MIDTERM EXAM - ANSWER KEY

Problem 1: Local vs Global Conditions. (15 points)

a) Let $f : \mathbb{R}^n \to \mathbb{R}$ and suppose that there exists $\bar{\mathbf{x}} \in \mathbb{R}^n$ and an open set U containing $\bar{\mathbf{x}}$ such that $f(\bar{\mathbf{x}}) \leq f(\mathbf{x})$, for all $\mathbf{x} \in U$. *Do not assume that f is a differentiable function.*

i) Identify a condition on f that is sufficient to ensure that $f(\bar{\mathbf{x}}) < f(\mathbf{x})$, for all $\mathbf{x} \in \mathbb{R}^n$. Prove that the condition is sufficient.

The condition is that f is a strictly convex function. One definition of strict convexity is that the epigraph of the function, is a strictly convex set. (The epigraph of a function, denoted epi(f), is the "stuff above the graph," i.e., epi$(f) = \{(\mathbf{x}, y) \in \mathbb{R}^n \times \mathbb{R} : y \geq f(\mathbf{x})\}$. To prove that strict convexity is sufficient for the property specified in i) assume that there exists $\bar{\mathbf{x}} \in \mathbb{R}^n$ and an open set U containing $\bar{\mathbf{x}}$ such that $f(\bar{\mathbf{x}}) \leq f(\mathbf{x})$, for all $\mathbf{x} \in U$, and $\mathbf{y} \in \mathbb{R}^n$, such that $f(\mathbf{y}) \leq f(\bar{\mathbf{x}})$. We will show that f is not strictly convex, by showing that the line segment joining $\big(\bar{\mathbf{x}}, f(\bar{\mathbf{x}})\big)$ and $\big(\mathbf{y}, f(\mathbf{y})\big)$ is not contained in the interior of epi(f). For $\lambda \in (0,1)$, let $\boldsymbol{\psi}(\lambda) = \big(\lambda\bar{\mathbf{x}} + (1-\lambda)\mathbf{y}, \lambda f(\bar{\mathbf{x}}) + (1-\lambda)f(\mathbf{y})\big)$. Since $f(\mathbf{y}) \leq f(\bar{\mathbf{x}})$, it follows that. for all $\lambda \in (0,1)$, $\psi_{n+1}(\lambda) = \lambda f(\bar{\mathbf{x}}) + (1-\lambda)f(\mathbf{y})\big) \leq f(\bar{\mathbf{x}})$. But since U is an open set, there exists $\epsilon > 0$, such that for $0 < \lambda < \epsilon$, $(\psi_1(\lambda), ...\psi_n(\lambda)) \in U$. By assumption, then, for $0 < \lambda < \epsilon$, $f(\bar{\mathbf{x}}) \leq f(\psi_1(\lambda), ...\psi_n(\lambda))$ and hence $\boldsymbol{\psi}(\lambda)$ does not belong to epi(f).

ii) Identify a second condition on f that is sufficient to ensure that $f(\bar{\mathbf{x}}) < f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$ and is a *strictly weaker* condition than the first one you identified. Prove that the condition is sufficient.

The condition is that f is a strictly quasi-convex function. A necessary condition for strict quasi-convexity is that every lower contour set of f is a strictly convex set. To prove that strict quasi-convexity is sufficient for the property specified in ii) assume that there exists $\bar{\mathbf{x}} \in \mathbb{R}^n$ and an open set U containing $\bar{\mathbf{x}}$ such that $f(\bar{\mathbf{x}}) \leq f(\mathbf{x})$, for all $\mathbf{x} \in U$, and $\mathbf{y} \in \mathbb{R}^n$, such that $f(\mathbf{y}) \leq f(\bar{\mathbf{x}})$. Let $LC(\bar{\mathbf{x}})$ denote the lower contour set of f corresponding to $f(\bar{\mathbf{x}})$. We will show that f is not strictly quasi-convex, by showing that the line segment joining $\bar{\mathbf{x}}$ and $\mathbf{y}$ is not contained in the interior of $LC(\bar{\mathbf{x}})$. By assumption, $\mathbf{y} \in LC(\bar{\mathbf{x}})$. For $\lambda \in (0,1)$, let $\boldsymbol{\psi}(\lambda) = \lambda\bar{\mathbf{x}} + (1-\lambda)\mathbf{y}$. Since U is an open set, there exists $\epsilon > 0$, such that for $0 < \lambda < \epsilon$, $\boldsymbol{\psi}(\lambda) \in U$. By assumption, then, for $0 < \lambda < \epsilon$, $f(\boldsymbol{\psi}(\lambda)) \geq f(\bar{\mathbf{x}})$, and hence $\boldsymbol{\psi}(\lambda)$ does not belong to the interior of $LC(\bar{\mathbf{x}})$. Hence we have established that the line segment joining $\bar{\mathbf{x}}$ and $\mathbf{y}$ is not contained in the interior of $LC(\bar{\mathbf{x}})$.

iii) Demonstrate with an example of a function from $\mathbb{R}^2$ to $\mathbb{R}$ that the second condition is strictly weaker than the first.

$f(\mathbf{x}) = -x_1x_2$ is strictly quasi-convex on $\mathbb{R}^2_+$ but not strictly convex. To see that it is not strictly convex observe that its Hessian is $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$, which is not positive definite. On the other hand, fix $\mathbf{x}$ and consider dx such that $\nabla f(\mathbf{x})\mathrm{dx} = (x_2,\ x_1) \cdot \mathrm{dx} = 0$. From Problem 5 (reversed), we need do show that $\mathrm{dx}'\mathrm{Hf}(\mathbf{x})\mathrm{dx} > 0$. Necessarily, $\mathrm{dx} = \lambda(-x_1,\ x_2)$, for some $\lambda \in \mathbb{R}$. Hence $\mathrm{dx}'\mathrm{Hf}(\mathbf{x})\mathrm{dx} = 2\lambda^2 x_1 x_2 > 0$.

Problem 2: Vector Spaces. (20 points)

Fix $n > 2$ and let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function and fix $\mathbf{x} \in \mathbb{R}^n$. For $i = 1, ...n$, define the row vector ψ^i by

$$\psi^i_j = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{otherwise} \end{cases}$$

where ψ^i_j denotes the j'th element of the row vector ψ^i. Let $g = \nabla f(\mathbf{x}) \in \mathbb{R}^n$ and consider the set of vectors $V = \{\mathbf{v}^1, \mathbf{v}^2, ..., \mathbf{v}^n\} \subset \mathbb{R}^{n+1}$, defined by $\mathbf{v}^i = (\psi^i, g_i) \in \mathbb{R}^{n+1}$, for each i.

a) for an arbitrary vector $\boldsymbol{\alpha} \in \mathbb{R}^n$, write down, in the most economical possible form (i.e., using the fewest symbols you can), the linear combination of the elements of V where the weight on the i'th element of V is α_i.

   $(\alpha_1, ...\alpha_n, \sum_{i=1}^n \alpha_i g_i)$. or alternatively. $(\alpha_1, ...\alpha_n, \nabla f(\mathbf{x}) \cdot \boldsymbol{\alpha})$.

b) Write down the vector space W which is *the span* of V. (Hint: the ideal answer to this question is in the following form: "$\mathbf{w} \in \mathbb{R}^?$ belongs to W if and only ???)".

   $\mathbf{w} \in \mathbb{R}^{n+1}$ belongs to W if and only $w_{n+1} = \nabla f(\mathbf{x}) \cdot (w_1, ..., w_n)'$.

c) What is the dimension of W? Formally support your answer.

   In our answer to the following question, we establish that V is a basis for W. Since, by definition, the dimension of a vector space is the number of elements in any basis for that space, and V has n elements, it follows that the dimension of W is n.

d) Verify that V is a basis for W.

   Consider the element $\mathbf{v}^i = (\psi^i, g_i) \in V$. Clearly, $v^i_{n+1} = g_i = \nabla f(\mathbf{x}) \cdot (\psi^i_1, ..., \psi^i_n)'$. Therefore $\mathbf{v}_i \in W$. Since i was chosen arbitrarily, $V \subset W$. Next note that the elements of V are linear independent: $\mathbf{v}^i$ is the unique member of V whose i element is non-zero, therefore $\mathbf{v}^i$ cannot be written as a linear combination of the other members of V. It now follows that V is a minimal spanning set, i.e., for arbitrarily chosen i, $V \sim \{\mathbf{v}^i\}$ does not span W. Therefore V is a basis for W.

e) Write down an different basis for W, and verify that it is indeed a basis.

   For $i = 1, ..., n$, let $\boldsymbol{\gamma}^i = 2\mathbf{v}^i$. The set $\Gamma = \{\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2, ..., \boldsymbol{\gamma}^n\} \subset \mathbb{R}^{n+1}$ is another basis for W. To verify that this set spans W, consider $\mathbf{w} \in W$. Since V is a basis for W, there exists $\mathbf{t} \in \mathbb{R}^n$ such that $w = \sum_{i=1}^n t_i \mathbf{v}^i$. Let $\boldsymbol{\tau} = 0.5\mathbf{t}$ and observe that

$$\sum_{i=1}^n \tau_i \boldsymbol{\gamma}^i = 0.5 \sum_{i=1}^n t_i \boldsymbol{\gamma}^i = 0.5 \sum_{i=1}^n t_i 2\mathbf{v}^i = \sum_{i=1}^n t_i \mathbf{v}^i = \mathbf{w}$$

   Hence Γ spans W. Also, Γ is a minimal spanning set, by an argument identical to the one that showed that V was a minimal spanning set. Hence Γ is a basis for W.

f) Write down a minimal spanning set for W that contains $n + 1$ elements. Verify that it is a minimal spanning set, and that it is not a basis.

For $i = 1, ..., n-1$, let $\boldsymbol{\gamma}^i = \mathbf{v}^i$. Define $\boldsymbol{\gamma}^n = (\psi^n, g_n - 1)$ and $\boldsymbol{\gamma}^{n+1} = (\psi^n, g_n + 1)$. Now let $\Gamma = \{\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2, ..., \boldsymbol{\gamma}^n, \boldsymbol{\gamma}^{n+1}\}$. Clearly $g_n - 1 \neq g_n = \nabla f(\mathbf{x}) \cdot (\psi_1^i, ..., \psi_n^i)'$. Hence $\boldsymbol{\gamma}^n \notin W$, so that Γ is not a basis for W. However, Γ does span W. To see this, consider $\mathbf{w} \in W$. Since V is a basis for W, there exists $\mathbf{t} \in \mathbb{R}^n$ such that $w = \sum_{i=1}^n t_i \mathbf{v}^i$. Define $\boldsymbol{\tau} \in \mathbb{R}^{n+1}$ by $\tau_i = t_i$, for $i = 1, ..., n-1$, and let $\tau_n = \tau_{n+1} = 0.5 t_n$. Clearly, $\tau_n \boldsymbol{\gamma}^n + \tau_{n+1} \boldsymbol{\gamma}^{n+1} = t_n \mathbf{v}^n$, so that $w = \sum_{i=1}^{n+1} \tau_i \boldsymbol{\gamma}^i$. Therefore, Γ spans W. Finally, we need to show that Γ is a minimal spanning set: for $i = 1, ..., n-1$ the argument is identical to the one we gave in 2d). Moreover $\{\boldsymbol{\gamma}^n, \boldsymbol{\gamma}^{n+1}\}$ is clearly a linear independent set. Hence Γ is a minimal spanning set.

g) Write down a two-dimensional subspace of W.

Let $\Gamma = \{\boldsymbol{\gamma} \in \mathbb{R}^{n+1} : \gamma_3 = ... = \gamma_n = 0; \gamma_{n+1} = \nabla f(\mathbf{x}) \cdot (\gamma_1, ..., \gamma_n)'\}$. Clearly, Γ is spanned by the set $\{\mathbf{v}^1, \mathbf{v}^2\}$. Moreover, $\Gamma \subset V$. Hence Γ is a two-dimensional subspace of V.

h) W corresponds to a familiar object in multivariable calculus. What is this object? Support your answer by relating W to the definition of the object you've identified.

W is the graph of the differential of f at $\mathbf{x}$. To see this, let $L^{f,\mathbf{x}}(\cdot)$ denote the differential of f at $\mathbf{x}$. For $\mathrm{dx} \in \mathbb{R}^n$, $L^{f,\mathbf{x}}(\mathrm{dx}) = \nabla f(\mathbf{x}) \cdot \mathrm{dx}$. Therefore $(\mathrm{dx}, \nabla f(\mathbf{x}) \cdot \mathrm{dx})$ belongs to the graph of $L^{f,\mathbf{x}}$. But from our answer to 2 b), $(\mathrm{dx}, \nabla f(\mathbf{x}) \cdot \mathrm{dx}) \in W$. Therefore the graph of $L^{f,\mathbf{x}}$ is contained in W. Similarly, any point $(\mathrm{dx}, \nabla f(\mathbf{x}) \cdot \mathrm{dx}) \in W$ belongs to the graph of $L^{f,\mathbf{x}}$. Therefore, W is contained in the graph of $L^{f,\mathbf{x}}$. Therefore the two sets are equal.

i) Let $\mathbf{w}$ be a weighted combination of the elements of V, with the property that the norm of the vector of weights is unity. The object $\mathbf{w}$ corresponds to another familiar object in multivariable calculus. What is this object? Support your answer by relating $\mathbf{w}$ to the definition of the object you've identified.

By assumption, there exists $\mathbf{w} \in \mathbb{R}^n$ such that $||\mathbf{w}|| = 1$ and $\mathbf{w} = \sum_{i=1}^n w_i \mathbf{v}^i$. We will show that w_{n+1} is the directional derivative of f at $\mathbf{x}$ in the direction $(w_1, ..., w_n)$. To verify this, note from equations (1) and (3) of lecture mathCalculus2, that the directional derivative of f at $\mathbf{x}$ in the direction h is given by $\nabla f(\mathbf{x}) \cdot \frac{\mathbf{h}}{||\mathbf{h}||}$. Letting $h = (w_1, ...w_n)$, observe that since $||h|| = 1$, $\mathbf{w}_{n+1} = \nabla f(\mathbf{x}) \cdot \mathbf{h} = \nabla f(\mathbf{x}) \cdot \frac{\mathbf{h}}{||\mathbf{h}||}$. We have established, then, that $\mathbf{w}$ is a pair, consisting of a *direction* $(w_1, ...w_n)$ and the directional derivative of f at $\mathbf{x}$ in that direction.

Problem 3: Calculus. (15 points)
After he gave away the Chocolate Factory to Charlie, Willie Wonka turned to calculus. Being an different kind of guy, he decided he would do calculus a little differently. For a function $g : \mathbb{R}^2 \to \mathbb{R}$ and each $\mathbf{x} \in \mathbb{R}^2$, Willy defined

- the *positive werivative* of g at $\mathbf{x}$, denoted $g_+(\mathbf{x})$, as $\lim_{|k|\to\infty} \frac{\left(g(\mathbf{x}+(1,1)/k)-g(\mathbf{x})\right)}{\sqrt{2}/k}$
- the *negative werivative* of g at $\mathbf{x}$, denoted $g_-(\mathbf{x})$, as $\lim_{|k|\to\infty} \frac{\left(g(\mathbf{x}+(1,-1)/k)-g(\mathbf{x})\right)}{\sqrt{2}/k}$
- the *gwadient* of g as $\Delta g(\cdot) = (g_+(\cdot), g_-(\cdot))$.
- the *wifferential* of g at $\mathbf{x}$ is the linear function $L^{g,\mathbf{x}}(\mathrm{dx}) = \Delta g(\mathbf{x}) \cdot \mathrm{dx}$.

a) Using the wifferential, write an expression for the partial derivatives of g.

Note that $(1,0) = \frac{1}{\sqrt{2}}\frac{(1,1)}{\sqrt{2}} + \frac{1}{\sqrt{2}}\frac{(1,-1)}{\sqrt{2}}$. That is, the vector $(1,0)$ is a linear combination of the vectors $\frac{(1,1)}{\sqrt{2}}$ and $\frac{(1,-1)}{\sqrt{2}}$, with weights $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. Hence, $f_1(\mathbf{x}) = L^{g,\mathbf{x}}(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) = \Delta\, g(\mathbf{x}) \cdot (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. Similarly, $(0,1) = \frac{1}{\sqrt{2}}\frac{(1,1)}{\sqrt{2}} - \frac{1}{\sqrt{2}}\frac{(1,-1)}{\sqrt{2}}$. so that $f_2(\mathbf{x}) = L^{g,\mathbf{x}}(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) = \Delta g(\mathbf{x}) \cdot (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$.

For the remainder of the question, let $f = xy^2$.

b) Write down the expression for $\Delta f(\cdot)$.

$\Delta f(\cdot) = (y^2 + 2xy, y^2 - 2xy)/\sqrt{2}$.

c) Write down the wifferential of f at $(2,3)$

The wifferential of f at $(2,3)$ is $L^{f,\mathbf{x}}(\mathrm{dx}) = (21, -3)/\sqrt{2} \cdot \mathrm{dx}$.

d) Using the wifferential, compute the partial derivatives of f.

The first partial derivative of f is the directional derivative in the direction (1,0). Now as noted in the answer to 2(a)

$$(1,0) = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) + \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$$

while

$$(0,1) = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) - \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$$

so that

$$
\begin{aligned}
f_1(\mathbf{x}) &= \frac{1}{\sqrt{2}} f_+(\mathbf{x}) + \frac{1}{\sqrt{2}} f_-(\mathbf{x}) \\
&= \frac{1}{\sqrt{2}} \frac{(y^2 + 2xy)}{\sqrt{2}} + \frac{1}{\sqrt{2}} \frac{(y^2 - 2xy)}{\sqrt{2}} \\
&= 2y^2/2 \quad = \quad y^2 \\
f_2(\mathbf{x}) &= \frac{1}{\sqrt{2}} f_+(\mathbf{x}) - \frac{1}{\sqrt{2}} f_-(\mathbf{x}) \\
&= \frac{1}{\sqrt{2}} \frac{(y^2 + 2xy)}{\sqrt{2}} - \frac{1}{\sqrt{2}} \frac{(y^2 - 2xy)}{\sqrt{2}} \\
&= 4xy/2 \quad = \quad 2xy
\end{aligned}
$$

e) Comment on the relationship between the wifferential of f at $(2, 3)$ and the differential of f at $(2, 3)$. (Hint: the ideal answer to this question includes a word starting with “v” and another starting with “b”.)

The wifferential of f at $(2, 3)$ is identical to the differential of f at $(2, 3)$. As we have seen from question 2 h), the graph of the differential is a vector space, in fact a vector subspace of $\mathbb{R}^3$. A basis for this vector space is the pair of vectors $(1, 0, f_1), (0, 1, f_2)$. Any pair of linear independent unit length vectors, together with their directional derivatives, forms an alternative basis for this space. The vectors $(1, 1, f_+), (1, -1, f_-)$ are such a pair. Thus, the wifferential and the differential are two alternative representations of the same vector subspace of $\mathbb{R}^3$.

Problem 4: Taylor Theory. (25 points)

Consider the CES production function $f(x,y) = x^\rho + y^\rho$, where $\rho \in (0,1)$. (Actually this is a transformation of a CES production function, but never mind.)

a) Write down the gradient and the Hessian of this function

$$\nabla f(x,y;\rho) = \begin{bmatrix} \rho x^{\rho-1} \\ \rho y^{\rho-1} \end{bmatrix}; \; Hf(x,y;\rho) = \begin{bmatrix} \rho(\rho-1)x^{\rho-2} & 0 \\ 0 & \rho(\rho-1)y^{\rho-2} \end{bmatrix};$$

b) Write down the second order Taylor expansion of f at $(x,y;\rho)$. *Do not use matrix notation, i.e., multiply out the matrix.* Factor out as many terms as possible.

$$\tfrac{\rho}{2}\Big(x^{\rho-2}dx\,(2x + dx(\rho-1)) \quad + \quad y^{\rho-2}dy\,(2y + dy(\rho-1)) \Big)$$

c) Now let $y = x$,

i) express in the simplest possible way the expression for the second order Taylor expansion of $f(\cdot,\cdot;\rho)$ at (x,x).

$$\begin{aligned} & \frac{\rho x^{\rho-2}}{2}\Big(dx\,(2x + dx(\rho-1)) \quad + \quad dy\,(2x + dy(\rho-1)) \Big) \\ = \quad & \frac{\rho x^{\rho-2}}{2}\Big(2x\big(dx + dy\big) \quad + \quad (\rho-1)\big(dx^2 + dy^2\big) \Big) \qquad (1) \end{aligned}$$

ii) Characterize the conditions on (dx, dy) under which the first order Taylor expansion has the same sign as the second order expansion? (Hint: there are two cases to consider). (Hint2: I really meant (dx, dy) not (dx, dx)).

Consider expression (1). Since $\rho < 1$, the second term is always negative. Therefore, for *all* $dx + dy < 0$, the first order Taylor expansion has the same sign as the second order one. If $dx + dy = 0$, the first order Taylor expansion is zero and the second order expansion is negative, so that the two expansions never have the same sign. Now suppose that $dx + dy > 0$, so that the first order Taylor expansion is positive. Alternatively, let $dy = \alpha dx$, where $\alpha \in (-1, \infty)$. In this case, the first and second order expansions will have the same sign iff

$$0 \quad < \quad 2x(1+\alpha)dx \quad + \quad (\rho-1)(1+\alpha^2)dx^2$$

or

$$dx \quad < \quad \frac{2x(1+\alpha)dx}{(1-\rho)(1+\alpha^2)} \quad = \quad \frac{x}{(1-\rho)\left(\frac{1+\alpha}{2} + \frac{1}{1/\alpha+1}\right)}$$

iii) Fix x, ρ and an element $\mathbf{v}$ of the unit circle. Let $\bar\lambda(x,\mathbf{v},\rho)$ denote the largest λ such that for $0 < \theta < \lambda$ the first order Taylor expansion, i.e., $\nabla f(x,x)\cdot\theta\mathbf{v}$ has the same sign as $(f(x+\theta v_1, x+\theta v_2;\rho) - f(x,x;\rho))$. Discuss the comparative statics of $\bar\lambda(\cdot,\cdot,\cdot)$ with respect to $\mathbf{x}$, $\mathbf{v}$ and ρ. For the purposes this question only, I suggested that you should ignore the remainder term.

Let $\mathbf{v} = (dx,\ \alpha dx)$. From the answer to part (ii),

$$\bar{\lambda}(x,\mathbf{v},\rho) = \begin{cases} \infty & \text{if } \alpha < -1 \\ 0 & \text{if } \alpha = -1 \\ \frac{x}{(1-\rho)\left(\frac{1+\alpha}{2}+\frac{1}{1/\alpha+1}\right)} & \text{if } \alpha > -1 \end{cases}$$

(If $\alpha = -1$, the first order expansion is always zero, while $(f(x+\theta v_1, x+\theta v_2;\rho) - f(x,x;\rho)) < 0$. So in this case there exists no positive θ satisfying the required property. When $\alpha > -1$, we have

$$\begin{aligned} \frac{\partial\bar{\lambda}(x,\mathbf{v},\rho)}{\partial x} &> 0; \\ \frac{\partial\bar{\lambda}(x,\mathbf{v},\rho)}{\partial \rho} &> 0; \end{aligned}$$

Taking the derivative w.r.t. α requires some computation:

$$\frac{\partial\bar{\lambda}(x,\mathbf{v},\rho)}{\partial \alpha} > \frac{2x(\alpha^2+2\alpha-1)}{(\rho-1)(1+\alpha^2)^2}$$

It's easy to check that $\frac{\partial\bar{\lambda}(x,\mathbf{v},\rho)}{\partial \alpha}$ is positive, for $\alpha < \sqrt{2}-1$, and negative thereafter.

iv) for two values of ρ, preferably $1/3$ and $2/3$, sketch the level sets of $f(\cdot, \cdot; \rho)$ that pass through (1,1) and (2,2). Illustrate diagramatically the comparative statics properties you've identified in the previous sub-question.

In Fig. 1, the red-ish solid lines indicate the length of $\bar{\lambda}$, for $\rho = 1/3$, and the green-ish dotted lines indicate the length of $\bar{\lambda}$, for $\rho = 2/3$. $dy = -0.7dx$, for the thicker lines; $dy = -0.8dx$, for the thinner lines. Note that the lines are longer when $(x, x) = (2, 2)$ than when $(x, x) = (1, 1)$; longer when $\rho = 2/3$ than when $\rho = 1/3$; longer when $\alpha = -0.7$ than when $\alpha = -0.8$.

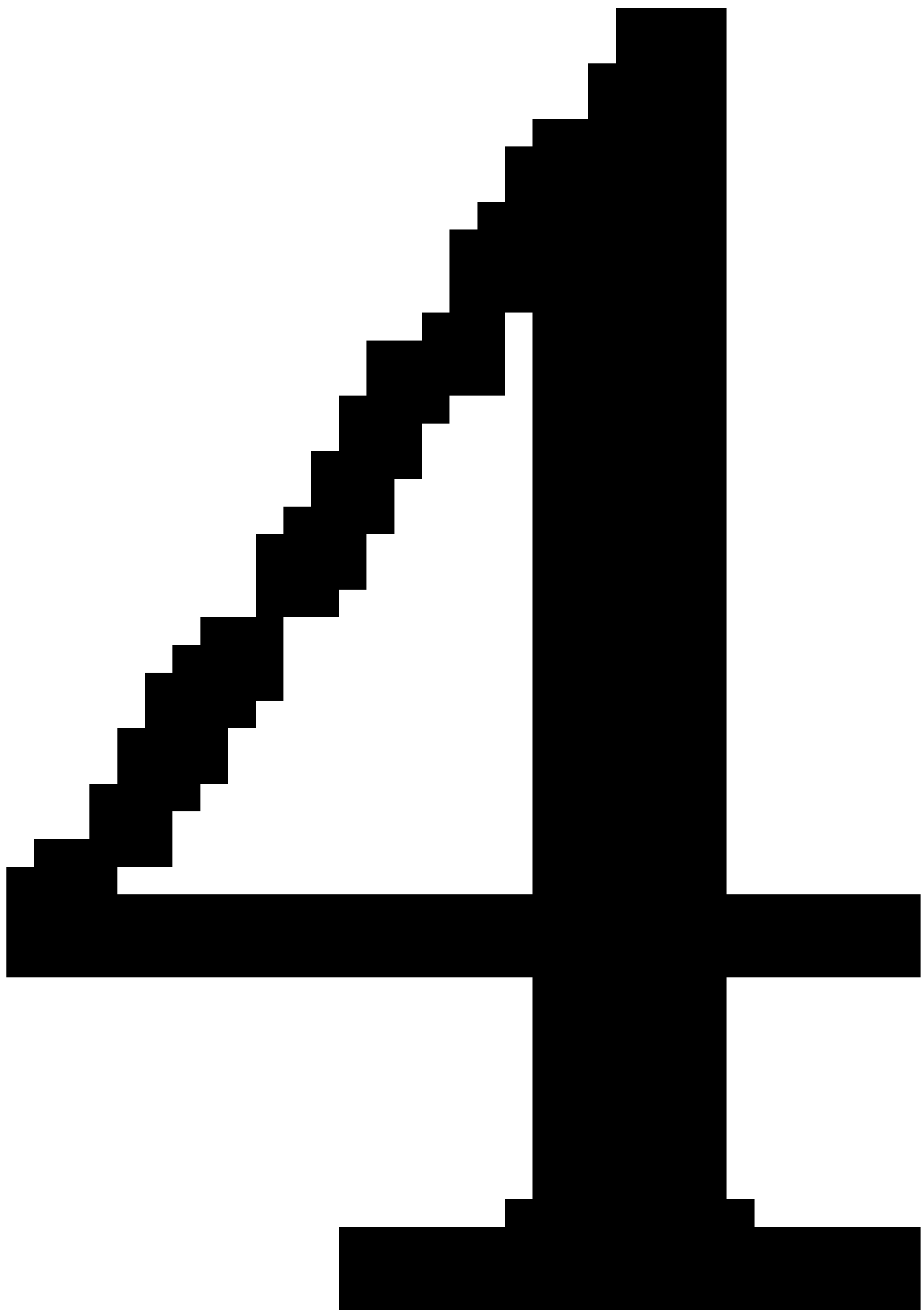

Figure 1. Answer to 4(c)(iv)

Problem 5: Second Order Conditions. (25 points)

a) For this part of the question, assume that f is twice continuously differentiable and that the gradient of f is never zero.

Let $\mathbb{T}(f, \mathbf{x})$ denote the plane that is tangent to the level set of f corresponding to $f(\mathbf{x})$ (that is, $\mathbb{T}(f, \mathbf{x})$ is the set of points that are perpendicular to $\nabla f(\mathbf{x})$). Then f is strictly quasi-concave if and only if for every $\mathbf{x}$, the set $\mathbb{T}(f, \mathbf{x}) \sim \{\mathbf{x}\}$ belongs to the strict lower contour set of f corresponding to $f(\mathbf{x})$. (The set $A \sim \{\mathbf{x}\}$ consists of all the elements of A excluding the element $\mathbf{x}$.)

Now suppose that f satisfies the following condition:

$$\text{for all } \mathbf{x} \text{ and all } \mathbf{dx} \text{ such that } \nabla f(\mathbf{x})'\mathbf{dx} = 0, \mathbf{dx}'\mathrm{Hf}(\mathbf{x})\mathbf{dx} < 0 \tag{A}$$

i) Use one of Taylor's theorems to prove that if f satisfies condition A then it satisfies the above definition of strict quasi-concavity.

Fix $\mathbf{x} \in \mathbb{R}^n$ such that condition A is satisfied at f. By Taylor Young's theorem, there exists $\epsilon > 0$, such that if $||\mathsf{dx}|| < \epsilon$, then the second order Taylor expansion of f about $\mathbf{x}$ in the direction dx has the same sign as $f(\mathbf{x} + \mathsf{dx}) - f(\mathsf{dx})$, specifically, for all $\mathsf{dx} \neq 0$ such that $\mathbf{x} + \mathsf{dx} \in \mathbb{T}(f, \mathbf{x})$,

$$\begin{aligned} \mathrm{sgn}(f(\mathbf{x}+\mathsf{dx}) - f(\mathbf{x})) &= \mathrm{sgn}\left(\nabla f(\mathbf{x}) \cdot \mathsf{dx} \quad + \quad \mathsf{dx}'\mathsf{Hf}(\mathbf{x})\mathsf{dx}/2\right) \\ &= \mathrm{sgn}\left(\mathsf{dx}'\mathsf{Hf}(\mathbf{x})\mathsf{dx}/2\right) \quad = \quad -1 \end{aligned}$$

This establishes that $f(\cdot)$ attains a strict local maximum on $\mathbb{T}(f, \mathbf{x})$ at $\mathbf{x}$, We will prove that $f(\cdot)$ attains a unique *global* maximum on $\mathbb{T}(f, \mathbf{x})$ at $\mathbf{x}$, which is equivalent to the statement that the set $\mathbb{T}(f, \mathbf{x}) \sim \{\mathbf{x}\}$ belongs to the strict lower contour set of f corresponding to $f(\mathbf{x})$. To accomplish this, we will show that if f is *not* globally maximized on $\mathbb{T}(f, \mathbf{x})$ at $\mathbf{x}$, then there must exist $\mathbf{y} \in \mathbb{R}^n$ such that condition A is violated at $\mathbf{y}$, proving the result. Suppose that f is *not* globally maximized on $\mathbb{T}(f, \mathbf{x})$ at $\mathbf{x}$, i.e., that there exists $\mathbf{x}' \neq \mathbf{x} \in \mathbb{T}(f, \mathbf{x})$ such that $f(\mathbf{x}') \geq f(\mathbf{x})$, and consider the line segment joining $\mathbf{x}$ and $\mathbf{x}'$, i.e., the set $\mathbb{L} = \{\mathbf{y} = \lambda\mathbf{x} + (1-\lambda)\mathbf{x}' : \lambda \in [0,1]\}$. Clearly $\mathbb{L} \subset \mathbb{T}(f, \mathbf{x})$. Moreover, $\mathbb{L}$ is clearly a compact set. Since f is continuous, by Weierstrass's theorem, there exists $\mathbf{y} \in \mathbb{L}$ that minimizes $f(\cdot)$ on $\mathbb{L}$. Moreover, since f attains a strict local maximum on $\mathbb{L} \subset \mathbb{T}(f, \mathbf{x})$ at $\mathbf{x}$, it follows that $f(\mathbf{y}) < f(\mathbf{x}) \leq f(\mathbf{x}')$, i.e., $\mathbf{y}$ lies on the *open* line segment joining $\mathbf{x}$ and $\mathbf{x}'$, (i.e., for some $\lambda \in (0,1)$, $\mathbf{y} = \lambda\mathbf{x} + (1-\lambda)\mathbf{x}'$). It now follows from Taylor Young's theorem that $\nabla f(\mathbf{y})\mathsf{dx} = 0$, for all dx such that $\mathbf{y} + \mathsf{dx} \in \mathbb{L}$. But this now implies that condition A fails at $\mathbf{y}$, since otherwise, by Taylor Young's theorem, there would exist some neighborhood of $\mathbf{y}$ such that

$$\begin{aligned} \mathrm{sgn}(f(\mathbf{y}+\mathsf{dx}) - f(\mathbf{y})) &= \mathrm{sgn}\left(\nabla f(\mathbf{y}) \cdot \mathsf{dx} \quad + \quad \mathsf{dx}'\mathsf{Hf}(\mathbf{y})\mathsf{dx}/2\right) \\ &= \mathrm{sgn}\left(\mathsf{dx}'\mathsf{Hf}(\mathbf{y})\mathsf{dx}/2\right) \quad = \quad -1 \end{aligned}$$

implying that $f(\cdot)$ is not minimized but *maximized* on $\mathbb{L}$.

ii) Provide an example to establish that condition (A) is not necessary for the above definition of strict quasi-concavity to hold.

$$f(\mathbf{x}), g(\mathbf{x})\mathbf{x}\{\mathbf{x} \in R^2_+ : g(\mathbf{x}) = b\}$$

FIGURE 2. Answer to 5(i)

Let $f(\mathbf{x}) = x_2 - x_1^4/4$. We have $\nabla f(\mathbf{x}) = (-x_1^3,\ 1)$, so that $\nabla f(\mathbf{x}) \cdot \mathrm{dx} = 0$ iff $x_2 = x_1^3$. Also, $\mathrm{Hf}(\mathbf{x}) = \begin{bmatrix} 3x_1^2 & 0 \\ 0 & 0 \end{bmatrix}$, so that $\mathrm{Hf}(0, x_2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, and condition (A) fails, for the vector (1 , 0). However, f satisfies the above definition of strict quasi-concavity. To see this, note that $\mathbf{y} \in \mathbb{T}(f, \mathbf{x})$ iff $\mathrm{dx} = \mathbf{y} - \mathbf{x}$ satisfies $dx_2 = x_1^3 dx_1$. But in this case,

$$\begin{aligned} f(\mathbf{x} + \mathrm{dx}) - f(\mathbf{x}) &= dx_1^3 - (x_1 + dx_1)^4/4 + x_1^4/4 \\ &= -\left(x_1^4 + dx_1^2(2x_1^2 + (2x_1 + dx_1)^2)\right)/4 \\ &< 0 \text{ provided } dx_1 \neq 0. \end{aligned}$$

This establishes that for every $\mathbf{y} \in \mathbb{T}(f, \mathbf{x})$, $f(\mathbf{y}) = f(\mathbf{x} + \mathrm{dx}) < f(\mathbf{x})$.

b) Consider the problem, maximize $f(\mathbf{x})$ s.t $g(\mathbf{x}) \leq b$, where f and g both map $\mathbb{R}^n$ to $\mathbb{R}$. Assume that f and g are both *concave* functions. In this question, we explore conditions on g which ensure the following property

$$\text{if } \mathbf{x} \text{ satisfies the Kuhn Tucker conditions, then } \mathbf{x} \text{ solves the max problem.} \tag{S}$$

i) Show graphically that if g is everywhere *less concave* than f, then property (S) holds. (Hint: concave functions are quasi-concave).

When g has less curvature than f, as in Fig. 2, then when the KT conditions are satisfied, the strict upper contour set of f corresponding $f(\mathbf{x})$ has an empty intersection with the *lower* contour set of g corresponding to $g(\mathbf{x})$.

ii) Use one of Taylor's theorems to show that the following mathematical condition does *not* capture the notion of "less concave." Specifically, demonstrate that the condition (B) below *does not* imply that condition (S) is satisfied:

$$\text{for all } \mathbf{x} \text{ and all } \mathbf{dx}, \mathbf{dx}'\mathrm{Hf}(\mathbf{x})\mathbf{dx} < \mathbf{dx}'\mathrm{Hg}(\mathbf{x})\mathbf{dx} < 0 \tag{B}$$

(Hint: let $\nabla f(\mathbf{x}) = \begin{bmatrix} 1 & 1 \end{bmatrix}$, $\mathrm{Hf}(\mathbf{x}) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$, $\mathrm{Hg}(\mathbf{x}) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. Now all you have to do is find $\nabla g(\mathbf{x})$ and pick a vector $\mathbf{dx}$ such that one of the second order Taylor expansions is positive, and the other is negative. Of course, you can't just stop there: you have to explain why these properties answer the question.)

As per the hint suppose that for some functions f and g and some vector $\mathbf{x} \in \mathbb{R}^2$, we have $\nabla f(\mathbf{x}) = \begin{bmatrix} 1 & 1 \end{bmatrix}$, $\nabla g(\mathbf{x}) = \begin{bmatrix} 0.1 & 0.1 \end{bmatrix}$, $\mathrm{Hf}(\mathbf{x}) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$, $\mathrm{Hg}(\mathbf{x}) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. Indeed, we will suppose that the two Hessians are *independent* of $\mathbf{x}$. Clearly, the KT conditions are satisfied at $\mathbf{x}$ since $\nabla f(\mathbf{x}) = 10\nabla g(\mathbf{x})$. However, $\mathbf{x}$ does not solve the constrained optimization problem, since for $\mathrm{dx} = (0.1,\ -0.075)$ and some

$\lambda \in [0,1]$.

$$f(\mathbf{x}+\mathsf{dx}) - f(\mathbf{x}) \quad = \quad \nabla f(\mathbf{x}) \cdot \mathsf{dx} \quad + \quad \mathsf{dx}'\mathsf{Hf}(\mathbf{x}+\lambda\mathsf{dx})\mathsf{dx}$$

which, since the hessian is independent of $\mathbf{x}$

$$\begin{aligned} &= \nabla f(\mathbf{x}) \cdot \mathsf{dx} \quad + \quad \mathsf{dx}'\mathsf{Hf}(\mathbf{x})\mathsf{dx} \\ &= 0.0094 \end{aligned}$$

$$g(\mathbf{x}+\mathsf{dx}) - g(\mathbf{x}) \quad = \quad \nabla g(\mathbf{x}) \cdot \mathsf{dx} \quad + \quad \mathsf{dx}'\mathsf{Hg}(\mathbf{x}+\lambda\mathsf{dx})\mathsf{dx}$$

which, since the hessian is independent of $\mathbf{x}$

$$\begin{aligned} &= \nabla g(\mathbf{x}) \cdot \mathsf{dx} \quad + \quad \mathsf{dx}'\mathsf{Hg}(\mathbf{x})\mathsf{dx} \\ &= -0.0053 \end{aligned}$$

Hence, we've established that by adding dx to $\mathbf{x}$, you increase the objective while decreasing the constraint. Hence $\mathbf{x}$ is not a solution to the constrained maximization problem. Note that by making the Hessian independent of $\mathbf{x}$, we could use the *exact* version of Taylor's theorem and not have to worry about how small is small.

iii) Modify condition (B) so that it *does* capture the notion of "less concave". (Hint: your condition should exhibit the property that whenever one of the second order Taylor expansions is positive then the other one is also. Of course, you can't just stop there: you have to explain why these properties answer the question.)

The following (strong) condition is an easy solution that allows us to switch from Taylor-Young to Taylor-Lagrange. Let $\lambda(\mathbf{x}) = \frac{||\nabla f(\mathbf{x})||}{||\nabla g(\mathbf{x})||}$. Now define condition (B') as follows:

$$\begin{aligned} &\mathsf{Hf}(\cdot) \text{ and } \mathsf{Hg}(\cdot) \text{ are independent of } \mathbf{x}; \\ &\text{for all } \mathbf{x} \text{ and all } \mathbf{dx}, \mathbf{dx}'\mathsf{Hf}\mathbf{dx} < \lambda(\mathbf{x})\mathbf{dx}'(\mathbf{x})\mathsf{Hg}\mathbf{dx} < 0 \end{aligned} \tag{B'}$$

Now suppose that $\mathbf{x}$ satisfies the KT conditions. In this case, $\nabla f(\mathbf{x}) = \lambda(\mathbf{x})\nabla g(\mathbf{x})$. Moreover, since Hf and Hg are independent of $\mathbf{x}$,

$$\begin{aligned} f(\mathbf{x}+\mathsf{dx}) - f(\mathbf{x}) &= \nabla f(\mathbf{x}) \cdot \mathsf{dx} \quad + \quad \mathsf{dx}'\mathsf{Hfdx} \\ &= \lambda(\mathbf{x})\nabla g(\mathbf{x}) \cdot \mathsf{dx} \quad + \quad \mathsf{dx}'\mathsf{Hfdx} \\ &< \lambda(\mathbf{x})\nabla g(\mathbf{x}) \cdot \mathsf{dx} \quad + \quad \lambda(\mathbf{x})\mathsf{dx}'\mathsf{Hgdx} \\ &= \lambda(\mathbf{x})\left(\nabla g(\mathbf{x}) \cdot \mathsf{dx} \quad + \quad \mathsf{dx}'\mathsf{Hgdx}\right) \\ &= \lambda(\mathbf{x})\left(g(\mathbf{x}+\mathsf{dx}) - g(\mathbf{x})\right) \end{aligned}$$

It now follows immediately that $f(\mathbf{x}+\mathsf{dx}) - f(\mathbf{x}) > 0$ implies $g(\mathbf{x}+\mathsf{dx}) - g(\mathbf{x}) > 0$, proving that $\mathbf{x}$ maximizes f on the constraint set.