# Calculus and optimization

These notes provide some calculus background. Throughout the course we will discuss why economists use calculus.[1]

## 1 Functions of a single variable

In an earlier set of notes we defined the concept of a function and discussed the concept of slope. For a line, calculating the slope was fairly straightforward as it is a constant and equal to the coefficient on the $x$ coordinate. With the function $y = 10 - 4x$, the slope is $-4$. For functions with curvature the slope is calculated in a similar manner (at least intuitively) but depends on the specific point along the curve. In the notes on algebra one example used a parabola with the equation $y = 3x^2 + 7x - 4$. The point $(2, 22)$ is on the graph of the parabola as is the point $(-4, 16)$. But the slopes at those points are certainly not the same – for starters, the half of the parabola on which $(-4, 16)$ lies is decreasing while the half on which $(2, 22)$ lies is increasing. What we want to do is find the line that is tangent to the curve at each point. A line is tangent to a curve if it touches the curve at a point but does not intersect the curve at that point. The reason we want to find the line that is tangent at that point is because the slope of the tangent line will be the slope at that point of the curve. I am going to "guess" that the line that is tangent at $(2, 22)$ is $y = 19x - 16$. I am also going to "guess" that the line that is tangent at $(-4, 16)$ is $y = -17x - 52$. Figure 1 shows the parabola, the two points plotted, and the two "guesses" at the tangent lines.

At a basic level, the slope tells us the change in $y$ for a specific unit change in $x$, or $\frac{\Delta y}{\Delta x}$. When finding the slope of a particular point on a curve, we want the denominator, $\Delta x$, to be as small as possible, essentially an infinitesimal (extremely small) or instantaneous change. The derivative function, $f'(x)$, tells us the slope of that line and that is one reason derivatives are important in economics. I am certainly not good enough that I could guess the equations of those tangent lines; I used the derivative function of $3x^2 + 7x - 4$ to find the slope at that point and then created the line because I had the slope and I knew a point on the line. Generally we are concerned with finding the slope of the tangent line at a particular point and not concerned with plotting the actual tangent line, but I plotted them for reference purposes.

A function $f$ is differentiable if it is continuous and "smooth" with no breaks or kinks; in economics we generally structure our models by assuming that the functions have these properties. A derivative, $f'(x)$, gives the slope or instantaneous rate of change in $f(x)$ at $x$.[2] Table 1 provides some common rules of differentiation for functions, and combinations of functions, of single variables. Note that $a$ in Table 1 is a constant. I think the use of letters to represent both constants, such as $a$, and choice variables, such as $x$, causes confusion at times but it is necessary to have that flexibility.

For the function $y = 3x^2 + 7x - 4$, the derivative is $f'(x) = 6x + 7$. The function $3x^2 + 7x - 4$ is really three "separate" functions, $3x^2$, $7x$, and $-4$, added together. We are using the Sums rule, the Power rule, and the Constants rule in determining its derivative. By the power rule the derivative of $3x^2$ is $6x$. By the power rule the derivative of $7x$ is 7. By the constants rule the derivative of $-4$ is 0. By the sums rule we can add those three derivatives together to get $6x + 7 + 0$ or $6x + 7$. With the parabola we wanted to find

[1] There is a more mathematical set of background notes available at: https://belkcollegeofbusiness.charlotte.edu/azillant/wp-content/uploads/sites/846/2014/12/ECON6202_msmicroI_math2.pdf

[2] If you have taken calculus, there is a discussion about using limits to determine the derivative of a function. We will not go through those details, but if you are extremely curious, the formal definition of the derivative function is:

$$f'(x) = \lim_{t \to x} \frac{f(t) - f(x)}{t - x}$$

While this may look a little complicated, remember that $f(\cdot)$ is really just $y$. So essentially it is just the formula for a slope, $(y_1 - y_0) / (x_1 - x_0)$, where we want the difference between $x_1$ and $x_0$ to be essentially zero.

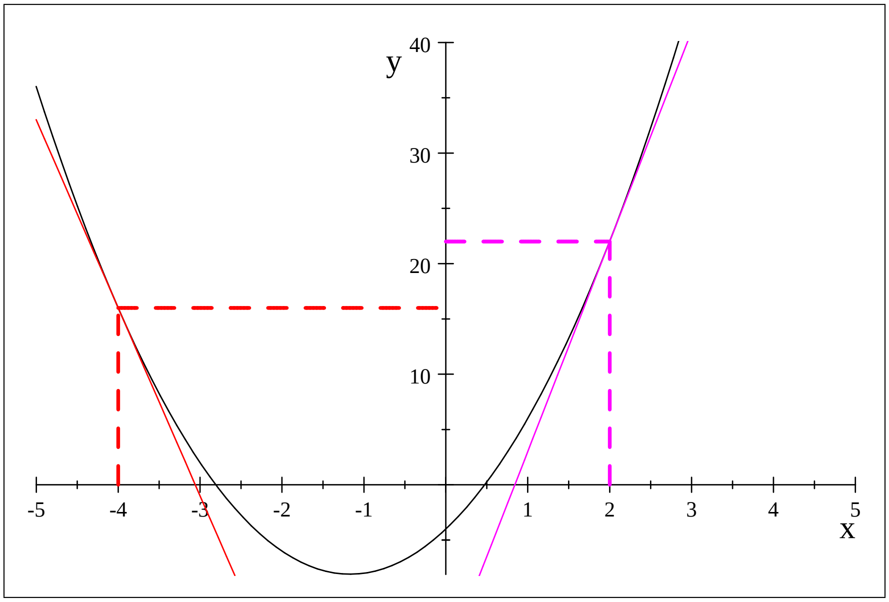


Figure 1: The plot of the parabola given by the equation $y = 3x^2 + 7x - 4$. The lines tangent to the points $(2, 22)$ and $(-4, 16)$ are given in magenta and red, respectively.

| | |
|---|---|
| Constants, $\alpha$ | $\frac{d}{dx}(\alpha) = 0$ |
| Sums | $\frac{d}{dx}(f(x) \pm g(x)) = f'(x) \pm g'(x)$ |
| Power rule | $\frac{d}{dx}(\alpha x^n) = n\alpha x^{n-1}$ |
| Product rule | $\frac{d}{dx}(f(x)\,g(x)) = f(x)\,g'(x) + f'(x)\,g(x)$ |
| Quotient rule | $\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{g(x)f'(x) - f(x)g'(x)}{(g(x))^2}$ |
| Chain rule | $\frac{d}{dx}(f(g(x))) = f'(g(x))\,g'(x)$ |
| ln rule | $\frac{d(\alpha \ln x)}{dx} = \frac{\alpha}{x}$ |

Table 1: Common rules of differentiation

the slope when $x = 2$ and when $x = -4$. Using the derivative function the slope when $x = 2$ is 19 and the slope when $x = -4$ is $-17$, which are the slopes of the respective tangent lines in Figure 1.

## 2 Functions of several variables

The previous section examined a function of a single variable, $x$. For some problems in economics, such as a firm choosing its quantity to maximize profit, a function of a single variable is sufficient. However, there are times when we want to examine functions of multiple variables. In consumer theory individuals have utility functions which are functions of the quantities consumed of various goods; in producer theory firms have cost functions which are functions of the quantities of inputs used to produce a good. When there are multiple variables, it is at times convenient to think in terms of the slope of one particular variable, holding the other variables constant. One of the key starting phrases in an economist's vocabulary is "holding all else constant" because we want to examine the effect of one variable changing.

The partial derivative of a function is the derivative of the function with respect to one variable while holding all other variables constant.[3] If there is a consumer who has the utility function $f(q_1, q_2) = q_1^2 + 3q_1q_2 + 4q_2^2$, where $q_1$ and $q_2$ represent quantities of the goods chosen, we may want to understand the effect of changing the quantity consumed of $q_1$ while holding $q_2$ constant. When taking the partial derivative with respect to $q_1$ we simply apply the Constants rule to $q_2$ and treat it like a number. Similarly, when taking the partial derivative with respect to $q_2$, we treat $q_1$ like a constant.

$$\begin{aligned} \frac{\partial f}{\partial q_1} &= 2q_1 + 3q_2 \\ \frac{\partial f}{\partial q_2} &= 3q_1 + 8q_2 \end{aligned}$$

When taking the partial derivative with respect to $q_1$, the entire term $4q_2^2$ is treated like a constant because it does not contain $q_1$ so it drops out of the derivative as if it was just a number like 9 or $-232$. As we are essentially treating our function of several variables as a single variable, the same rules of differentiation apply when taking partial derivatives. So for $\frac{\partial f}{\partial q_1}$, we can use the power rule to find that the partial derivative of the first term, $q_1^2$, is $2q_1$; the power rule to find that the partial derivative of the second term, $3q_1q_2$, is $3q_2$ (we retain the $q_2$ because this term has a $q_1$ and so we treat $q_2$ exactly like we would the number 3); and the constants rule to find that the partial derivative of the third term, $4q_2^2$, is 0. Then using the sums rule we have $2q_1 + 3q_2 + 0 = 2q_1 + 3q_2$.

## 3 Optimization

Finding derivatives is an intermediary step on the path to finding an equilibrium in our models, albeit an intermediate step that yields some important results. Ultimately we would like to use the derivatives to find a solution, which will be an optimal point. That optimal point may be a maximum (if we are maximizing utility or profit) or it may be a minimum (if we are minimizing costs). There are methods of using the second derivative of the original function to determine whether one is finding a local maximum, local minimum, or inflection point. There are also methods for determining whether a maximum or minimum is local or global. We can think of a local maximum as Mount Mitchell, which is the highest point in the state of North Carolina, but not the highest point in the U.S. or the world. The same is true for functions: sometimes they have a local maximum or minimum within a specified domain of the function, and sometimes they have a global maximum or minimum for the entire domain of the function. Economists tend to structure problems in a way such that these concerns are handled with the model assumptions to ensure we are finding the relevant optimal point.

Consider our initial function, $f(x) = 3x^2 + 7x - 4$, in Figure 1, We know that parabolas are U-shaped and we can see the (global) minimum in the picture. How do we determine the exact point at which the function is at the minimum? We still want to find the slope of the tangent line but at the specific minimum point. Intuitively, if we are looking for a tangent line at the top of a hill (or bottom of a valley), the line

[3] The standard is to use the symbol $\partial$ to denote a partial derivative where using the letter $d$ denotes the derivative.

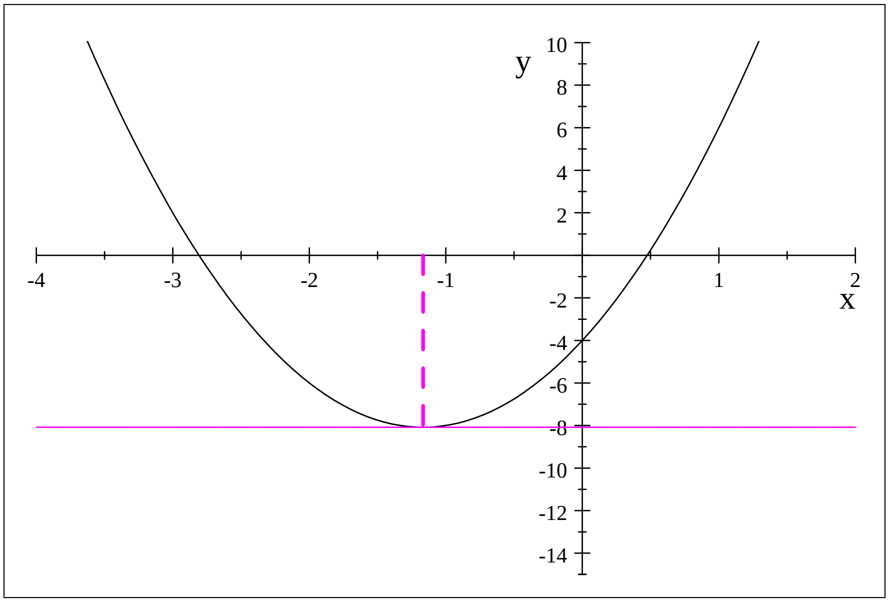


Figure 2: The plot of the parabola given by the equation $y = 3x^2 + 7x - 4$. The tangent line at the minimum is also plotted.

should be flat or perfectly horizontal. A perfectly horizontal line has a slope of 0 because there is no "rise" to the function – the $y$ values are always the same. If the derivative represents the slope of the function at any point, we want to set the slope equal to 0 and then we can solve for the $x$-coordinate of the minimum point using that equation. In this case we know the slope but want to find the specific point. We know that $f'(x) = 6x + 7$. Setting that equal to 0 we have:

$$\begin{aligned} 6x + 7 &= 0 \\ 6x &= -7 \\ x &= \frac{-7}{6} \end{aligned}$$

Using the function itself, when $x = \frac{-7}{6}$ we have:

$$\begin{aligned} f(x) &= 3x^2 + 7x - 4 \\ f\left(\frac{-7}{6}\right) &= 3\left(\frac{-7}{6}\right)^2 + 7\left(\frac{-7}{6}\right) - 4 \\ f\left(\frac{-7}{6}\right) &= 3 * \frac{49}{36} - \frac{49}{6} - 4 \\ f\left(\frac{-7}{6}\right) &= \frac{49}{12} - \frac{49}{6} - 4 \\ f\left(\frac{-7}{6}\right) &= \frac{49}{12} - \frac{98}{12} - \frac{48}{12} \\ f\left(\frac{-7}{6}\right) &= -\frac{97}{12} \end{aligned}$$

The point at which our parabola given by the equation $f(x) = 3x^2 + 7x - 4$ reaches its minimum is $\left(\frac{-7}{6}, \frac{-97}{12}\right)$. Figure 2 shows the function and the tangent line at the minimum point. Note that the tangent line is given by the equation $y = \frac{-97}{12}$ because it is a horizontal line.

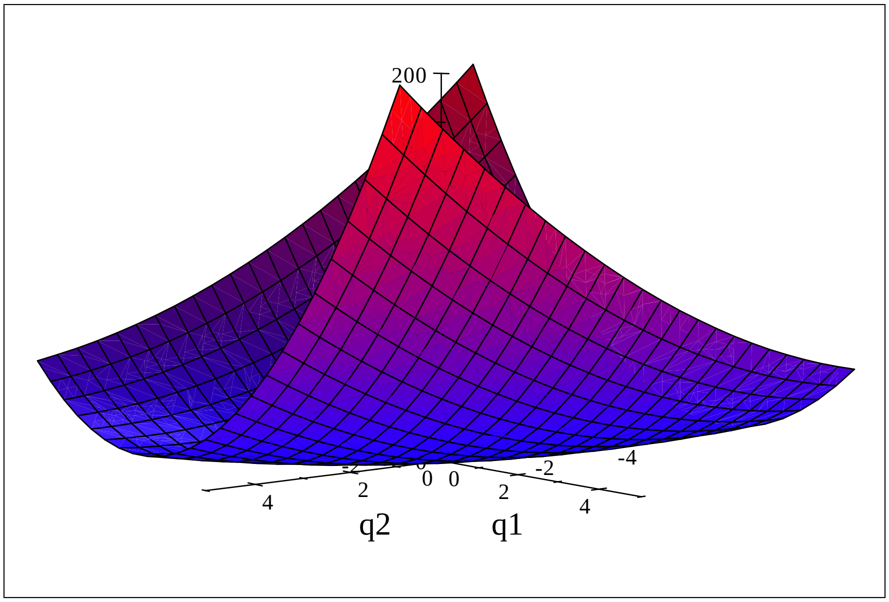


Figure 3: Plot of $q_1^2 + 3q_1q_2 + 4q_2^2$.

## 3.1 Optima for functions of multiple variables

Suppose that we have a function of $n$ variables. To find the local minimum or maximum we follow similar steps as we did with a function of a similar variable. First find the partial derivative with respect to all $n$ variables. Next set those partial derivatives equal to zero. Finally, there will be a system of $n$ equations and $n$ unknowns – solve that system of equations to find the critical values.

Figure 3 shows the plot of function $f(q_1, q_2) = q_1^2 + 3q_1q_2 + 4q_2^2$. We know that the partial derivatives are

$$\begin{aligned}\frac{\partial f}{\partial q_1} &= 2q_1 + 3q_2 \\ \frac{\partial f}{\partial q_2} &= 3q_1 + 8q_2\end{aligned}$$

Setting those equal to zero and solving we find:

$$\begin{aligned}2q_1 + 3q_2 &= 0 \\ 2q_1 &= -3q_2 \\ q_1 &= -\frac{3}{2}q_2\end{aligned}$$

Now:

$$\begin{aligned}\frac{\partial f}{\partial q_2} &= 3q_1 + 8q_2 \\ 0 &= 3q_1 + 8q_2 \\ 0 &= 3\left(-\frac{3}{2}q_2\right) + 8q_2 \\ 0 &= -\frac{9}{2}q_2 + 8q_2 \\ 0 &= q_2\left(-\frac{9}{2} + 8\right) \\ 0 &= q_2\end{aligned}$$

Now substituting into $q_1 = -\frac{3}{2} q_2$ we have that $q_1 = -\frac{3}{2} * 0 = 0$. To find the third coordinate we find $f(0,0)$ which is also zero. In this case, the minimum point is at the coordinate $(0,0,0)$, which seems consistent with the picture. Note that there are three coordinates in this ordered pair because we have a $q_1$ axis, a $q_2$ axis, and an $f(q_1, q_2)$ axis (the vertical axis).

## 3.2 Constrained optimization

Thus far we have focused on unconstrained optimization. However, in many problems there are constraints which must be met. In our study of consumer choice, we assume that consumers receive some additional utility for each unit of a good they consume, so if they were not constrained in some manner they would choose an infinite amount of every good. If we set up a model like that it would be a poor predictor of behavior.

They may be equality constraints, such that we are optimizing a particular function $f(x_1, x_2)$ subject to the equality constraint $g(x_1, x_2) = 0$. They may be constraints as simple as nonnegativity constraints, so that $x_1 \geq 0$ and $x_2 \geq 0$. They may be more complex inequality constraints, such that we are optimizing the function $f(x_1, x_2)$ subject to $g(x_1, x_2) \geq 0$. The function $g(x_1, x_2)$ may be linear or nonlinear.

### 3.2.1 Equality constraints

Formally, an optimization with an equality constraint is set up as:

$$\max_{x_1, x_2} \; f(x_1, x_2) \text{ s.t. } g(x_1, x_2) = 0$$

The function which we are optimizing, $f$ in this instance, is called the objective function. The variables that are being chosen, $x_1$ and $x_2$ in this problem, are the choice variables. The function $g(x_1, x_2)$ is called the constraint function.

With equality constrained optimization problems the problem can be solved by substitution. If we solve for one of the variables in the constraint function, say $x_2$, we would have a new function:

$$x_2 = \widetilde{g}(x_1)$$

Substitute this directly into the objective function and the problem becomes:

$$\max_{x_1} \; f(x_1, \widetilde{g}(x_1))$$

and then we are maximizing a function of a single variable. We then set the first derivative equal to zero and find the critical value for $x_1$. This first order condition is:

$$\frac{\partial f(x_1^*, \widetilde{g}(x_1^*))}{\partial x_1} + \frac{\partial f(x_1^*, \widetilde{g}(x_1^*))}{\partial x_2} \frac{d\widetilde{g}(x_1^*)}{dx_1} = 0$$

Once $x_1^*$ is known we find $x_2^*$ by using $x_2^* = \widetilde{g}(x_1^*)$. For simple problems this substitution method works well; when the constraint functions are complex or there are multiple choice variables (or constraints) this method becomes tedious. While a straightforward method of finding a solution, this method of direct substitution does not lend itself to building intuition about the problem being solved.

### 3.2.2 Lagrange's method

Solving unconstrained optimization problems is (relatively) easy. The idea that Lagrange had was to turn constrained optimization problems into unconstrained optimization problems. Our problem from before is:

$$\max_{x_1, x_2} \; f(x_1, x_2) \text{ s.t. } g(x_1, x_2) = 0$$

Now, multiply the constraint by the variable $\lambda$ (why? that will be explained later in the inequality constraints section). Add that product to the objective function and we have created a new function called the Lagrangian function (or Lagrangian). The Lagrangian is:

$$\mathcal{L}(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2)$$

Note that there are now three choice variables: $x_1$, $x_2$, and $\lambda$. The first-order necessary conditions for optimizing the Lagrangian are the set of partial derivatives set equal to zero:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x_1} &= \frac{\partial f\left(x_1^*, x_2^*\right)}{\partial x_1} + \lambda^* \frac{\partial g\left(x_1^*, x_2^*\right)}{\partial x_1} = 0 \\
\frac{\partial \mathcal{L}}{\partial x_2} &= \frac{\partial f\left(x_1^*, x_2^*\right)}{\partial x_2} + \lambda^* \frac{\partial g\left(x_1^*, x_2^*\right)}{\partial x_2} = 0 \\
\frac{\partial \mathcal{L}}{\partial \lambda} &= g\left(x_1^*, x_2^*\right) = 0
\end{aligned}$$

The idea of Lagrange's method is that if we solve these three equations simultaneously for $x_1^*$, $x_2^*$, and $\lambda^*$ we will have found a critical point of $f\left(x_1, x_2\right)$ along the constraint $g\left(x_1, x_2\right) = 0$. Note that while the function $f\left(x_1, x_2\right)$ and the Lagrangian $\mathcal{L}\left(x_1, x_2, \lambda\right)$ are different (the latter has the constraint tacked on to it), because $g\left(x_1, x_2\right) = 0$, we know that $\lambda g\left(x_1, x_2\right)$ so we are essentially adding zero to the original function. Intuitively, we are maximizing $f\left(x_1, x_2\right) + 0$ which is like maximizing $f\left(x_1, x_2\right)$. We will not go through all of the details to prove why this method works; there are plenty of math books that provide a nice discussion. Also, as of now we do not know whether or not these critical values are maxima or minima, but some comments on this topic will be made shortly.

Note that Lagrange's method can be used for any number of variables ($n$) and any number of constraints ($m$) so long as the number of constraints is less than the number of variables ($m < n$). There is still the question of whether or not a solution actually exists for a problem and whether or not the $\lambda$ variable exists. While these are important questions a general discussion of these concepts is more than is needed for our purposes.

**Inequality constraints** In some problems we have inequality constraints with which we must contend. There are many technical details that we will not cover because they are not that important for our purposes. An inequality constrained problem is very similar to an equality constrained problem it is just that the constraint may or may not hold with equality.

$$\max_{x_1, x_2} f\left(x_1, x_2\right) \text{ s.t. } g\left(x_1, x_2\right) \geq 0$$

This problem is a nonlinear programming problem, as a linear programming problem we have a linear function which is optimized subject to linear equality constraints. We make no such restrictions here. We can construct the Lagrangian to find:

$$\mathcal{L}\left(x_1, x_2, \lambda\right) = f\left(x_1, x_2\right) + \lambda\left[g\left(x_1, x_2\right)\right]$$

The necessary conditions to find a solution are similar to what we have seen earlier:

$$\begin{aligned}
\frac{\partial f}{\partial x_1} + \lambda \frac{\partial g}{\partial x_1} &= 0 \\
\frac{\partial f}{\partial x_2} + \lambda \frac{\partial g}{\partial x_2} &= 0 \\
\lambda g\left(x_1, x_2\right) &= 0 \\
\lambda &\geq 0, \; g\left(x_1, x_2\right) \geq 0
\end{aligned}$$

The first two equations are just the partial derivatives with respect to the choice variables $x_1$ and $x_2$. The third equation ensures that we are not really "changing" the original objective function. That last row with the two inequalities are technical conditions that need to be satisfied. Focusing on the third equation, either $\lambda = 0$ and $g\left(x_1, x_2\right) \geq 0$ or $\lambda > 0$ and $g\left(x_1, x_2\right) = 0$ or (in rare cases) both equal zero. If $\lambda = 0$ and $g\left(x_1, x_2\right) > 0$, then that result would mean that the constraint does not hold with equality; in other words, the constraint does not matter (typically we would say that the constraint does not bind) because the objective function is maximized at some feasible point that is not along the constraint. If $\lambda > 0$ and $g\left(x_1, x_2\right) = 0$, then the constraint is binding (it is being set to equality). The purpose of $\lambda$ is to show the marginal value of relaxing the constraint. If $\lambda = 0$, then there is no value in relaxing the constraint because the constraint does not bind. But if $\lambda > 0$ then the constraint does not bind and the solution would change. Putting this concept into an economic model, when consumers are maximizing utility we assume that they have limited income (a budget constraint). If $\lambda = 0$, then it would be the case that the budget constraint did not bind so it did not affect their choices.[4]

Economists typically use this type of solution concept to solve consumer optimization problems; when the only constraint is the budget constraint, economists assume it binds (people do not throw away money that they could use to buy some good) and so equality constrained optimization methods could be used. However, suppose there is a requirement that a certain amount of some good, such as some minimum level of insurance, be consumed. For some consumers that constraint may be binding (they would have chosen less than the required amount of insurance if they could) but for other consumers that constraint may not be binding (they were going to choose more than the minimum required amount of insurance anyway). In those types of problems whether or not the constraint binds is important.

[4] When $g(x_1, x_2) = 0$ and $\lambda = 0$ we have the odd case that the optimal solution just happens to fall on the constraint. So while the solution falls on the constraint, technically the constraint is not binding.