Equality constraints. Lagrange multipliers

1. (a) Lagrangian $L(x, y, \lambda) = x^2 + y^2 - xy + x + y - 4 - \lambda(x + y + 3)$, a convex function of (x, y) for any λ, so a local minimum is the absolute one.

   Critical points: $2x - y + 1 - \lambda = 0$, $2y - x + 1 - \lambda = 0$, $x + y = -3$. Subtracting the second equation from the first one yields $x = y$, so the solution $x = y = -3/2, \lambda = -1/2$ is easy to find. Thus it is the minimum sought, where the objective $z = -4.75$.

   On the other hand, you could express y from the constraint, plug it into the objective function and deal with the case of one variable only.

   (b) Lagrangian $L(x, y, \lambda) = 1/x + 1/y - \lambda(x + y - 2)$, a convex function of nonzero (x, y) given λ.

   Critical points: $-1/x^2 - \lambda = 0$, $-1/y^2 - \lambda = 0$, $x + y = 2$, subtracting the first two yields $x = \pm y$, then by the third one $x = y$, and thus $x = y = 1, \lambda = -1$ by the first one. It is a minimum, by convexity, where the objective $z = 2$. Note that as x or y go to zero along the line $x + y = 2$, the objective goes to infinity.

   On the other hand, you could express y from the constraint, plug it into the objective function and deal with the case of one variable only.

   (c) The objective function is the Euclidean distance, sought on the ellipsoid, so the minimum and the maximum will correspond to the smallest and the largest semi-axes. Maxima: $x = \pm 4$, $y = z = 0$, where the objective equals 16, minima $z = \pm 2$, $x = y = 0$, where the objective equals 4. Saddle points $y = \pm 3$, $x = z = 0$.

   Of course, it can all be done with the Lagrange multipliers. You get

$$L(x, y, z, \lambda) = x^2(1 - \lambda/16) + y^2(1 - \lambda/9) + z^2(1 - \lambda/4) + \lambda.$$

   First, restrict yourself to $x, y, z \geq 0$, by symmetry: all the functions involved are quadratic, thus even.

   Lagrange equations:

$$x(-\lambda + 16) = 0, \quad y(-\lambda + 9) = 0, \quad z(-\lambda + 4) = 0; \text{ plus } x^2/16 + y^2/9 + z^2/4 = 1.$$

   Solutions (note that $(x = y = z = 0)$ is unfeasible):

$$\lambda = 16, y = z = 0; \text{ or } \lambda = 9, x = z = 0; \text{ or } \lambda = 4, y = x = 0.$$

   From the constraint, the first one yields $x = \pm 4$, the second $y = \pm 3$, the third one $z = \pm 2$.

   To classify the critical points found as local minima/maxima, or saddles, observe that whenever a constraint is satisfied the Lagrangian equals the objective function. To each critical point found, there corresponds its own value of λ, and therefore in order to use he second derivative test for the given critical point, one should fix the λ corresponding to it in the Lagrangian, and then look at the second derivative in the variables (x, y, z).

   Substitution of, say the value of $\lambda = 16$ into L eliminates the variable x from it, leaving $-7y^2/9 - 3z^2 + 16$, this is clearly a concave function, with a maximum at $(y, z) = (0, 0)$.

   The other values of λ get treated in the same way, and you conclude that $\lambda = 4$ yields a maximum and $\lambda = 9$ yields a saddle.

2. Suppose, x is the width of the tub, y is the length, and z is *twice* the height (to symmetrise the problem). Then the optimisation problem is to minimize $xz + yz + xy$, given $xyz = 2V$.

   Now comes the easiest solution: By the Young inequality $xz + yz + xy \geq 3\sqrt[3]{xzyzxy} = 3\sqrt[3]{(2V)^2}$. The left-hand side is minimum when the equality occurs, and we know that this is only possible when $xy = yz = xz$, so $x = y = z = \sqrt[3]{2V}$.

   With the Lagrange multipliers, let us proceed as follows.

The Lagrangian is $L(x,y,z,\lambda) = xz + yz + xy - \lambda(xyz - 2V)$, the critical point equation is $y + z - \lambda yz = 0$, the same for other pairs of variables. Thus $\lambda > 0$ and

$$\frac{y+z}{yz} = \frac{x+z}{xz} = \frac{y+x}{xy}, \text{ i.e. } \frac{1}{y} + \frac{1}{z} = \frac{1}{z} + \frac{1}{x} = \frac{1}{x} + \frac{1}{y}, \text{ so } x = y = z = \sqrt[3]{2V}.$$

It's will be obvious that this is a minimum if we show that whenever one of the coordinates goes to infinity (given the volume V), then the surface are also goes to infinity. In other words, the feasible set can be effectively treated as closed and bounded, so that the Extreme value theorem applies. Then the critical point we have found cannot be anything but the minimum we're after.

Here is one way to see it. First off, the surface area is bounded from below by zero, and the domain is $x, y, z > 0$. As one of the variables, say z goes to zero, to compensate for the nonzero volume V, one is forced to have a huge area for the bottom of the tub. So the area is infinite if either x, y or z goes to zero. On the other hand, suppose z goes to infinity. Then the bottom of the tub has a very small area V/z. But how about the area of the sides? It is z times the perimeter of the base, and the latter will be about $\sqrt{V/z}$ so the side area will be about $z\sqrt{V/z}$ and will go to infinity as z goes to infinity. So, the total surface area we are looking at is bounded from below, and goes to infinity as either of the measurements x, y, z goes to zero or infinity. Then it must have a minimum somewhere inside the domain $x, y, z = 0$. This minimum shall be given by some critical point of the Lagrangian, and we have one critical point only. Thus, the critical point we have found IS this minimizer.

Here is how to do it using the second derivative test, by the way. You can eliminate the constraint by expressing $z = \frac{2V}{xy}$, and plug it into the objective function, which will become $2V\left(\frac{1}{x} + \frac{1}{y}\right) + xy$. The Hessian of this function of two variables is equal to $\begin{pmatrix} \frac{4V}{x^3} & 1 \\ 1 & \frac{4V}{y^3} \end{pmatrix}$. With $x = y = \sqrt[3]{2V}$ it is positive definite, so we have found the above critical point is a local minimum of the objective function, where the value is $3\sqrt[3]{4V^2}$. As there are no other critical points and the surface area is bounded form below by zero, it is possible to conclude that this is an absolute minimum.

In terms of the original problem it corresponds to the tub with the square base, and the height equal to one half the width.

Kuhn-Tucker conditions

Sketches are given on the last page.

1. (a) The feasible set F is Sketch 2 on the Figure below, the maximizer is the point $\boldsymbol{x} = (1,0)$, where the objective $x_1 = 1$.

   (b) At $\boldsymbol{x} = (1,0)$ the constraints $x_2 \geq 0$, $(1-x_1)^5 - x_2 \geq 0$ are tight, the functions in the left-hand side have nonzero gradients at $\boldsymbol{x}$, directed vertically upwards and downwards, respectively. Hence the set $\texttt{FD}(\boldsymbol{x})$ of *feasible directions* at $(\boldsymbol{x})$ (direction vectors $\boldsymbol{v}$ forming angles of no more than 90 degrees with all the gradient involved); consists of two vectors $(\pm 1, 0)$. The feasible direction $(1,0)$ in not a *true feasible* one: there is no path contained in the feasible set, to which $(1,0)$ would be tangent at $\boldsymbol{x}$. Thus at the point $(1,0)$ the CQ non-degeneracy assumption is not satisfied.

   (c) Let us denote the Lagrange multipliers, corresponding to $\boldsymbol{x} \geq 0$ as $\boldsymbol{\mu}$. The Lagrangian is $L(\boldsymbol{x}, \lambda, \boldsymbol{\mu}) = x_1 + \lambda[(1-x_1)^5 - x_2] + \mu_1 x_1 + \mu_2 x_2$. Differentiate the Lagrangian w.r.t. x_1. Then $1 - 5\lambda(1-x_1)^4 + \mu_1 = 0$ for some $\lambda, \boldsymbol{\mu} \geq 0$.

   Let us show that KT conditions above have no solutions.

$$\begin{array}{rcl} 1 - 5\lambda(1-x_1)^4 + \mu_1 &=& 0, \\ \mu_2 &=& \lambda, \\ \lambda[(1-x_1)^5 - x_2 = \mu_1 x_1 = \mu_2 x_2 = 0. && \end{array} \tag{1}$$

   We cannot have $\lambda = 0$, as the first equation then would yield $\mu_1 = -1$, unfeasible. If $\lambda > 0$, then $\mu_2 > 0$ and the corresponding two constraints must be tight: $(1-x_1)^5 - x_2 = x_2 = 0$. Which means again, $\mu_1 = -1$, unfeasible. Hence, due to the fact that CQ are not satisfied at $(1,0)$, KT equations for the *max* x_1 problem have no solution.

   (d) Adding the constraint $x_1 \leq 1$ adds the additional term $+\nu(1-x_1)$ to the Lagrangian, with the new Lagrange multiplier $\nu \geq 0$. So (1) now has ν in the right-hand side of the first equation, in addition $\nu(1-x_1) = 0$. We know that there are no solutions for $\nu = 0$, so must have $x_1 = 1$, and then, from the first equation, $\nu = 1$. Which means the new constraint must be tight: $x_1 = 1$, then $\mu_1 = 0$ and $x_2 = 0$ (from the constraints themselves). The Lagrange multipliers $\lambda = \mu_2$ can be anything in $\mathbb{R}_+$.

   CQ are now satisfied, because the vector $(1,0)$ is no longer a feasible direction at $\boldsymbol{x}$: the arrival of the new constraint requires $\boldsymbol{v} \cdot (-1, 0) \geq 0$ for a feasible direction $\boldsymbol{v}$, which $\boldsymbol{v} = (1,0)$ fails to satisfy.

   (e) Changing the objective to Max x_2 in the initial problem gives KT conditions

$$\begin{array}{rcl} -5\lambda(1-x_1)^4 + \mu_1 &=& 0, \\ 1 + \mu_2 &=& \lambda, \\ \lambda[(1-x_1)^5 - x_2 = \mu_1 x_1 + \mu_2 x_2 = 0. && \end{array}$$

   The second equation necessitates $\lambda > 0$ (because $\boldsymbol{\mu} \geq 0$). So $(1-x_1)^5 = x_2$. From the first equation, if $\mu_1 = 0$, we must have $x_1 = 1$, so $x_2 = 0$, $\mu_2 \geq 0$ and $\lambda = 1 + \mu_2$. This is a one-parameter family of solutions of KT conditions, which does not give us a maximizer, however. Instead, it gives us the point $(1,0)$, where as we know CQ fails.

   Nothing unusual about it: The KT theorem is only a necessary condition: it says, the maximizer must be *a solution* of Kuhn-Tucker/Lagrange conditions (provided that the non-degeneracy assumption CQ is satisfied).

   Consider then $\mu_1 > 0$, then we must have $x_1 = 0$, so as the constraint is tight (and we still must have $\lambda > 0$) $x_2 = 1$. We then get another, in fact two-parameter family of solutions $x_1 = 0, x_2 = 1, \mu_1 > 0, \mu_2 \geq 0, \lambda = 1 + \mu_2$. This family gives us the true maximizer $\boldsymbol{x} = (0,1)$.

2. (a) The feasible set F is Sketch 3 on the Figure below, the minimiser is the origin $\boldsymbol{x} = (0,0)$, where the objective is zero.

   (b) As there is only one constraint, for $\lambda \geq 0$, let $L(x, y, \lambda) = x - \lambda(x^3 - y^2)$. KTL conditions:

$$\begin{array}{l} 0 = L_x = 1 - 3\lambda x^2, \; 0 = L_y = 2\lambda y, \\ \lambda(x^3 - y^2) = 0, \; \lambda \geq 0, \; x^3 - y^2 \geq 0. \end{array}$$

If $\lambda = 0$, the first equation is absurd. Then $\lambda > 0$, whence $y = 0$ from the second equation, so $x = 0$ by the third one. As a result, the first equation is absurd again. Contradiction!

This tells us that the mimimiser (which must exist, since the constraint requires $x \geq 0$ and the objective does not decrease as x goes to infinity, but increases, so we can treat the feasible set as if it were closed and bounded.) may occur *only* where the non-degeneracy condition CQ fails. At this point, since there is only one constraint, we must have its gradient vanish. The gradient of $g(x, y) = x^3 - y^2$ is zero only at the origin. So, the conclusion is the origin *must be* the mimimiser. CQ is not satisfied at the origin, because the gradient of the constraint at $(0, 0)$ is the zero vector. To this end, *any* direction is feasible at $(0, 0)$ (indeed, $\boldsymbol{v} \cdot (0, 0) \geq 0$ for any $\boldsymbol{v}$), while only $\boldsymbol{v} = (1, 0)$ is a *true feasible*direction.

(c) Now, with the extra constraint $x \geq 0$, we have $L(x, y, \lambda) = x - \lambda(x^3 - y^2) - \mu x$, with $\mu \geq 0$. CQ are still not satisfied: any $\boldsymbol{v} = (\alpha, \beta)$, with $\alpha \geq 0$ any β - the only requirement for $\boldsymbol{v}$ to be in $FD(0, 0)$ is $\boldsymbol{v} \cdot (1, 0) \geq 0$.

As for the formal KT equations: The first line of the KT above changes to

$$0 = L_x = 1 - 3\lambda x^2 - \mu, \ 0 = L_y = 2\lambda y.$$

If $\lambda = 0$, get $\mu = 1, x = 0, y = 0$.

If $\lambda > 0$, again $\mu = 1, x = 0, y = 0$. Over all, $\lambda \geq 0$ becomes a parameter. So the solutions of KT are given by $(x, y, \lambda, \mu) = (0, 0, \lambda, 1)$, for any $\lambda \geq 0$. The opitmizer $(x, y) = (0, 0)$ is thereby unambiguously fixed, but λ is undetermined.

So, failure of CQ does not *necessarily* fail KT completely, but it very well may.

Note however that if the constraints are rewritten in a slightly different, yet geometrically equivalent form $y \leq |x|^{3/2}$, $y \geq -|x|^{3/2}$, $x \geq 0$, then CQ will be satisfied in problem P', but not P.

3. The sketch is essentially the same as Sketch 3 below. The constraints are $x^5 - y \geq 0$ and $x^5 + y \geq 0$. Their gradients at the origin are $(0, -1)$ and $(0, 1)$, pointing in the opposite vertical directions. The vector $\boldsymbol{v} = (1, 0)$ gives the only true feasible direction. The feasible directions without the constraint $x \geq 0$ are both $\boldsymbol{v}$ and $-\boldsymbol{v}$, so CQ is not satisfied. It is satisfied, however with the additional constraint $x \geq 0$, because this requires $\boldsymbol{v} \cdot (1, 0) \geq 0$, ruling the vector $-\boldsymbol{v}$ out.

   Without the constraint $x \geq 0$ the Lagrangian is

$$L = x - \lambda_1(x^5 - y) - \lambda_2(x^5 + y), \quad \lambda_{1,2} \geq 0$$

   and the KT conditions

$$1 - 5x^4\lambda_1 - 5x^4\lambda_2 = 0, \ \ \lambda_1 = \lambda_2, \ \ \lambda_1(x^5 - y) + \lambda_2(x^5 + y) = 0$$

   are absurd, because the last two require that either x or $\lambda_1 = \lambda_2$ be zero, and in both cases this means $1 = 0$ in the first equation.

   With the constraint $x \geq 0$ the Lagrangian is

$$L = x - \lambda_1(x^5 - y) - \lambda_2(x^5 + y) - \mu x, \quad \lambda_{1,2}, \, \mu \geq 0$$

   and the KT conditions are

$$1 - 5x^4\lambda_1 - 5x^4\lambda_2 = \mu, \ \ \lambda_1 = \lambda_2, \ \ \lambda_1(x^5 - y) + \lambda_2(x^5 + y) = 0, \ \mu x = 0.$$

   This is fine now, because as it has been shown one must have $\mu > 0$ (otherwise we are in the situation above), so $x = 0$, then $y = 0$ because this is the only feasible y corresponding to $x = 0$, and $\lambda_1 = \lambda_2$ can be any non-negative number.

4. The Lagrangian is $L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \frac{1}{2}\boldsymbol{x} \cdot C\boldsymbol{x} + \boldsymbol{c} \cdot \boldsymbol{x} + \boldsymbol{\lambda} \cdot (A\boldsymbol{x} - \boldsymbol{b}) + [-]\boldsymbol{\mu} \cdot \boldsymbol{x}$ with $\boldsymbol{\lambda} \in \mathbb{R}^m$ (equality constraints $A\boldsymbol{x} = \boldsymbol{b}$ impose no sign constraint on $\boldsymbol{\lambda}$) and $\boldsymbol{\mu} \geq 0$ (square brackets refer to the `min` problem).

   One can assume that C is symmetric, since $\boldsymbol{x} \cdot C\boldsymbol{x} = 0$ for an antisymmetric C. (In other words $\boldsymbol{x} \cdot C\boldsymbol{x}$ is a quadratic form in $\boldsymbol{x}$.)

(a) Rewrite KT as inequalities, using $\boldsymbol{\mu}, \boldsymbol{x} \geq 0$:

$$C\boldsymbol{x} + \boldsymbol{c} + A^T\boldsymbol{\lambda} = \pm\boldsymbol{\mu}, \quad \boldsymbol{x} \cdot C\boldsymbol{x} + \boldsymbol{c} \cdot \boldsymbol{x} + \boldsymbol{\lambda} \cdot A\boldsymbol{x} = 0, \quad A\boldsymbol{x} = \boldsymbol{b}, \quad \boldsymbol{x} \geq 0.$$

The second equation is, in fact $\boldsymbol{\mu} \cdot \boldsymbol{x} = 0$. These equations are central for *quadratic* programming.

(b) KT is sufficient for the maximizer [minimizer] that $L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ be a concave [convex] function of $\boldsymbol{x}$ alone, i.e whenever the matrix C is negative [positive] definite.

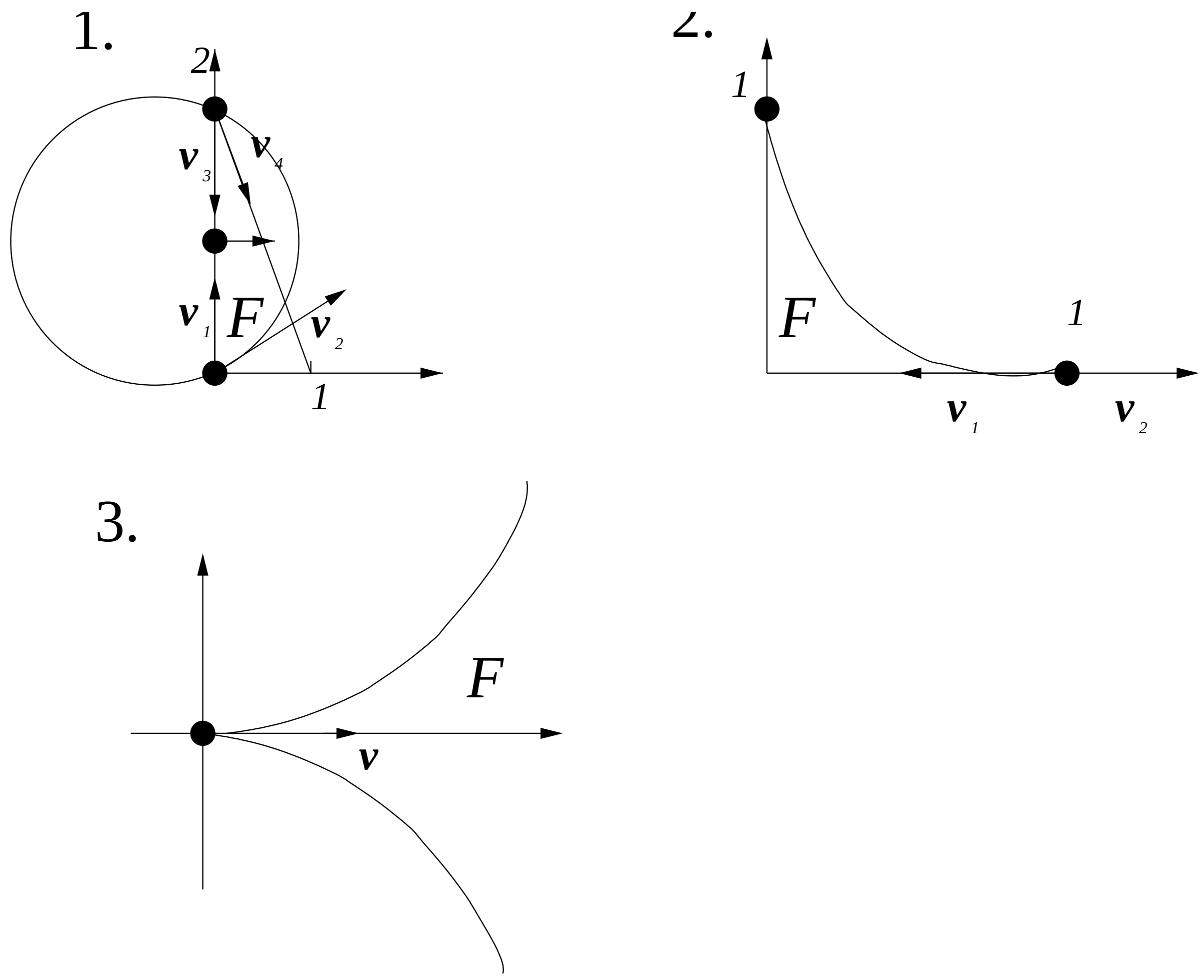

Figure 1: Sketches for Kuhn-Tucker Problems 1,2. Sketch 1 is old stuff, disregard it. Sketches 2,3 correspond to Problems 1,2.