ARE211, Fall2015

NPP2: THU, SEP 17, 2015 PRINTED: AUGUST 25, 2015 (LEC# 8)

Contents

3. Nonlinear Programming Problems and the Kuhn Tucker conditions (cont)

Key points: Sufficient conditions for a solution to the NPP

- Quasi-concavity and semi-definiteness on a subspace:

  $\mathbf{dx}'Hf(\mathbf{x})\mathbf{dx} \leq 0$, for all $\mathbf{dx}$ s.t. $\bigtriangledown f(\mathbf{x})\mathbf{dx} = 0$.
- The principal minor representation of strict quasi-concavity:

  $\forall \mathbf{x}$, and all $k = 1, ..., n$, the sign of the k'th leading principal minor of the bordered matrix $\begin{bmatrix} 0 & \bigtriangledown f(\mathbf{x})' \\ \bigtriangledown f(\mathbf{x}) & \mathrm{Hf}(\mathbf{x}) \end{bmatrix}$. must have the same sign as $(-1)^k$, where the k'th leading principal minor of this matrix is the det of the top-left $(k+1) \times (k+1)$ submatrix.
- Understanding the problem of the vanishing gradient
- Defn of pseudo-concavity: f is pseudo-concave if

  $$\forall \mathbf{x}, \mathbf{x}' \in X, \text{ if } f(\mathbf{x}') > f(\mathbf{x}) \text{ then } \bigtriangledown f(\mathbf{x}) \cdot (\mathbf{x}' - \mathbf{x}) > 0.$$

- Pseudo-concavity and its relationship to quasi-concavity:

  if f is $\mathbb{C}^2$ then f is pseudo-concave iff f is quasi-concave and if $\bigtriangledown f(\cdot) = 0$ at $\mathbf{x}$ implies $f(\cdot)$ attains a global max at $\mathbf{x}$.

- Sufficient conditions for a solution to the NPP:

  If f is pseudo-concave and the g^j's are quasi-convex, then a sufficient condition for a solution to the NPP at $\bar{\mathbf{x}} \in \mathbb{R}^m_+$ is that there exists a vector $\bar{\lambda} \in \mathbb{R}^m_+$ such that

  $$\bigtriangledown f(\bar{\mathbf{x}})^T = \lambda^T Jg(\bar{\mathbf{x}})$$

  and $\bar{\lambda}$ has the property that $\bar{\lambda}_j = 0$, for each j such that $g^j(\bar{\mathbf{x}}) < b_j$.

- understanding the role of second order conditions in the sufficiency argument

3.2. **Sufficient conditions for a solution to an NPP: Preliminaries**

So far we've only established necessary conditions for a solution to the NPP. Not surprisingly, without further restrictions, the KTT conditions aren't sufficient for a solution. They may be satisfied at a *minimum* on the constraint set, or else at a local but not global max. In this lecture we focus on identifying restrictions we can impose on the objective and constraint functions which ensure that the KKT conditions will be both necessary *and* sufficient for a solution. A good place to start, in our search for restrictions is to assume that objective function f is strictly quasi-concave while the constraint functions are quasi-convex. (Since the lower contour sets of quasi-convex functions are convex, and the intersection of convex sets is convex, and the constraint set is an intersection of lower contour sets, the condition that the constraint functions are quasi-convex implies that the constraint *set* is a convex set.) This isn't quite good enough, as we will see.

Figure Fig. 1 illustrates why quasi-concavity of the objective and convexity of the constraint set are required as restrictions. It also illustrates why they are not enough to ensure that the a solution to the KKT is sufficient for a solution to the NPP.

- the left panel has a constraint set that isn't generated by quasi-convex functions; level set represents a strictly quasi-concave objective function; the KKT conditions are satisfied at a

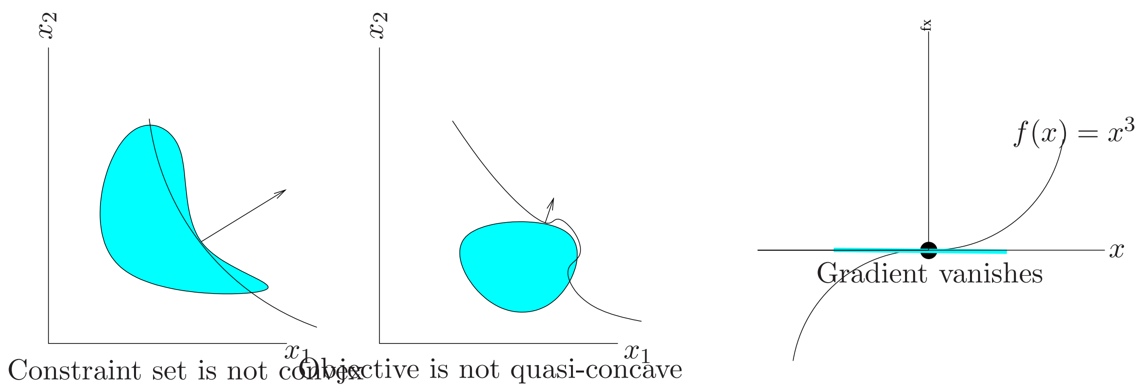

FIGURE 1. Three examples where KKT conditions are not sufficient for a soln

local tangency, but it's a local *minimum* on the north east boundary of the constraint set. there are points further away that give higher values for the objective function.

- the middle panel has a convex constraint set, but the objective isn't quasi-concave. In this case we have a local max on the constraint set that isn't a solution to the problem. By a solution we mean a *global* max on the constraint set.
- the right panel is a more subtle problem. The constraint set, $[-1, 1]$ is convex and compact. It's given by $g^1(x) = x \leq 1$, $g^2(x) = -x \leq -1$. The objective function $f(x) = x^3$ is strictly quasi-concave, but at the origin (represented by the big dot, the KKT is satisfied: i.e., $f'(x) = 0 = [0, 0] \cdot [1, -1] = 0$. But zero is clearly not a solution to the NPP. The problem in this example is referred to as the "vanishing gradient" problem, because the gradient vanishes at an x value that is not a global maximum.

These examples make clear that we cannot say that if the objective and constraint functions have the right "quasi" properties, then satisfying the KKT conditions is sufficient for a max. We will have to strengthen quasi-concavity just enough so that it excludes function such as $f(x) = x^3$.

3.2.1. *First preliminary: the problem of the vanishing gradient.* To avoid the problem in the right panel of Fig. 1, we could simply assume that f has a non-vanishing gradient. But this restriction throws the baby out with the bath-water: e.g., the problem $\max \mathbf{x}(1 - \mathbf{x})$ s.t. $\mathbf{x} \in [0, 1]$ has a global max at 0.5, at which point the gradient vanishes. More generally, the non-vanishing gradient assumption excludes *any* differentiable function that attains a global maximum.

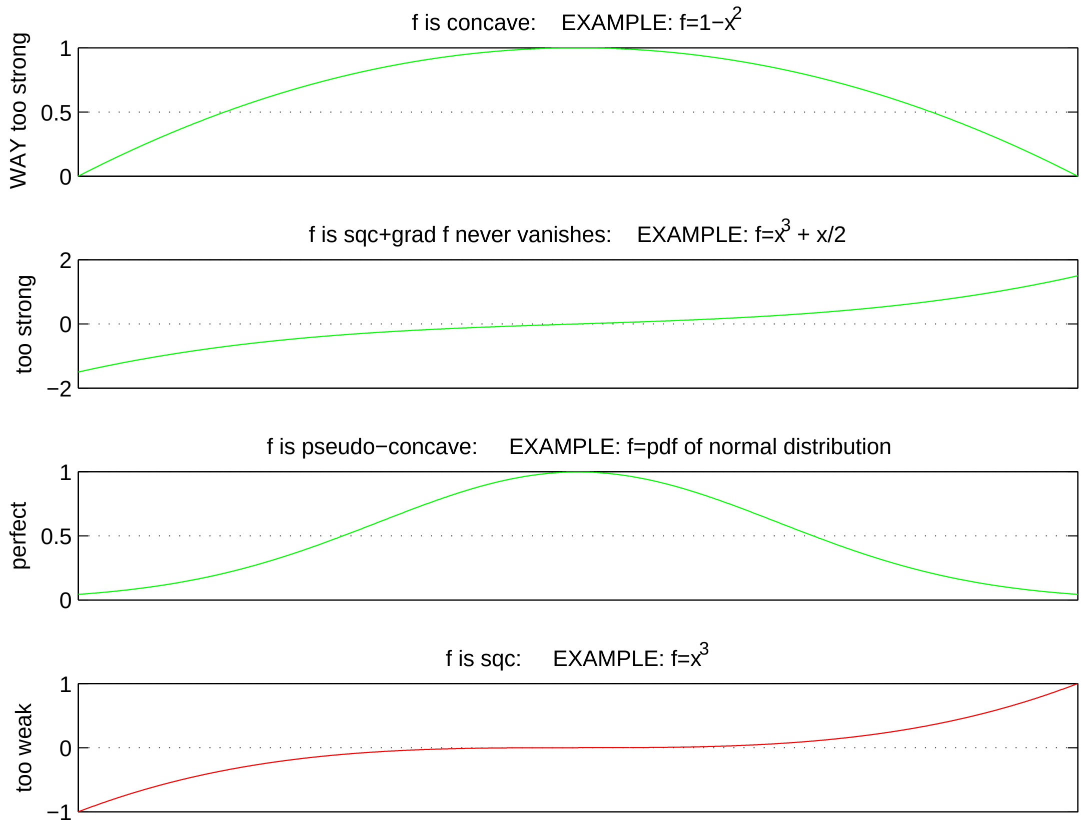

FIGURE 2. Weakest condition such that solution to KKT $\Longrightarrow$ soln to NNP

So we need a condition that implies quasi-concavity and also has the property that the gradient vanishes at $\mathbf{x}$ *only* if $\mathbf{x}$ is a global maximizer of the function. The following condition on f—called *pseudoconcavity* in S&B (the original name) and M.K.9 in MWG—does just precisely this

$$\forall \mathbf{x}, \mathbf{x}' \in X, \text{ if } f(\mathbf{x}') > f(\mathbf{x}) \text{ then } \bigtriangledown f(\mathbf{x}) \cdot (\mathbf{x}' - \mathbf{x}) > 0. \tag{1}$$

Fig. 2 summarizes the issue: going from the top panel to the bottom, the assumptions on f get progressively weaker. All but the bottom panel, give us what we want, i.e., a condition that is sufficient to ensure that a solution to the KKT $\Longrightarrow$ a soln to NNP.

Note that (1) says a couple of things. First, it says that a *necessary* condition for $f(\mathbf{x}') > f(\mathbf{x})$ is that $dx = (\mathbf{x}' - \mathbf{x})$ makes an acute angle with the gradient of f. (This looks very much like quasi-concavity). Second, it implies that

$$\text{if } \bigtriangledown f(\cdot) = 0 \text{ at } \mathbf{x} \text{ then } f(\cdot) \text{ attains a global max at } \mathbf{x} \tag{2}$$

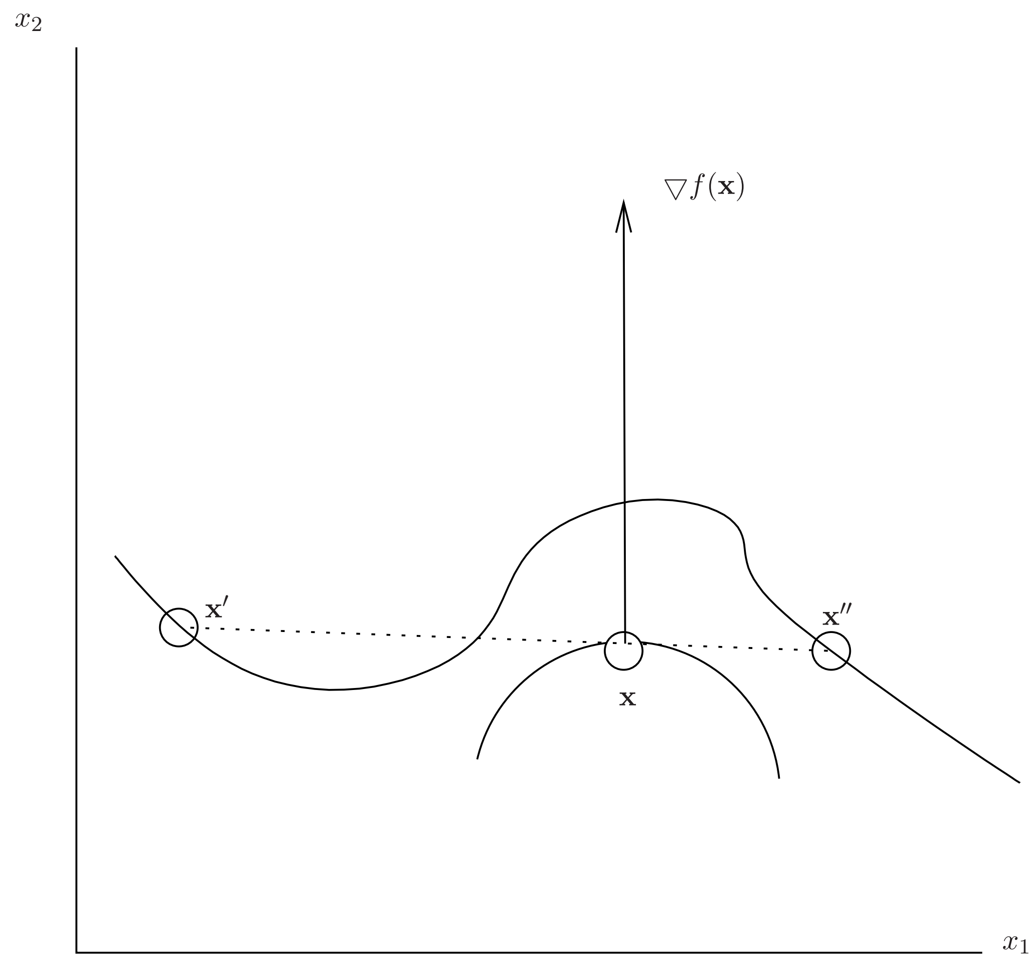

FIGURE 3. ¬-pseudo-concavity implies ¬-quasi-concavity

since if not then there would necessarily exist $\mathbf{x}, \mathbf{x}' \in X$ s.t. $f(\mathbf{x}') > f(\mathbf{x})$, and $\triangledown f(\mathbf{x}) \cdot (\mathbf{x}' - \mathbf{x}) = \mathbf{0} \cdot (\mathbf{x}' - \mathbf{x}) = 0$, violating (1).

Our next result establishes precisely the relationship between pseudo-concavity and quasi-concavity:

$$\text{if } f \text{ is } \mathbb{C}^2 \text{ then } f \text{ is pseudo-concave iff } f \text{ is quasi-concave and satisfies (2)} \tag{3}$$

To prove the $\Longrightarrow$ direction of (3), we'll show that (2) together with ($\neg$ quasi-concavity) implies ($\neg$ pseudo-concavity).

Suppose that f is not quasi-concave, i.e., there is an upper contour set that is not convex, i.e., $\exists \mathbf{x}', \mathbf{x}'', \mathbf{x} \in X$ such that $\mathbf{x} = \lambda \mathbf{x}' + (1-\lambda)\mathbf{x}''$, for some $\lambda \in (0,1)$ and $f(\mathbf{x}'') \geq f(\mathbf{x}') > f(\mathbf{x})$. ($\mathbf{x}$ exists by Weierstrass.) Assume w.l.o.g. that $f(\cdot)$ is minimized on $[\mathbf{x}', \mathbf{x}'']$ at $\mathbf{x}$. In this case, by the

KKT necessary conditions, either $\bigtriangledown f(\mathbf{x}) = 0$ or $\bigtriangledown f(\mathbf{x})$ is perpendicular at $\mathbf{x}$ to the level set of f corresponding to $f(\mathbf{x})$. In either case, $\bigtriangledown f(\mathbf{x}) \cdot (\mathbf{x}' - \mathbf{x}) = 0$. Since $\mathbf{x}' > \mathbf{x}$ this contradicts (1).

To prove the $\Leftarrow$ direction of (3), we'll show that (2) together with ($\neg$ pseudo-concavity) implies ($\neg$ quasi-concavity). Assume that there exists $\mathbf{x}, \mathbf{x}'$ such that $f(\mathbf{x}') > f(\mathbf{x})$ but $\bigtriangledown f(\mathbf{x}) \cdot (\mathbf{x}' - \mathbf{x}) \leq 0$.. Since $\mathbf{x}$ is not a global maximizer of f, (2) implies that $\bigtriangledown f(\mathbf{x}) \neq 0$. By continuity, we can pick $\epsilon > 0$ sufficiently small that for $\mathbf{y} = \mathbf{x}' - \epsilon \bigtriangledown f(\mathbf{x})$, $f(\mathbf{y}) > f(\mathbf{x})$. We'll show that a portion of the line segment joining $\mathbf{y}$ and $\mathbf{x}$ does not belong to the upper contour set of f corresponding to $f(\mathbf{x})$, proving that f is not quasi-concave. We have

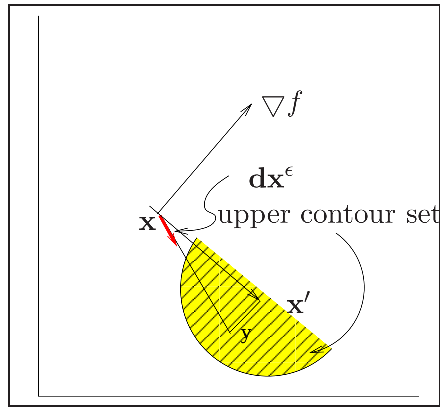

FIGURE 4. pseudo-concavity rules this out

$$\begin{aligned}
\bigtriangledown f(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) &= \bigtriangledown f(\mathbf{x}) \cdot ((\mathbf{x}' - \epsilon \bigtriangledown f(\mathbf{x})) - \mathbf{x}) \\
&= \underbrace{\bigtriangledown f(\mathbf{x}) \cdot (\mathbf{x}' - \mathbf{x})}_{\leq 0 \text{ by assumption}} - \epsilon \underbrace{|| \bigtriangledown f(\mathbf{x})||^2}_{>0} \quad < \quad 0
\end{aligned}$$

Let $\mathbf{dx}^\epsilon = \epsilon(\mathbf{y} - \mathbf{x})$. For all ϵ, $\bigtriangledown f(\mathbf{x}) \cdot \mathbf{dx}^\epsilon = \epsilon \bigtriangledown f(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) < 0$. Now, by Taylor-Young's theorem, if $\mathbf{dx} \neq 0$ is sufficiently small, then the sign of $(f(\mathbf{x} + \mathbf{dx}^\epsilon) - f(\mathbf{x}))$ is the same as the sign of $\bigtriangledown f(\mathbf{x}) \cdot \mathbf{dx}^\epsilon$, i.e., $f(\mathbf{x} + \mathbf{dx}^\epsilon) < f(\mathbf{x})$. We have now established that a portion of the line segment joining $\mathbf{x}$ and $\mathbf{y}$ does not belong to the upper contour set of f corresponding to $f(\mathbf{x})$. □

Conclude that pseudo-concavity is a much weaker assumption than the non-vanishing gradient condition, and will give us just enough to ensure that the KT conditions are not only necessary but sufficient as well. In particular, pseudo-concavity admits the possibility that our solution to the NPP may be unconstrained, whereas quasi-concavity plus "the gradient never vanishes" excludes this possibility.

3.2.2. *Second preliminary: quasi-concavity and the Hessian of f.* Recall from earlier that a condition sufficient to ensure that f is strictly (weakly) concave is to require that the Hessian of f be everywhere negative (semi) definite. Analogously, as we've seen in the Calculus section, a sufficient condition for f is strict (weak) quasi-concavity is a weaker "definiteness subject to constraint" property for f. The following result, which is a tiny bit stronger than the one we proved earlier, gives a sufficient condition for strict quasi concavity. Since we've already proved the earlier theorem, we won't bother to prove this variant.

Theorem (SQC): A sufficient condition for $f : \mathbb{R}^n \to \mathbb{R}$ to be strictly quasi-concave is that for all $\mathbf{x}$ and all $\mathbf{dx}$ such that $\bigtriangledown f(\mathbf{x})'\mathbf{dx} = 0$, $\mathbf{dx}'\mathrm{Hf}(\mathbf{x})\mathbf{dx} < 0$.

The following condition on the leading principal minors of the *bordered* hessian of f is equivalent to this "definiteness subject to constraint" property. *for all* $\mathbf{x}$, and all $k = 1, ..., n$, the sign of the k'th leading principal minor of the following bordered matrix must have the same sign as $(-1)^k$: $\begin{bmatrix} 0 & \bigtriangledown f(\mathbf{x})' \\ \bigtriangledown f(\mathbf{x}) & \mathrm{Hf}(\mathbf{x}) \end{bmatrix}$. where the k'th leading principal minor of this matrix is the determinant of the top-left $(k+1) \times (k+1)$ submatrix. We emphasize yet again that strict quasi-concavity is a *global* property, so that this leading principal minor property has to hold *for all* $\mathbf{x}$ in the domain of the function in order to guarantee strict quasi-concavity.

Note also that the above condition isn't *necessary* for strict quasi-concavity: the usual example, $f(x) = x^3$, establishes this: f is strictly quasi-concave, but at $\bar{\mathbf{x}} = 0$, and all $\mathbf{dx}$, $\bigtriangledown f(\bar{\mathbf{x}})\mathbf{dx} = 0$, while $\mathbf{dx}'\mathrm{Hf}(\bar{\mathbf{x}})\mathbf{dx} = 0$.

3.2.3. *Third preliminary: "Definiteness subject to constraint" and sufficiency.* A sufficient condition for strict concavity is that for all $\mathbf{x}$, $\mathbf{dx}'\mathrm{Hf}(\bar{\mathbf{x}})\mathbf{dx} < 0$, and all $\mathbf{dx} \neq 0$. For strict quasi-concavity, we only require this property of the Hessian holds for vectors that are *orthogonal to* $\bigtriangledown(\mathbf{x})$. Similarly, for g to be strictly quasi-convex, we only require that $\mathbf{dx}'\mathrm{Hg}(\bar{\mathbf{x}})\mathbf{dx} > 0$, and all $\mathbf{dx} \neq 0$ such that $\bigtriangledown g(\mathbf{x}) = 0$. These conditions are much weaker, infinitely weaker in fact, than the conditions for concavity and convexity. However, they are not quite as weak as they look: the condition on orthogonal vectors also has implications for $\mathbf{dx} \neq 0$'s that are *almost* orthogonal to $\bigtriangledown(\mathbf{x})$, and we need these implications in order to prove that the KT conditions are sufficient for a solution. Specifically, if $\bigtriangledown f(\bar{\mathbf{x}})\mathbf{dx} = 0$ implies $\mathbf{dx}'\mathrm{Hf}(\bar{\mathbf{x}})\mathbf{dx} < 0$, then by continuity, there exists $\epsilon > 0$ such that for any $\mathbf{dx} \neq 0$,

$$\text{if } |\bigtriangledown f(\bar{\mathbf{x}})\mathbf{dx}| < \epsilon, \text{ then } \mathbf{dx}'\mathrm{Hf}(\bar{\mathbf{x}})\mathbf{dx} < 0 \tag{4a}$$

similarly, if $\bigtriangledown g(\bar{\mathbf{x}})\mathbf{dx} = 0$ implies $\mathbf{dx}'\mathrm{Hg}(\bar{\mathbf{x}})\mathbf{dx} > 0$, then

$$\text{if } |\bigtriangledown g(\bar{\mathbf{x}})\mathbf{dx}| < \epsilon, \text{ then } \mathbf{dx}'\mathrm{Hg}(\bar{\mathbf{x}})\mathbf{dx} > 0 \tag{4b}$$

Why is (4) so important?

a) to show that the KKT conditions give us a solution to the NPP, we need to show that

$$\text{for any } \mathbf{dx}, \qquad \text{if } f(\mathbf{x}+\mathbf{dx}) - f(\mathbf{x}) > 0 \qquad \text{then} \qquad g(\mathbf{x}+\mathbf{dx}) - g(\mathbf{x}) > 0 \tag{5}$$

b) We know that $\forall \epsilon > 0$, $\exists \delta > 0$ s.t. if $||\mathbf{dx}|| < \delta$, then the first order Taylor approximations to both $(f(\mathbf{x}+\mathbf{dx}) - f(\mathbf{x}))$ and $(g(\mathbf{x}+\mathbf{dx}) - g(\mathbf{x}))$ have the same signs as the true differences, except (possibly) if $|\cos\theta^{\mathbf{dx}}| < \epsilon$, where $\theta^{\mathbf{dx}}$ is angle between $\mathbf{dx}$ and $\bigtriangledown f$. Since $\bigtriangledown f$ and $\bigtriangledown g$ are colinear, it follows that θ is *also* the angle between $\mathbf{dx}$ and $\bigtriangledown g$.

c) It follows from point b) that condition (5) is established for all $\mathbf{dx}$ such that $||\mathbf{dx}|| < \delta$ and $\theta^{\mathbf{dx}} \geq \epsilon$. Specifically, point b) establishes that $(f(\mathbf{x}+\mathbf{dx}) - f(\mathbf{x})) > 0$ implies $(g(\mathbf{x}+\mathbf{dx}) - g(\mathbf{x})) > 0$.

d) Now pick $\epsilon > 0$ such that conditions (4) hold and consider $\mathbf{dx}$ such that $||\mathbf{dx}|| < \delta$ and $\theta^{\mathbf{dx}} < \epsilon$. There are now two cases to consider

   i) $\bigtriangledown f(\mathbf{x})\mathbf{dx} \geq 0$: Necessarily $\bigtriangledown g(\mathbf{x})\mathbf{dx} \geq 0$. But from (4b), we know that the remainder term is positive also. Conclude that $(g(\mathbf{x}+\mathbf{dx}) - g(\mathbf{x})) > 0$ establishing (5) for this $\mathbf{dx}$.

   ii) $\bigtriangledown f(\mathbf{x})\mathbf{dx} < 0$: From (4a), we know that the remainder term is negative also. Conclude that $(f(\mathbf{x}+\mathbf{dx}) - f(\mathbf{x})) < 0$ so that for this $\mathbf{dx}$, condition (5) does not apply.

Here's the above argument in more detail. Now as we've discussed over and over again, you can't find this ball just by using first order conditions. You need your *second order conditions* to be cooperative in the region where the first order conditions fail you. If they are sufficiently uncooperative, i.e., if the signs of the quadratic terms in (4) are all reversed, then for any given ϵ-ball, there are going to be $\mathbf{dx}$'s that

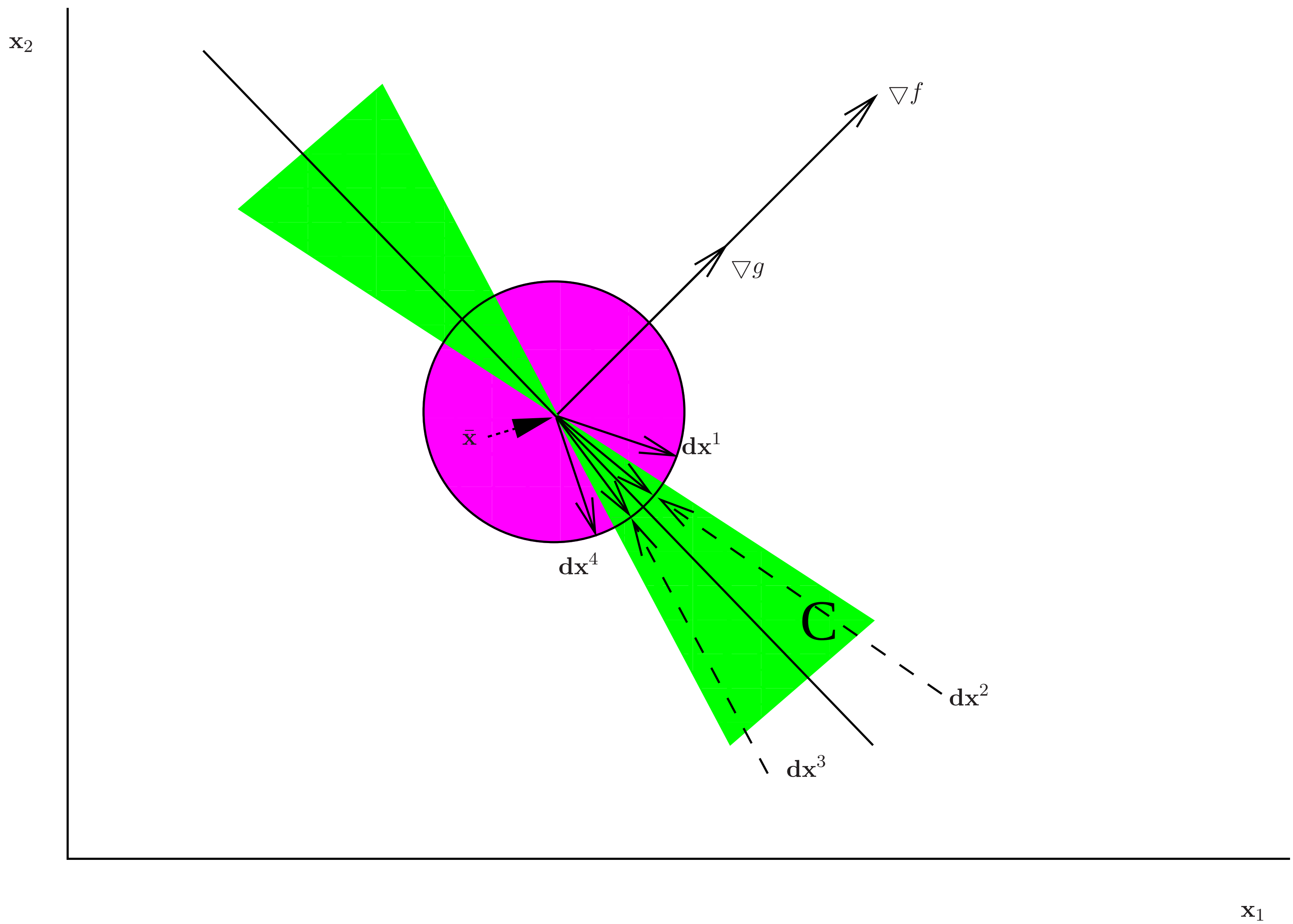

Figure 5. Second order conditions and sufficiency

- make an angle close to 90° with both $\bigtriangledown f(\mathbf{x})$ and $\bigtriangledown g(\mathbf{x})$, resulting in almost zero inner products $\bigtriangledown f(\mathbf{x})\mathbf{dx}$ and $\bigtriangledown g(\mathbf{x})\mathbf{dx}$, which are dominated by
- a positive second order term for f, resulting in a net increase in f, and
- a negative second order term for g, resulting in a net *decrease* in g, so you remain in the constraint set.
- in which case you don't have a maximum on the constraint set.

On the other hand, suppose that the point $\mathbf{x}$ in Fig. 5 satisfies the KKT conditions for the canonincal NPP with one constraint, i.e., $\max f(\mathbf{x})$ s.t. $g(\mathbf{x}) \leq b$. We will make the following assumptions:

(1) the KKT conditions are satisfied;
(2) the constraint is satisfied with equality;
(3) the two parts of (4) are satisfied for any vector $\mathbf{dx}$ in the double cone labeled C.
(4) for any vector $\mathbf{dx}$ in the shaded circle but *not* in the set C, the first order terms in the Taylor expansions of f and g dominate, i.e., the signs of the first order terms $\bigtriangledown f(\mathbf{x})\mathbf{dx}$ and $\bigtriangledown g(\mathbf{x})\mathbf{dx}$ agree with, respectively, $f(\mathbf{x}+\mathbf{dx}) - f(\mathbf{x})$ and $g(\mathbf{x}+\mathbf{dx}) - g(\mathbf{x})$..

The following four part argument is a very informal sketch of the proof that $\mathbf{x}$ solves the constrained maximization problem.

(1) For a vector such as $\mathbf{dx}^1$ which makes an acute angle with $\bigtriangledown g(\mathbf{x})$, but does not belong to C, the positive first order term in the expansion of g dominates, so that $g(\mathbf{x}+\mathbf{dx}) > g(\mathbf{x})$, implying that $\mathbf{x}+\mathbf{dx}$ does not belong to the constraint set.
(2) For a vector such as $\mathbf{dx}^4$ which makes an obtuse angle with $\bigtriangledown f(\mathbf{x})$, but does not belong to C, the negative first order term in the expansion of f dominates, so that $f(\mathbf{x}+\mathbf{dx}) < f(\mathbf{x})$.
(3) For a vector such as $\mathbf{dx}^2 \in C$ which makes a near 90° acute angle with $\bigtriangledown g(\mathbf{x})$, the second order term $0.5\mathbf{dx}'\mathrm{Hg}(\bar{\mathbf{x}})\mathbf{dx} > 0$ reinforces rather than offsets the negligible first order term, ensuring that $\mathbf{x}+\mathbf{dx}$ does not belong to the constraint set.
(4) For a vector such as $\mathbf{dx}^3 \in C$ which makes a near 90° obtuse angle with $\bigtriangledown f(\mathbf{x})$, the second order term $0.5\mathbf{dx}'\mathrm{Hf}(\bar{\mathbf{x}})\mathbf{dx} < 0$ reinforces rather than offsets the negligible first order term, ensuring that $f(\mathbf{x}+\mathbf{dx}) < f(\mathbf{x})$.

How small is "sufficiently small" in (4)? The following example shows that the requirement for sufficiently small gets tougher and tougher, the less concave is f. Example: Consider the function $f(\mathbf{x},\mathbf{y}) = (\mathbf{xy})^\beta$, which is strictly quasi-concave but not concave for $\beta > 0.5$. We'll illustrate that regardless of the value of β, $\mathbf{dx}'\mathrm{Hf}^\lambda\mathbf{dx} < 0$, for any vector that is *almost* orthogonal to $\bigtriangledown f$, but that the criterion of "almost" gets tighter and tighter as β gets larger. That is, the higher is β (i.e., the less concave is f), the closer to orthogonal does $\mathbf{dx}$ have to be in order to ensure that

$\mathbf{dx}'\mathrm{Hf}^{\lambda}\mathbf{dx}$ is negative.

We have $\bigtriangledown f(\mathbf{x}_1,\mathbf{x}_2) = \beta\left(\mathbf{x}_1^{\beta-1}\mathbf{x}_2^{\beta}, \mathbf{x}_1^{\beta}\mathbf{x}_2^{\beta-1}\right)$ and $\mathrm{Hf}(\mathbf{x}_1,\mathbf{x}_2) = \beta\begin{bmatrix} (\beta-1)\mathbf{x}_1^{\beta-2}\mathbf{x}_2^{\beta} & \beta\mathbf{x}_1^{\beta-1}\mathbf{x}_2^{\beta-1} \\ \beta\mathbf{x}_1^{\beta-1}\mathbf{x}_2^{\beta-1} & (\beta-1)\mathbf{x}_1^{\beta}\mathbf{x}_2^{\beta-2} \end{bmatrix}$.

Evaluated at $(\mathbf{x}_1,\mathbf{x}_2) = (1,1)$, we have $\mathrm{Hf}(1,1) = \beta^2\begin{bmatrix} \frac{\beta-1}{\beta} & 1 \\ 1 & \frac{\beta-1}{\beta} \end{bmatrix} = \beta^2\begin{bmatrix} \lambda & 1 \\ 1 & \lambda \end{bmatrix}$, where $\lambda = \frac{\beta-1}{\beta}$.

Note that $\lambda \to 1$ as $\beta \to \infty$.

Now choose a unit length vector $\mathbf{dx}$ and consider

$$\mathbf{dx}'\mathrm{Hf}^{\lambda}\mathbf{dx} = \lambda(dx_1^2+dx_2^2)+2dx_1dx_2 = (dx_1+dx_2)^2-(1-\lambda)(dx_1^2+dx_2^2) = (dx_1+dx_2)^2-(1-\lambda)$$

For $\mathbf{dx}$ such that $\mathbf{x}_1 = -\mathbf{x}_2$, $\mathbf{dx}'\mathrm{Hf}^{\lambda}\mathbf{dx} < 0$, for all $\lambda < 1$, verifying that f is strictly quasi-concave.

However, the closer is λ to unity, the smaller is the set of unit vectors for which $\mathbf{dx}'\mathrm{Hf}^{\lambda}\mathbf{dx} < 0$.

3.3. Sufficient Conditions for a solution to the NPP: the theorem

The following theorem gives *sufficient* conditions for a solution (not necessarily unique) to the NPP.

Theorem (S): (*Sufficient* conditions for a solution to the NPP): If f is pseudo-concave and the g^j's are quasi-convex, then a sufficient condition for a solution to the NPP at $\bar{\mathbf{x}} \in \mathbb{R}^m_+$ is that there exists a vector $\bar{\lambda} \in \mathbb{R}^m_+$ such that

$$\bigtriangledown f(\bar{\mathbf{x}})^T = \lambda^T Jg(\bar{\mathbf{x}})$$

and $\bar{\lambda}$ has the property that $\bar{\lambda}_j = 0$, for each j such that $g^j(\bar{\mathbf{x}}) < b_j$.

Note that Theorem (S) doesn't guarantee that a solution exists. Need compactness for this. Note also that the sufficient conditions are like the necessary conditions, except that you don't need the CQ but do need pseudo-concavity and quasi-convexity. (MWG's version of Theorem (S)—Theorem M.K.3—is just like mine except that they *do* include the constraint qualification. This addition is unnecessary (they're not wrong, they just have a meaningless additional condition). The C.Q. says that you can have a maximum without the non-negative cone condition holding. If you assume as in (S) that the nonnegative cone property holds, then, obviously, you don't need to worry that perhaps it mightn't hold!

Proof of Theorem (S):

(1) suppose $\bar{\mathbf{x}}$ does not solve the NPP, i.e., for some $\mathbf{dx}$;
  - $f(\bar{\mathbf{x}} + \mathbf{dx}) > f(\bar{\mathbf{x}})$ & $g^j(\bar{\mathbf{x}} + \mathbf{dx}) \leq b_j$, $\forall j$.
  - since f is pseudo-concave, $\bigtriangledown f(\bar{\mathbf{x}})\mathbf{dx} > 0$.
  - we'll show $\mathbf{x}$ fails the KKT:
    - i.e., consider *any* $\boldsymbol{\lambda} \in \mathbb{R}^m_+$ s.t. $g^j(\bar{\mathbf{x}}) < b^j \Longrightarrow \lambda^j = 0$
    - we'll show $\bigtriangledown f(\bar{\mathbf{x}}) \neq \boldsymbol{\lambda}^T Jg(\bar{\mathbf{x}})$
    - that is,
      - ◇ the KKT necessary conditions state that in order for $\bar{\mathbf{x}}$ to solve the NPP, *there must exist* a $\boldsymbol{\lambda}$ vector satisfying a certain property (the complementary slackness condition), and also satisfies equation (6).
      - ◇ we'll show that this condition cannot be satisfied *for any* $\boldsymbol{\lambda}$ vector in the specified class.

(2) since for all j, $g^j(\bar{\mathbf{x}}) \leq b_j$ and $g^j(\bar{\mathbf{x}} + \delta\mathbf{dx}) \leq b_j$, it follows from the convexity of g^j's lower contour sets that $\forall \delta < 1$, $g^j(\bar{\mathbf{x}} + \delta\mathbf{dx}) \leq b_j$.

(3) there are now two cases to consider
  (a) suppose $g^j(\bar{\mathbf{x}}) = b_j$: in this case, $\forall \delta < 1$ $g^j(\bar{\mathbf{x}} + \delta\mathbf{dx}) - g^j(\mathbf{dx}) \leq 0$.
    - Local Taylor implies that for such a j, $\bigtriangledown g^j(\bar{\mathbf{x}})\mathbf{dx} \leq 0$,
      - suppose instead that for such a j, $\bigtriangledown g^j(\bar{\mathbf{x}})\mathbf{dx} > 0$;
        - ◇ then Local Taylor implies: for $\delta \approx 0$, $g^j(\bar{\mathbf{x}} + \delta\mathbf{dx}) - g^j(\mathbf{dx}) > 0$, a contradiction.

  (b) suppose $g^j(\bar{\mathbf{x}}) < b_j$: in this case,
    - it's not necessarily true that $\forall \delta < 1$ $g^j(\bar{\mathbf{x}} + \delta\mathbf{dx}) - g^j(\mathbf{dx}) \leq 0$.

      (we can't rule out the possibility that $g^j(\mathbf{dx}) \ll b^j$, while $g^j(\bar{\mathbf{x}} + \mathbf{dx}) = b^j$.)
    - so we *can't* conclude $\bigtriangledown g^j(\bar{\mathbf{x}})\mathbf{dx} \leq 0$.
    - *but* we have, by assumption, that $\lambda^j = 0$.

(4) Whichever case hold in part (3), we have that $\lambda^j \bigtriangledown g^j(\bar{\mathbf{x}}) \cdot \mathbf{dx} \leq 0$:
  - for (3a), it holds because $\lambda^j \geq 0$ and $\bigtriangledown g^j(\bar{\mathbf{x}}) \cdot \mathbf{dx} \leq 0$:
  - for (3b), it holds because $\lambda^j = 0$.

(5) That is, $(\boldsymbol{\lambda}^T Jg(\bar{\mathbf{x}})) \cdot \mathbf{dx} = \sum_{j=1}^m \lambda^j \bigtriangledown g^j(\bar{\mathbf{x}}) \cdot \mathbf{dx} \leq 0$

(6) Combining (1) and (5), and factoring out $\mathbf{dx}$, we have $\left(\bigtriangledown f(\bar{\mathbf{x}}) - \boldsymbol{\lambda}^T Jg(\bar{\mathbf{x}}))\right) \cdot \mathbf{dx} > 0$.

(7) Hence, the vector $\left(\bigtriangledown f(\bar{\mathbf{x}}) - \boldsymbol{\lambda}^T Jg(\bar{\mathbf{x}}))\right) \neq 0$.

(8) But since KKT requires equality of the expression in (7) to be zero, $\bar{\mathbf{x}}$ fails the KKT.

3.4. **Second Order conditions Without Quasi-Concavity**

This is a hard section that I'm not going to teach this year. Rather than include it in the notes, I've suppressed the entire section to save paper. If anybody is interested I'll print off the section.