HW1 ECE 719 Optimal systems

Spring 2016
Electrical engineering department
University of Wisconsin, Madison

Instructor: Professor B Ross Barmish

By

Nasser M. Abbasi

December 30, 2019

Contents

List of Figures

List of Tables

0.1 Problem 1

PROBLEM DESCRIPTION

Barmish

ECE 719 – Homework Multilinear

Suppose U is a hypercube in $\mathbf{R}^n$ and $J : U \to \mathbf{R}$ is multilinear function. Argue that the maximum of $J(u)$ over U is attained at a vertex of U. **Remarks**: If your argument involves working with one coordinate at a time, I suggest you review your solution to this problem to see if it is consistent with your solution to the next problem called "Homework Multilinear Revisited." Also note that the minumum of J is also attained at a vertex.

SOLUTION

A multilinear function $f(x_1, \cdots, x_n)$ is one which is linear with respect to each of its independent variables taken one at a time. In other words, when fixing all the independent variables except for one, then it reduces to a linear function in the one variable which is free to change. For an example, for $n = 2$,

$$\boxed{f\left(x,y\right) = ax + by + cxy}$$

is a multilinear function in x, y. When fixing x to some specific value x_0 in $\Re$, the above becomes

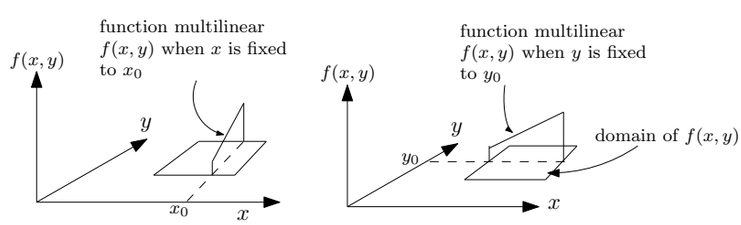
$$\begin{aligned} f\left(x,y\right)\Big|_{x=x_0} &= ax_0 + by + cx_0y \\ &= y\left(b + cx_0\right) + ax_0 \\ &= Ay + B \end{aligned}$$

Where all the constants a, b, c, x_0 have been combined into A and B. Similarly when fixing $y = y_0$, then

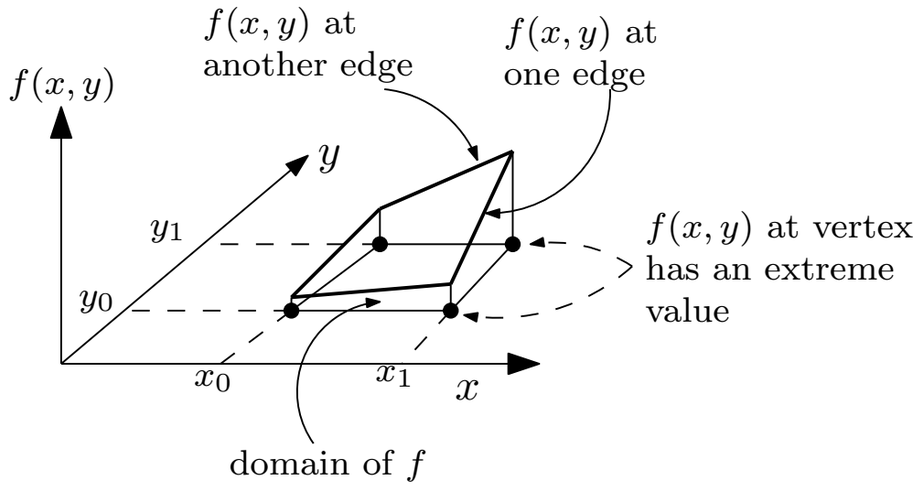
$$f\left(x,y\right)\Big|_{y=y_0} = Cx + D$$

A linear function has it extreme values at the start or at the end of its allowed range of values (The function can be either increasing or decreasing or a constant), this shows that $f(x_1, \cdots, x_n)$ will have its extreme values at one end of the boundaries of each of its variables $x_1, \cdots, x_n$.

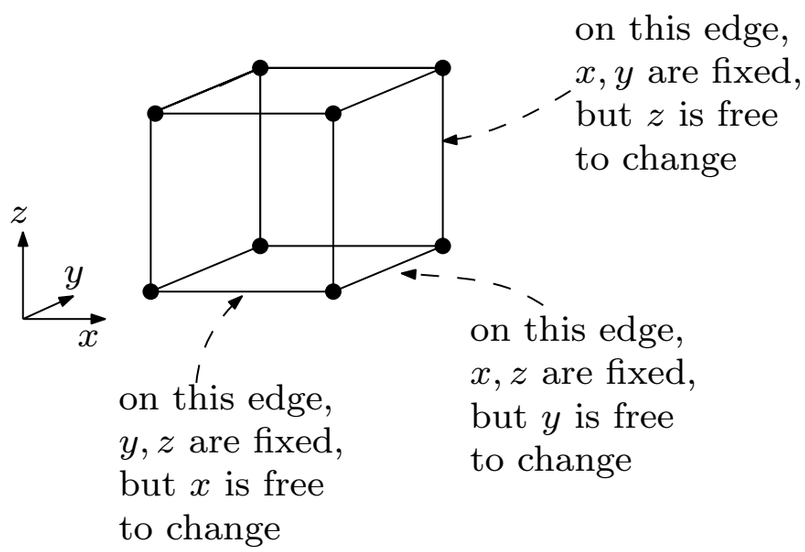
To illustrate what was said so far, taking $n = 2$ and fixing $x = x_0$, then the function will be $f\left(x,y\right)\Big|_{x=x_0}$ and when fixing $y = y_0$ the function will be $f\left(x,y\right)\Big|_{y=y_0}$



To show that the extreme points must be at a "corner" or a vertex, is now straight forward. From the above, the extreme value of the multilinear function must be on an edge. But on any edge, only one of the coordinates is free to change while all the others are fixed. Therefore on any edge of the hypercube (when in higher dimensions) the multilinear function is linear in only one of the parameters at an edge. Hence the function must be either increasing or decreasing on that edge (or be constant). Therefore the function extreme values on the edge is where the edge starts or ends, which is a vertex node. This is illustrated by this diagram for the $\Re^2$case.



The following illustrates the case for $\Re^3$ by showing few edges and (the function value $f(x,y,z)$ is hard to show here, since we would need fourth dimension).



This process carry over to higher dimensions hypercube.

0.2 Problem 2

PROBLEM DESCRIPTION

Barmish

ECE 719 – Homework "Multilinear Revisited"

For the three-variable multilinear function

$$J(u) = 8u_1u_2u_3 - 4u_1u_2 - 4u_1u_3 - 4u_2u_3 + 2u_1 + 2u_2 + 2u_3 - 1$$

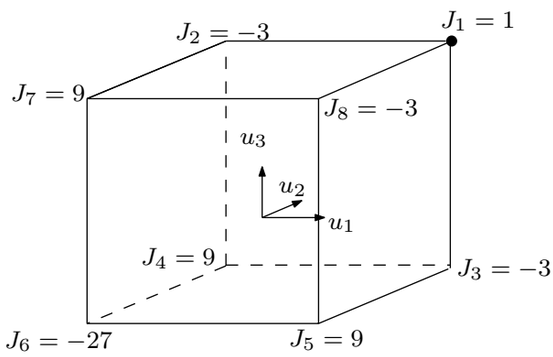
with constraints $|u_i| \leq 1$ for $i = 1, 2, 3$, let

$$u^0 = (1, 1, 1)$$

be an initial guess for the minimizer. Now carry out a sequence of one-variable optimizations beginning with u_1 and obtain successive refinements of the solution; i.e., hold $u_2 = u_3 = 1$ and optimize u_1 to obtain $\hat{u}_1$. Then with $u_1 = \hat{u}_1$ and $u_3 = 1$ held fixed, optimize u_2. Continue this process to optimize u_3. Finally, begin additional one-variable optimization cycles by returning to u_1, etc. Does this process converge to a minimizer u^*? Discuss.

SOLUTION

The function $J(u)$ is defined on a cube. Since it is multilinear in u_1, u_2, u_3, the minimizer point must be at one of the 8 vertices of the cube as shown in problem one. The optimization method the problem asks to perform *does not converge* to a minimizer u^* in general. And it does not in this problem. The value of J at each corner are[1] as shown below



In the first stage, we select between J_1 and J_2, (this is the result of fixing $u_2 = u_3 = 1$ and optimizing $J(u)$ to find $\hat{u}_1$). We find that J_2 wins, since it is smaller. Then we select between J_2 and J_7 and find that J_2 still wins. Then we select between J_2 and J_4, and find that J_2 also wins. So at the end of the first phase we mark J_2 as the winner (the minimum so far).

Now we go back to J_1 but select the second edge leaving this node, and select between J_1 and J_8, and find that J_8 wins. Now we select between J_8 and J_7, and find that J_8 wins. Then select between J_8 and J_5, and J_8 still wins. This ends the second phase.

Between $J_8 = -3$ and $J_2 = -3$ (winner of phase one), there is a tie so far.

We go back to J_1 (this is the final stage) and now take the third edge leaving this node, and select between J_1 and J_3. J_3 wins. Then select between J_3 and J_4 and J_3 wins. Then select between J_3 and J_5 and $J_3 = -3$ still is the winner.

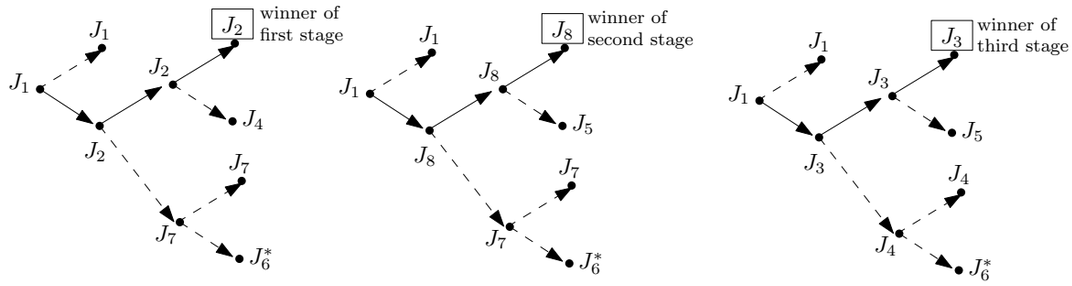
From vertex J_1 we have followed this algorithm over all the three edges leaving it, and found that overall minimum is $J(u) = -3$. If we continue this process, it will just repeat all

[1] small matlab script in the appendix

over.

But $J(u) = -3$ is not the correct value for the global minimum, since the global minimum is at vertex $J_6^* = -27$, which we never got the chance to compare with due to the nature of how this algorithm works. If, for example, J_7 had been smaller than J_2, say -4, then we would had the chance to select J_7 and then compare J_6 with J_7 to find that it is the overall minimum. So this algorithm is *not guaranteed to converge to the global minimum* in general.

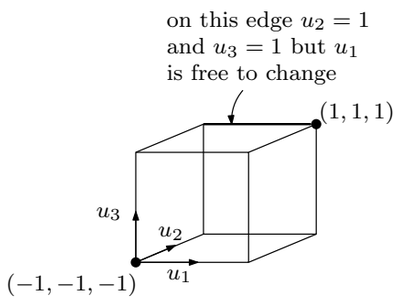
Here is the decision tree used for the above process.



We see that the global minimum at vertex $J_6 = -27$ was not visited. Wrong turn was taken in each stage.

0.2.1 Appendix

This appendix can be skipped. It shows the basic calculations for the first phase for illustration, and Matlab code used. Starting at (1,1,1), and fixing $u_2 = u_3 = 1$.



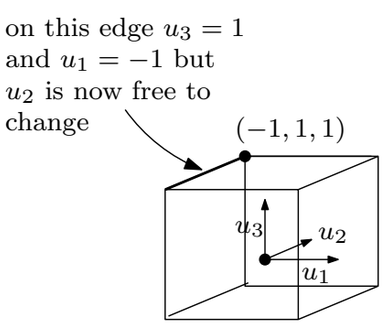
Hence on the first edge above, we have

$$J(u_1,u_2,u_3) = 8u_1u_2u_3 - 4u_1u_2 - 4u_1u_3 - 4u_2u_3 + 2u_1 + 2u_2 + 2u_3 - 1$$

Fixing $u_2 = u_3 = 1$ gives

$$\begin{aligned} J(u_1,1,1) &= 8u_1 - 4u_1 - 4u_1 - 4 + 2u_1 + 2 + 2 - 1 \\ &= 2u_1 - 1 \end{aligned}$$

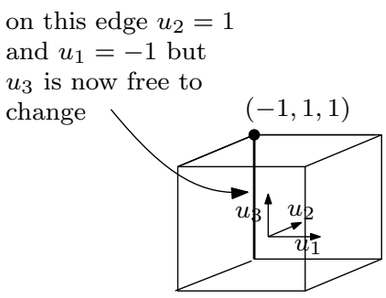
This is minimum when $u_1 = -1$. Hence $\hat{u}_1 = -1$. Now we see that on the above edge, J is smaller on vertex (–1,1,1) than on (1,1,1). Now we look at the next edge, where $u_1 = -1$ and $u_3 = 1$



Fixing u_3 at 1 and $u_1 = \hat{u}_1 = -1$ then

$$\begin{aligned} J(\hat{u}_1, u_2, u_3) &= J(-1, u_2, 1) \\ &= 8\hat{u}_1 u_2 u_3 - 4\hat{u}_1 u_2 - 4\hat{u}_1 u_3 - 4u_2 u_3 + 2\hat{u}_1 + 2u_2 + 2u_3 - 1 \\ &= -8u_2 + 4u_2 + 4 - 4u_2 - 2 + 2u_2 + 2 - 1 \\ &= 3 - 6u_2 \end{aligned}$$

Hence $J(\hat{u}_1, u_2, 1)$ is minimum when $u_2 = 1$. Therefore $\hat{u}_2 = 1$. This tells us that J at vertex $(-1,1,1)$ is smaller than on $(-1,1,-1)$. Traveling down the edge between $(-1,1,1)$ and $(-1,1,-1)$, which is done by fixing u_2, u_1 and changing u_3 gives



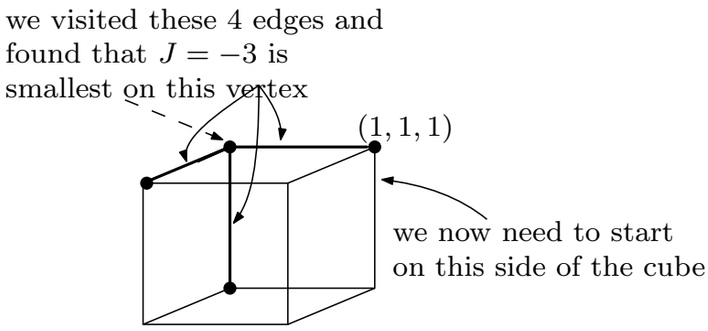
Now we need to find $\hat{u}_3$

$$\begin{aligned} J(\hat{u}_1, \hat{u}_2, u_3) &= J(-1, 1, u_3) \\ &= 8\hat{u}_1 \hat{u}_2 u_3 - 4\hat{u}_1 \hat{u}_2 - 4\hat{u}_1 u_3 - 4\hat{u}_2 u_3 + 2\hat{u}_1 + 2\hat{u}_2 + 2u_3 - 1 \\ &= -8u_3 + 4 + 4u_3 - 4u_3 - 2 + 2 + 2u_3 - 1 \\ &= 3 - 6u_3 \end{aligned}$$

This is minimum at $u_3 = 1$, Therefore $\hat{u}_3 = 1$. This means that J is still smallest at vertex $(-1,1,1)$. We have so far visited 3 edges, and looked at 4 vertices and found that J is smallest at $(-1,1,1)$.

$$\begin{aligned} J(-1,1,1) &= 8u_1 u_2 u_3 - 4u_1 u_2 - 4u_1 u_3 - 4u_2 u_3 + 2u_1 + 2u_2 + 2u_3 - 1 \\ &= -8 + 4 + 4 - 4 - 2 + 2 + 2 - 1 \\ &= -3 \end{aligned}$$

Here is a diagram to illustrate what we did so far



We now repeat the process starting from (1,1,1). Fixing $u_1 = 1$, $u_3 = 1$, but vary u_2. The calculation is similar to the above, and will not be shown. A small Matlab script is given below that was used to verify the results.

```
function nma_HW1_problem2_ECE719
%function to evaluate J at corner of the cube and do
%some syms caluclations.
%1/22/16

syms u1 u2 u3;
J = 8*u1*u2*u3-4*u1*u2-4*u1*u3-4*u2*u3+2*u1+2*u2+2*u3-1;

%first find J value at all the corners, the coordinates  are
a      = {[-1 1] [-1 1] [-1 1]};
```

```
coords = allcomb(a{:});

%function to evalute J at each coordinate
f = @(c)subs(J,{u1,u2,u3},c);

%print values at each corner of the cube
vpa(arrayfun(@(i) f(coords(i,:)),1:length(coords),'Uni',false))

J0    = subs(J,{u2,u3},{1,1});
u1Hat = getuHat(J0,u1)
J0    = subs(J,{u1,u3},{u1Hat,1});
u2Hat = getuHat(J0,u2)
J0    = subs(J,{u1,u2},{u1Hat,u2Hat});
u3Hat = getuHat(J0,u3)

end

function uHat = getuHat(J0,u)
uHat = 1;
u0  = subs(J0,u,-1);
if u0<subs(J0,u,1)
    uHat = -1;
end
end
```

Output of the above is

```
>> nma_HW1_problem2_ECE719
[ -27.0, 9.0, 9.0, -3.0, 9.0, -3.0, -3.0, 1.0]
u1Hat =
-1
u2Hat =
1
u3Hat =
1
```

0.3 Problem 3

PROBLEM DESCRIPTION

Barmish

ECE 719 – Homework Quotient

Suppose $U \subset \mathbf{R}^n$ is a hypercube and let

$$J(u) \doteq \frac{N(u)}{D(u}$$

where $N(u)$ and $D(u)$ are multilinear functions with $D(u)$ non-vanishing on U. Argue that the maximum of $J(u)$ over U is achieved by an extreme point of U. (Note: The same result holds for the minimum).

SOLUTION

Since $N(u)$ is multilinear, its maximum and minimum values will be on a vertex. Similarly for $D(u)$. Therefore, we only need to compare the ratios $\frac{N(u)}{D(u)}$ on the vertices to find the largest ratio. For example, for $n = 2$, we look at the four ratios, $\frac{N_1}{D_1}, \frac{N_2}{D_2}, \frac{N_3}{D_3}, \frac{N_4}{D_4}$. Where N_i means the value of $N(u)$ at vertex i, and similarly for D_i.

Since one of these N_i will be the largest value that $N(u)$ can take, and one of these D_i

values will be the smallest $D(u)$ can take, then the maximum of $J(u) = \frac{N(u)}{D(u)}$ will be one of these four values.

It can not be at any other u_i location. Proof by contradiction: Let us assume there is a point somewhere else in the domain of $J(u)$, say an internal point u_i where $\frac{N_i}{D_i}$ was the largest. This would imply that N_i is so large as to make $\frac{N_i}{D_i}$ larger than any value at the vertices regardless of what D_i happened to be at u_i, which means that N_i is the maximum of $N(u)$, but this is not possible since the maximum of $N(u)$ must be at a vertex.

Or it could mean that D_i is so small such that $\frac{N_i}{D_i}$ is larger than any value at the vertices regardless of what N_i happened to be, which means that D_i is the minimum of $D(u)$, but this is also not possible, since the minimum of $N(u)$ is at a vertex. Therefore the maximum of $J(u)$ must be at a vertex and can not be at any other point.

0.4 Problem 4

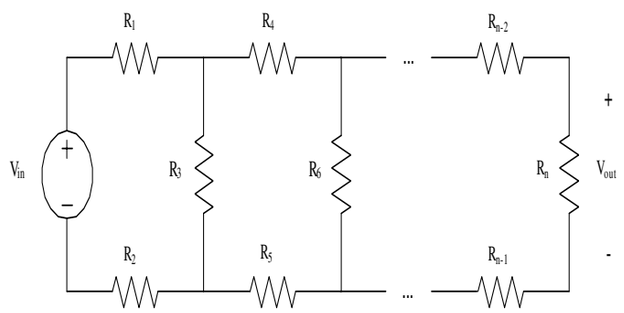
PROBLEM DESCRIPTION

Barmish

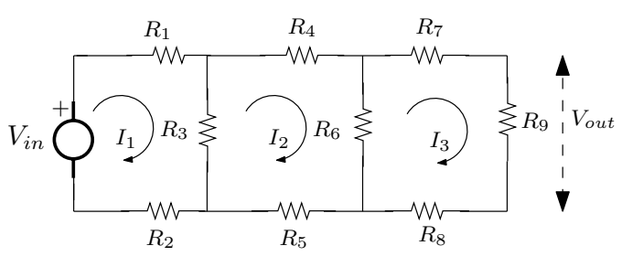
ECE 719 – Homework Ladder

For the ladder network below with $n = 9$ and input voltage $V_{in} = 1$, use symbolic manipulation in Matlab to find the output voltage $V_{out}(R)$ as a function of the R_i.

For resistor values with bounds $90 \leq R_i \leq 110$ for $i = 1, 2, \ldots, 9$, find the maximum value of the output voltage $V_{out}(R)$.



SOLUTION



The total network resistance (Input impedance) is

$$Z_{in} = R_1 + R_2 + R_3 || (R_4 + R_5 + R_6 || (R_7 + R_9 + R_8))$$

Using $X||Y = \frac{XY}{X+Y}$ since in parallel the above become

$$\begin{aligned} Z_{in} &= R_1 + R_2 + \frac{R_3 (R_4 + R_5 + R_6 || (R_7 + R_9 + R_8))}{R_3 + (R_4 + R_5 + R_6 || (R_7 + R_9 + R_8))} \\ &= R_1 + R_2 + \frac{R_3 \left(R_4 + R_5 + \frac{R_6 (R_7 + R_9 + R_8)}{R_6 + (R_7 + R_9 + R_8)} \right)}{R_3 + \left(R_4 + R_5 + \frac{R_6 (R_7 + R_9 + R_8)}{R_6 + (R_7 + R_9 + R_8)} \right)} \end{aligned}$$

The above is the overall ladder network resistance. Let $X = R_4 + R_5 + \frac{R_6 (R_7 + R_9 + R_8)}{R_6 + (R_7 + R_9 + R_8)}$ to simplify the equation

$$Z_{in} = R_1 + R_2 + \frac{X R_3}{R_3 + X}$$

The output impedance is

$$\begin{aligned} Z_{out} &= R_3 \| \left(R_4 + R_5 + R_6 \| \left(R_7 + R_9 + R_8\right)\right) \\ &= \frac{XR_3}{R_3 + X} \end{aligned}$$

Hence V_{out} is now found, using $V_{in} = 1$, from

$$\begin{aligned} \frac{V_{out}}{V_{in}} &= \frac{Z_{out}}{Z_{in}} \\ V_{out} &= \frac{\frac{XR_3}{R_3+X}}{R_1 + R_2 + \frac{XR_3}{R_3+X}} = \frac{XR_3}{R_1\left(R_3 + X\right) + R_2\left(R_3 + X\right) + XR_3} \\ &= \frac{XR_3}{X\left(R_1 + R_2 + R_3\right) + R_1R_3 + R_2R_3} \end{aligned}$$

Since V_{out} multilinear function in $R_i, i = 1 \cdots 9$, the maximum and minimum will occur at the end range values of each resistance, which is 90 and 110. so there are 2^9 different cases to check. A small script is below which calculate these vertex values and shows the maximum V_{out} found. The maximum is

$$V_{out_{\max}} = 0.3147 \text{ volt}$$

Using R_1 and R_2 at 90 ohm, and the rest of the resistors using 110 ohm.

```
function nma_HW1_problem4_ECE719
%function to evaluate Vout at corners of the R^9
%Nasser M. Abbasi
%1/23/16

a = repmat({[90 110]},9,1);
v = allcomb(a{:});
r = arrayfun(@(i) vOut(v(i,:)),1:size(v,1));

%done. Now print min and max, and the vertix at each
[maxValue,indx] = max(r);
fprintf('max is %f at U=\n',maxValue);
v(indx,:)

[minValue,indx] = min(r);
fprintf('min is %f at U=\n',minValue);
v(indx,:)

end
function r = vOut(R)
%evaluate objective function at vertex. See HW
X = R(4)+R(5)+R(6)*(R(7)+R(9)+R(8))/(R(6)+(R(7)+R(9)+R(8)));
r = X*R(3)/(X*(R(1)+R(2)+R(3))+R(1)*R(2)+R(2)*R(3));
end
```

Which generates when run:

```
>> nma_HW1_problem4_ECE719
max is 0.314732 at U=
90    90   110   110   110   110   110   110   110
min is 0.225627 at U=
110   110    90    90    90    90    90    90    90
```

0.5 Problem 5

PROBLEM DESCRIPTION

Barmish

ECE 719 – Homework Common Sense

For constraint set $U \subseteq \mathbf{R}^n$ and $J : U \to \mathbf{R}$ suppose that for $k = 1, 2, \ldots, n$, the following condition is satisfied: With all u_i frozen except for $i = k$, the resulting one variable J function is minimized over U by $u_k = u_k^*$, independently of the frozen values of the remaining u_i.

(a) Argue that $u^* = (u_1^*, u_2^*, \ldots, u_n^*)$ minimizes $J(u)$ over U with all u_i being allowed to vary simultaneously.

(b) Give an example with $n = 2$ for which the result in Part (a) applies.

(c) Give an example with $n = 2$ for which the result in Part (a) does not apply.

SOLUTION

0.5.1 Part(a)

Since u_k^* minimizes $f(u)$ when all u_i are fixed except for u_k, then this is the same as saying that solving for $\frac{\partial f}{\partial u_k} = 0$ gives u_k^*. Since this is the necessary condition for an extreme point from unconstrained calculus (we still need to check the Hessian for u_k^* being the minimum or the maximum, but we are told here it is a minimizer).

But $\frac{\partial f}{\partial u_k}$ is partial derivative of $f(u)$ w.r.t. to u_k when all other u_i are fixed (by definition). For $\Re^n$ this carries over and becomes the gradient. Therefore

$$\nabla f = 0$$

$$\begin{pmatrix} \frac{\partial f}{\partial u_1} \\ \frac{\partial f}{\partial u_2} \\ \vdots \\ \frac{\partial f}{\partial u_n} \end{pmatrix} = 0$$

Leads to minimum being at $\begin{pmatrix} u_1^* \\ u_2^* \\ \vdots \\ u_n^* \end{pmatrix}$ since we are told that each $\frac{\partial f}{\partial u_k} = 0$ results in u_k^* as the solution.

0.5.2 Part(b)

An example where the above applies is $J\left(x, y\right) = xy$ on $U = [0,1]$ Since

$$\nabla f = 0$$

$$\begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Hence a minimizer is $u^* = \begin{pmatrix} x = 0 \\ y = 0 \end{pmatrix}$ and $J(u^*) = 0$ is the global minimum.

0.5.3 Part(c)

An example where part (a) does not apply is $J\left(x, y\right) = xy$ on $U = [-1,1]$. Now $u^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is not the minimizer, since if $x = -1$ and $y = 1$ or if $x = 1$ and $y = -1$, we find $J(u^*) = -1 < 0$.

The following is a plot of $J(x,y) = xy$ of different sets of constraints to illustrate part(b) and (c).