *value.*

2. *If $D_f(x^*, y^*) > 0$ and $f_{xx}(x^*, y^*) < 0$, then $f^* = f(x^*, y^*)$ is a local* ***maximum*** *value.*

3. *If $D_f(x^*, y^*) < 0$, then $f^* = f(x^*, y^*)$ is neither a local minimum value nor a local maximum value.*

The conditions in **1.** and **2.** are called the ***second order conditions*** for a local minimum value or a local maximum value, respectively, at (x^*, y^*).

3.1 *Three variables*

In the case of a critical point (x^*, y^*, z^*) of a function of three variables, $g(x, y, z)$, Taylor's approximation, centered at (x^*, y^*, z^*) reduces to

$$\begin{aligned} g(x, y, z) - g(x^*, y^*, z^*) \approx\ & a(x - x^*)^2 + b(y - y^*)^2 + c(z - z^*)^2 \\ & + d(x - x^*)(y - y^*) + e(x - x^*)(z - z^*) + f(y - y^*)(z - z^*), \end{aligned}$$

where

$$a = \frac{g_{xx}(x^*, y^*, z^*)}{2}, \quad b = \frac{g_{yy}(x^*, y^*, z^*)}{2} \text{ and } c = \frac{g_{zz}(x^*, y^*, z^*)}{2},$$

and

$$d = g_{xy}(x^*, y^*, z^*), \quad e = g_{xz}(x^*, y^*, z^*) \text{ and } f = g_{yz}(x^*, y^*, z^*),$$

as you can check.

This means that

$$g(x, y, z) - g(x^*, y^*, z^*) \approx Q_g(u, v, w) = au^2 + bv^2 + cw^2 + duv + euw + fvw,$$

where $u = (x - x^*)$, $v = (y - y^*)$ and $w = (z - z^*)$.

The reasoning we employed in the two variable case is just as valid in this one, namely

Q1. If $Q_g(u, v, w) > 0$ for all points $(u, v, w) \neq (0, 0, 0)$, then

$$g(x, y, z) > g(x^*, y^*, z^*)$$

for all points (x, y, z) that are *sufficiently close* to (x^*, y^*, z^*), in which case $g^* = g(x^*, y^*, z^*)$ is a *relative minimum value.*

Q2. If $Q_g(u, v, w) < 0$ for all points $(u, v, w) \neq (0, 0, 0)$, then

$$g(x, y, z) < g(x^*, y^*, z^*)$$

for all points (x, y, z) that are *sufficiently close* to (x^*, y^*, z^*), in which case $g^* = g(x^*, y^*, z^*)$ is a *relative maximum value.*

Q3. If $Q_g(u, v, w)$ takes both positive and negative values, then there will be points (x, y, z) close to (x^*, y^*, z^*) such that $g(x, y, z) > g(x^*, y^*, z^*)$ and there will also be points (x, y, z) close to (x^*, y^*, z^*) such that $g(x, y, z) < g(x^*, y^*, z^*)$. In this case $g^* = g(x^*, y^*, z^*)$ is neither a relative minimum value nor a relative maximum value.

Motivated by these observations and Fact 3, we form the discriminants

$$D_2(x, y, z) = g_{xx}(x, y, z)g_{yy}(x, y, z) - g_{xy}(x, y, z)^2$$

and

$$\begin{aligned} D_3(x, y, z) = {} & g_{xx}(x, y, z)g_{yy}(x, y, z)g_{zz}(x, y, z) + 2g_{xy}(x, y, z)g_{xz}(x, y, z)g_{yz}(x, y, z) \\ & - g_{xx}(x, y, z)g_{yz}(x, y, z)^2 - g_{yy}(x, y, z)g_{xz}(x, y, z)^2 - g_{zz}(x, y, z)g_{xy}(x, y, z)^2. \end{aligned}$$

We can now state the second derivative test for functions of three variables.

Fact 5.

Suppose that (x^*, y^*, z^*) satisfies the first order conditions

$$g_x(x^*, y^*, z^*) = 0, \quad g_y(x^*, y^*, z^*) = 0 \text{ and } g_z(x^*, y^*, z^*) = 0.$$

Then:

1. *If $D_3(x^*, y^*, z^*) > 0$, $D_2(x^*, y^*, z^*) > 0$ and $g_{xx}(x^*, y^*, z^*) > 0$, then $g^* = g(x^*, y^*, z^*)$ is a local minimum value.*

2. *If $D_3(x^*, y^*, z^*) < 0$, $D_2(x^*, y^*, z^*) > 0$ and $g_{xx}(x^*, y^*, z^*) < 0$, then $g^* = g(x^*, y^*, z^*)$ is a local maximum value.*

3. *If $D_2(x^*, y^*, z^*) < 0$, or if $D_2(x^*, y^*, z^*) > 0$ but $D_3(x^*, y^*, z^*) \cdot g_{xx}(x^*, y^*, z^*) < 0$, then $g(x^*, y^*, z^*)$ is neither a local minimum value nor a local maximum value.*

4. *Examples*

Example 3. Let $f(x, y) = 5x^3 - 4xy + y^2 + x + 2$. The first order conditions are

$$\begin{aligned} f_x &= 15x^2 - 4y + 1 = 0 \\ f_y &= -4x + 2y = 0. \end{aligned}$$

The condition $f_y = 0$ implies that $y = 2x$, and substituting this into the condition $f_x = 0$ gives

$$15x^2 - 8x + 1 = 0 \implies x = \frac{8 \pm \sqrt{64 - 60}}{30} \implies x = \frac{1}{3} \text{ or } x = -\frac{1}{5}.$$

This means that the critical points are $(x_1, y_1) = (1/3, 2/3)$ and $(x_2, y_2) = (-1/5, -2/5)$.

The second derivatives of f are $f_{xx} = 30x$, $f_{yy} = 2$ and $f_{xy} = -4$, so the discriminant is

$$D_f(x, y) = 60x - 16.$$

At the critical point $(x_1, y_1) = (1/3, 2/3)$, we have

- $D_f(1/3, 2/3) = 20 - 16 = 4 > 0$ and
- $f_{xx}(1/3, 2/3) = 10 > 0$,

so $f(1/3, 2/3) = 56/27$ is a relative minimum value.

At the critical point $(x_2, y_2) = (-1/5, -2/5)$, we have

- $D_f(-1/5, -2/5) = -12 - 16 = -28 < 0$,

so $f(-1/5, -2/5) = 1.6$ is neither a minimum value nor a maximum value.

Example 4. For quadratic functions, the second derivative test is not strictly necessary,** but it can save time. For example, let

$$g(x, y, z) = 2x^2 - 4xy + 5y^2 - 2yz + z^2 - 2x - 4y + 6.$$

The first order conditions are

$$\begin{array}{lcrcl} g_x = 0 & \Longrightarrow & 4x - 4y - 2 & = & 0 \\ g_y = 0 & \Longrightarrow & -4x + 10y - 2z - 4 & = & 0 \\ g_z = 0 & \Longrightarrow & -2y + 2z & = & 0 \end{array}$$

The condition $g_z = 0$ implies that $z = y$, and substituting this into the equation $g_y = 0$ gives

$$-4x + 10y - 2y - 4 = 0 \implies 8y = 4x + 4 \implies y = \frac{x+1}{2}.$$

Substituting this into the equation $g_x = 0$ gives

$$4x - \frac{4(x+1)}{2} - 2 = 0 \implies 2x = 4 \implies x^* = 2,$$

which means that

$$z^* = y^* = \frac{x^* + 1}{2} = \frac{3}{2}.$$

The second derivatives of $g(x, y, z)$ are

$$g_{xx} = 4,\ g_{xy} = -4,\ g_{xz} = 0,\ g_{yy} = 10,\ g_{yz} = -2 \text{ and } g_{zz} = 2,$$

so the discriminants are

$$D_2(x, y, z) = 40 - 16 = 24 \quad \text{and} \quad D_3(x, y, z) = 80 + 0 - 16 - 0 - 32 = 32.$$

Since $D_3 > 0$, $D_2 > 0$ and $g_{xx} > 0$, it follows that

$$g^* = g(2, 3/2, 3/2) = \frac{13}{4}$$

is the minimum value of this function.

**Because we can 'complete the square' in the different variables.