Some Review
10/21/04

Exercise 1 *Show that $S(A)$ and $N(A')$ are orthogonal subspaces, in the sense that $z \in S(A), u \in N(A') \Rightarrow z'u = 0$. Show further that $S(A) + N(A') = \mathbb{R}^m$, in the sense that for every $y \in \mathbb{R}^m$ there are vectors $z \in S(A)$ and $u \in N(A')$ such that $y = z + u$.*

Solution: For the first part, we just use the definitions of $S(A)$ and $N(A')$. $z \in S(A)$ means that $z = Ax$ for some $x \in \mathbb{R}^n$, and $u \in N(A')$ means that $A'u = 0$. So we have

$$z'u = x'A'u = 0.$$

For the second part, I will first simply show that such z and u exist - then show how we get them. It turns out that if we choose any $x \in S(A)$, then $z = x\frac{(x'y)}{(x'x)}$ and $u = y - z$ will satisfy $y = z + u$. It is clear that $z \in S(A)$; $\frac{(x'y)}{(x'x)}$ is simply a scalar. It is also clear that $y = z + u$, since $z + u = z + y - z = y$. The only thing to check is that $z'u = 0$, since we know that $u \in N(A') \Leftrightarrow u \perp S(A)$. Checking this, we have

$$\begin{aligned} z'u &= (x\frac{(x'y)}{(x'x)})'(y - x\frac{(x'y)}{(x'x)}) \\ &= x'y(\frac{(x'y)}{(x'x)}) - x'x(\frac{(x'y)}{(x'x)})^2 \\ &= x'y(\frac{(x'y)}{(x'x)}) - x'y(\frac{(x'y)}{(x'x)}) = 0 \end{aligned}$$

Now - how would we have found this z? You might have some idea if you notice that $x\frac{(x'y)}{(x'x)}$ is commonly called the **orthogonal projection of y on x.** The name comes from the picture that goes with it - suppose for intution that y and x are vectors in $\mathbb{R}^2$, and that they're linearly independent. Now extend x, drawing in its full span (which is a line). If we drop a line from y to the span of x such that the line is perpendicular to x, the vector which ends at the intersection of this line and the span of x is the orthogonal projection of y on x. Notice that precisely because the line was perpendicular to the span of x, we have found the vector in the span of x whose head is closest to the head of y. This is how ordinary least squares works in econometrics. You have a LHS variable y, and you're trying to describe it as a function of RHS variables X. In this example, there's just one RHS variable, x. Your best prediction of y is going to be the vector in x that is as close as possible to y. The difference is simply $y - x$, which is our u here, commonly called the residual. You want to choose a vector in the span of x so as to minimize the norm of the residual, which is a sum of squares - hence the name ordinary least squares. Going through this minimization reveals where the formula for the orthogonal projection comes from:

$$\begin{aligned}
&\min_{\beta} u'u \text{ s.t. } u = y - \beta x \\
&\min_{\beta} (y - \beta x)'(y - \beta x) \\
&\min_{\beta} y'y - 2\beta x'y + \beta^2 x'x \quad \text{(note that } x'y = y'x\text{)} \\
-2x\,'y + 2\beta x'x &= 0 \\
\beta &= \frac{x'y}{x'x}
\end{aligned}$$

So $z = x\frac{(x'y)}{(x'x)}$ is precisely the vector in the span of x whose head is closest to the head of y, which means that the segment connecting the two heads, which is the residual, must be perpendicular to it. Moreover, since this segment is simply $u = y - x\frac{(x'y)}{(x'x)}$, we know that $u + z = y$.

Exercise 2 ***(Optimization in*** $\mathbb{R}^n$***)*** *Let* $F(x, y) = x^2 + y^2 - 4x - 2y$ *and* $G(x, y) = x^2 + 2y^2 - 4x - 4y$.

1. *State the implicit function theorem. Find all points on the curve* $G(x, y) = 0$ *around which either* y *is not expressible as a function of* x *or* x *is not expressible as a function of* y. *Compute* $y'(x)$ *along the curve at the origin*

   Solution: *Let* $F(x_1, ..., x_n)$ *be a* C^1 *function around the point* $(x_1^*, ..., x_n^*, y^*)$ *such that* $\frac{\partial F}{\partial y}(x_1^*, ..., x_n^*, y^*) \neq 0$. *Denote* $c = F(x_1^*, ..., x_n^*, y^*)$. *Then there exists a* C^1 *function* $y = y(x_1, ..., x_n)$ *defined around* $(x_1^*, ..., x_n^*)$ *such that:*

   - $F((x_1, ..., x_n, y(x_1, ..., x_n)) = c$
   - $y^* = y(x_1^*, ..., x_n^*)$
   - $\frac{\partial y}{\partial x_i}(x_1^*, ..., x_n^*) = -\frac{\frac{\partial F}{\partial x_i}(x_1^*, ..., x_n^*, y^*)}{\frac{\partial F}{\partial y}(x_1^*, ..., x_n^*, y^*)}$.

   $\frac{\partial G}{\partial y}(x, y) = 4y - 4$, *so* y *is not expressible as a function of* x *when* $y = 1$. $\frac{\partial G}{\partial x}(x, y) = 2x - 4$, *so* x *is not expressible as a function of* y *when* $x = 2$.

$y'(x)$ *at the origin is* $-\frac{\frac{\partial G}{\partial x_i}(0,0)}{\frac{\partial G}{\partial y}(0,0)} = -\frac{-4}{-4} = -1$

*This is probably a good place to stop and think about what G looks like (especially if you've noticed the role it plays in parts c-e; always good to read the whole question before starting to work on it!). We've just established that $G = 0$ defines y as a function of x, except when $y = 1$ - so we have a curve defined in the x, y plane. What does it look like? I claim that it is an ellipse. To see this, note that we can split $G(x, y)$ into $g(x) = x^2 - 4x$ and $h(y) = 2y^2 - 4y$. Now let's 'complete the square for each of these. $g(x) = x^2 - 4x = x^2 - 4x + 4 - 4 = (x-2)^2 - 4$, and $h(y) = 2y^2 - 4y = 2y^2 - 4y + 2 - 2 = 2(y-1)^2 - 2$. So $G(x, y) = g(x) + h(y) = (x-2)^2 + 2(y-1)^2 - 6$. This is clearly the equation for a function in $\mathbb{R}^3$ which is symmetric about the lines $x = 2$, $y = 1$, and which is then simply shifted down by a constant (6, in this case)). And its level sets are clearly ellipses centered at $(2, 1)$. It takes some time to gather this information, but it will make the rest of the problem **much** easier to solve!*

2. *Find all unconstrained optima of F and G on $\mathbb{R}^2$. Is the Weierstraß theorem applicable?*

   Solution: *In the unconstrained problem, Weierstraß is not applicable, because the domain of the functions ($\mathbb{R}^2$) is not bounded.*

   Taking FOCs for F gives

$$2x - 4 = 0 \tag{1}$$
$$2y - 2 = 0 \tag{2}$$

   so $(2, 1)$ is the only critical point. Notice that we can split F into $g(x) = x^2 - 4x$ and $h(y) = y^2 - 2y$, so we can deal with second order conditions by examining each of these separately. Since both of these are convex ($g''(x) = h''(y) = 2$), F is convex, which means that $(2, 1)$ is a global minimum; there are no maxima.

   We can do exactly the same thing for G. The FOCs give

$$2x - 4 = 0 \tag{3}$$
$$4y - 4 = 0 \tag{4}$$

   which means that $(2, 1)$ is also the unique critical point of G. Again, G is clearly convex, so it is a global minimum.

3. *Maximize and minimize $F(x, y)$ subject to $G(x, y) = 0$. Is the Weierstraß theorem applicable?*

Solution: *To answer the part about Weierstraß, recall that the constraint set $G(x, y) = 0$ is an ellipse with its center at $(2, 1)$. Since such an ellipse is both closed and bounded, and F is continuous, Weierstraß is applicable, so we should be able to find both a minimum and a maximum on this set. To do so, we write our Lagrangian:*

$$\Lambda(x, y, \lambda) = x^2 + y^2 - 4x - 2y - \lambda(x^2 + 2y^2 - 4x - 4y) \tag{5}$$

And then take FOCs:

$$\frac{\partial \Lambda}{\partial x} = 2x - 4 - \lambda(2x - 4) = 0 \tag{6}$$

$$\frac{\partial \Lambda}{\partial y} = 2y - 2 - \lambda(2y - 4) = 0 \tag{7}$$

$$\frac{\partial \Lambda}{\partial \lambda} = x^2 + 2y^2 - 4x - 4y = 0 \tag{8}$$

If $x = 2$, (6) is satisfied, and (8) tells us that $y^2 - 2y - 2 = 0$, or $y = 1 \pm \sqrt{3}$. But if $x \neq 2$, then we have

$$\lambda = \frac{2x - 4}{2x - 4} = 1 = \frac{y - 1}{y - 2} \tag{9}$$

which isn't going to hold. So have we found all the possible solutions? This is where David Kreps would write, "Take out a pencil and put an x on the page to mark this spot." Sounds corny, but I'm not going to argue with David Kreps - take out a pencil and put an x on the page to mark this spot.

No, we haven't found all the possible solutions. We've found the only critical points (where the slope of the Lagrangian is zero), but we need to remember boundary points. What did we just do? We found x satisfying (6), and plugged it into the constraint, expressed in (8), so that the constraint was a function of y, and solved for y. But the constraint's not just a function of y; it's a function of x and y, unless we think of the level set described by the constraint as a function $y(x)$. That's essentially what we did; we plugged in a particular x and found the y implicitly defined by it. You might wonder how there can be boundary points on an ellipse, but we have to remember that $y(x)$ isn't defined on the whole ellipse. Back in part (1), we found that it's not defined when $y = 1$. We *can find $x(y)$ here, however - plug $y = 1$ into $G(x, y) = 0$ and we find $x = 2 \pm \sqrt{6}$.*

Now we check our critical points and boundary points. At the critical points, $(2, 1 \pm \sqrt{3})$, we have $F(x, y) = -2$. This is where you should go back to that 'x'. If we hadn't thought of boundary points, assuming that there aren't any on an ellipse, or if we hadn't thought about how our approach was related to implicit functions, we might have thought that these two points were the only possible solutions. And then, finding that the objective function takes the same value on both would have been disturbing. After all, we know Weierstraß applies, so we can find both maxima and minima. But the only way that these points could be both and give the same value for the objective function would be if the function were constant on the constraint set, in which case we shouldn't have singled out these two points as critical. Since we've already noticed that F can be split into two convex functions, one of x and one of y, it should be clear that the critical points are minima. But where are the maxima? Here is where we need to know that there are, in fact, boundaries - and that we should check them. $F(2 \pm \sqrt{6}, 1) = 0$, and these points are our maximizers. One other way to see that they must be the maxima is to note that $F(x, y) = x^2 + y^2 - 4x - 2y = (x-2)^2 + (y-1)^2 - 5$ is simply a function of the distance from the center of the ellipse described by the constraint set, minus a constant - and if you draw the ellipse described by $G(x, y) = 0$, it is clear that the distance from the curve to its center is maximized at the points $(2 \pm \sqrt{6}, 1)$.

4. *Maximize and minimize $F(x, y)$ subject to $G(x, y) \leq 0$. Is the Weierstraß theorem applicable?*

   Solution: *Yes, the theorem applies; since $G(x, y)$ is convex in both x and y, it is the interior and boundary of the ellipse which satisfies the constraint, and this is a compact set. We've already done all the work we have to do to maximize and minimize. From (2) we know that the minimum in the interior is at $(2, 1)$, which gives $F(2, 1) = -5$, less than the minima on the boundary, so this is our global max. We know that there is no maximum in the interior, so we have to look on the boundary; we found these points in (3) - $F(2 \pm \sqrt{6}, 1) = 0$.*

5. *Maximize and minimize $F(x, y)$ subject to $G(x, y) \geq 0$. Is the Weierstraß theorem applicable?*

   Solution: *Now Weierstraß does not apply; the constraint is sat-*

isfied on all of $\mathbb{R}^2$ except the interior of the ellipse, which is not a bounded set. And indeed, because the function is convex in both x and y, there are no maxima (another way to see this is to again think of F as giving the distance from the point $(2, 1)$, less a constant). Any minima will have to happen on the boundary of the ellipse, and we found these points in (3): $F(2, 1 \pm \sqrt{3}) = -2$.