

Solutions Homework Assignment 4, MATH 515, Spring 09

Problem 8) *Claim:* D convex, $x_1, \dots, x_n \in D$, $t_1, \dots, t_n > 0$ with $\sum_{j=1}^n t_j = 1$ then $\sum_{j=1}^n t_j x_j \in D$. *Proof:* Induction on n . The cases $n = 1, 2$ are clear by definition. For $n \geq 3$ let $t := t_1$ and $x := x_1 \in D$, and $y := \frac{1}{1-t_1}(t_2 x_2 + \dots + t_n x_n)$. Then $y \in D$ by induction hypothesis because $\frac{1}{1-t_1} \sum_{j=2}^n t_j = 1$. Thus an application with $n = 2$ to $D \ni tx + (1-t)y = t_1 x_1 + (1-t_1) \frac{1}{1-t_1} \sum_{j=2}^n t_j x_j = \sum_{i=1}^n t_i x_i$, which proves the claim. If $x = \sum_{t_i} x_i$ with $t_i > 0$ and $\sum t_i = 1$ and $y = \sum s_i x_i$ with $s_i > 0$ and $\sum s_i = 1$ then for $t > 0$ we get $tx + (1-t)y = t \sum t_i x_i + (1-t) \sum s_i x_i = \sum (tt_i + (1-t)s_i) x_i \in D$ because $\sum tt_i + (1-t)s_i = t \sum t_i + (1-t) \sum s_i = t + (1-t) = 1$, and for all i $tt_i + (1-t)s_i$ is a number between $t_i > 0$ and $s_i > 0$ and thus is > 0 .

Problem 9) (a) It suffices to show the claims for convexity. The concavity results follow because f' is increasing $\iff (-f)'$ is decreasing, and $f'' > 0 \iff (-f)'' = -f'' < 0$. *Claim.* $f : (a, b) \rightarrow \mathbb{R}$ differentiable, then f is convex $\iff f'$ is increasing. *Proof:* \implies : Let $x, y \in (a, b)$ and $z := tx + (1-t)y$. Then $f(z) \leq tf(x) + (1-t)f(y) \iff tf(z) + (1-t)f(z) \leq tf(x) + (1-t)f(y) \iff t(f(z) - f(x)) \leq (1-t)(f(y) - f(z))$. Now by the mean value theorem we know that $(f(z) - f(x)) = (z-x)f'(\xi)$ and $f(y) - f(z) = (y-z)f'(\eta)$ for $\xi \in (x, z)$ and $\eta \in (z, y)$ (see e. g. Rudin 5.10). Thus $f'(\xi) \leq f'(\eta)$ and $t(z-x) = (1-t)(y-z)$ ($z = tx + (1-t)y \implies z-x = (1-t)(y-x)$ and $y-z = t(y-x)$). show that f' increasing implies that f is convex. It also shows: $f'' > 0 \implies f'$ strictly increasing $\implies f$ strictly convex. \impliedby : f is convex implies with the notation above (take the quotient of $t(f(z) - f(x)) \leq (1-t)(f(y) - f(z))$ by $t(z-x) = (1-t)(y-z)$): $\frac{f(z)-f(x)}{z-x} \leq \frac{f(y)-f(z)}{y-z}$. If we let $z \rightarrow x$ we get $f'(x) \leq \frac{f(y)-f(x)}{y-x}$, and if we let $z \rightarrow y$ we get $\frac{f(y)-f(x)}{y-x} \leq f'(y)$, thus proving $f'(x) \leq f'(y)$ for $x < y$ in (a, b) .

(b) *Claim: Jensen's inequality:* For D a convex subset of a vector space E and $f : D \rightarrow \mathbb{R}$ a concave function show that

$$\sum_{i=1}^n t_i f(x_i) \leq f\left(\sum_{i=1}^n t_i x_i\right)$$

whenever $x_1, \dots, x_n \in D$, $t_1, \dots, t_n \in (0, 1)$ and $\sum_{i=1}^n t_i = 1$. Furthermore, if f is strictly concave then equality holds if and only if $x_1 = \dots = x_n$. *Proof:* This is very similar to Problem 8 and is done by induction. For $n = 1, 2$ is claim is true by definition. For $n \geq 3$ let $t'_i := \frac{t_i}{1-t_1}$ so that $\sum_{i=2}^n t'_i = 1$. Then the induction hypothesis for $n-1$ and 2 implies:

$\sum_{i=1}^n t_i f(x_i) = t_1 f(x_1) + (1-t_1) \sum_{i=2}^n t'_i f(x_i) \leq t_1 f(x_1) + (1-t_1) f(\sum_{i=2}^n t'_i x_i) \leq f(t_1 x_1 + (1-t_1) \sum_{i=2}^n t'_i x_i) = f(\sum_{i=1}^n t_i x_i)$. If f is strictly concave and $n \geq 3$ and not all x_i are equal then we may assume that not all of x_2, \dots, x_n are equal (otherwise use $n = 2$). Then $(1-t_1) \sum_{t'_i} f(x_i) < (1-t_1) f(\sum_{i=2}^n t'_i x_i)$ shows that the inequality is strict.

(c) Note that for $f(x) = \log(x)$ we have $f''(x) = -\frac{1}{x^2} < 0$ and thus f is strictly concave on $(0, \infty)$ by (a). Then by (b) we get with $p_i = t_i$ and $a_i = x_i$: $\sum_{i=1}^n p_i \log(a_i) = \sum_{i=1}^n \log(a_i^{p_i}) \leq \log(\sum_{i=1}^n p_i a_i)$. Now we can apply the strictly increasing exponential function on both sides to get the result:

$$e^{\sum_{i=1}^n \log(a_i^{p_i})} = \prod_{i=1}^n e^{\log(a_i^{p_i})} = \prod_{i=1}^n a_i^{p_i} \leq \sum_{i=1}^n p_i a_i = e^{\log(\sum_{i=1}^n p_i a_i)}.$$

Equality only if $a_1 = \dots = a_n$ because of (b) and the exponential function strictly increasing.

(d) *Claim:* For $a, b \geq 0$ and $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$ it follows $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ with equality only if $a^p = b^q$. *Proof:* A special case of (c) is $x_1^{p_1} x_2^{p_2} \leq p_1 x_1 + p_2 x_2$ with equality only if $x_1 = x_2$. Then put $x_1 = a^p$ and $x_2 = b^q$, $p_1 = 1/p$ and $p_2 = 1/q$ to get

$$ab = (a^p)^{1/p} (b^q)^{1/q} = x_1^{p_1} x_2^{p_2} \leq p_1 x_1 + p_2 x_2 = \frac{a^p}{p} + \frac{b^q}{q},$$

with equality if and only if $x_1 = x_2 \iff a^p = b^q$.

Problem 10 *Claim: Hölder's inequality:* Suppose $p, q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then for complex numbers $a_1, \dots, a_n, b_1, \dots, b_n$ we have

$$\left| \sum_{k=1}^n a_k b_k \right| \leq \left(\sum_{k=1}^n |a_k|^p \right)^{1/p} \left(\sum_{k=1}^n |b_k|^q \right)^{1/q}$$

with equality if and only if all a_k are 0 or $|b_k|^q = t|a_k|^p$ and $a_k b_k = e^{i\theta} |a_k b_k|$ for all k and some t and θ . *Proof:* If all $a_k = 0$ or all $b_k = 0$ then equality holds. Otherwise we can consider $a'_k := a_k / ((\sum |a_k|^p)^{1/p})$ and $b'_k := b_k / ((\sum |b_k|^q)^{1/q})$, which satisfy $\sum |a'_k|^p = \sum |b'_k|^q = 1$, and proving in this case that $|\sum_{k=1}^n a'_k b'_k| \leq 1$ proves the claim. But from (d) above for each k we get:

$$\left| \sum_{k=1}^n a'_k b'_k \right| \leq \sum_{k=1}^n |a'_k| |b'_k| \leq \sum_{k=1}^n \left(\frac{|a'_k|^p}{p} + \frac{|b'_k|^q}{q} \right) = \frac{1}{p} + \frac{1}{q} = 1$$

If the conditions hold on $a_1, \dots, a_n, b_1, \dots, b_n$ then obviously we have equality. If equality holds then $|a'_k|^p = |b'_k|^q$ for all k by (d) above. This implies $|a_k|^p = t|b_k|^q$ for some t and all k . Note that $t = 0$ covers the case of all $b_k = 0$, in which case equality also holds. For equality we also need $|\sum_{k=1}^n a_k b_k| = \sum_{k=1}^n |a_k| |b_k|$. Now the absolute value of a sum of complex numbers can only be equal to the sum of the absolute values if they all have the same polar angle. (Note that $|z + w| < |z| + |w|$ in the triangle inequality except if z is a *real* multiple w .) So the results follows from $\sum_{k=1}^n e^{i\theta} a_k b_k = e^{i\theta} \sum_{k=1}^n a_k b_k = |\sum_{k=1}^n a_k b_k| = \sum_{k=1}^n |a_k| |b_k|$ for some θ .