Integration and Fourier Theory

Lecture 11

Morten Grud Rasmussen

March 15, 2013

1 Convex functions and inequalities

Definition 1.1. A function $\varphi\colon (a,b) \to \mathbb{R}$, where $-\infty \le a < b \le \infty$, is called *convex* if

$$\varphi(\lambda x + (1-\lambda)y) \le \lambda\varphi(x) + (1-\lambda)\varphi(y) \tag{1}$$

holds for all $x, y \in (a,b)$, $0 \le \lambda \le 1$ $\big[$or, equivalently,

$$\frac{\varphi(t)-\varphi(x)}{t-x} \le \frac{\varphi(y)-\varphi(t)}{y-t}, \tag{2}$$

for any $a < x < t < y < b\big]$.

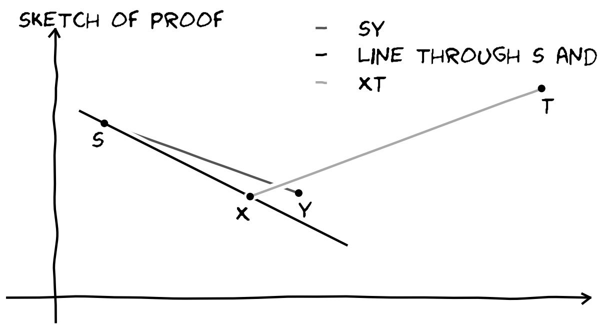
Remark 1.2. That (a,b) is open is important as Example 1.5 below will show.

Proposition 1.3. *Let $\varphi\colon (a,b) \to \mathbb{R}$ be differentiable. Then φ is convex if and only if $\varphi'(x) \le \varphi'(y)$ for all $a < x \le y < b$, i.e. the derivative is an increasing function.*

Proof. This follows easily from (2) and the mean value theorem. □

Theorem 1.4. *A convex function is continuous.*

Proof. For $a < s < x < y < t < b$ and $\psi\colon (a,b) \to \mathbb{R}$ convex, let S denote $(s, \varphi(s))$ and likewise for X, Y and T. The convexity of φ implies that X lies on or below SY so Y lies on or above the line through S and X. Likewise, Y lies on or below XT. This means that, as $y \downarrow x$, Y necessarily approaches X (y approaches x and Y has to lie between XT and the line through S and X), giving continuity from the right. Continuity from the left is handled analogously, and continuity follows. □



Example 1.5. The assumption that (a, b) is open is particularly important for the conclusion of the previous theorem. In fact, let φ be any bounded, convex function on (a, b). We can now define a class of functions φ_c on $[a, b]$ by setting $\varphi_c(a) = \varphi_c(b) = \sup_{x\in(a,b)} \varphi(x) + c$, and $\varphi_c(x) = \varphi(x)$ for $a < x < b$. It is now easy to see that the inequalities (1) and (2) are satisfied for φ_c if $c \geq 0$. However, φ_c may be discontinuous at a or b for $c = 0$ and is easily seen to necessarily be discontinuous at a and b if $c > 0$.

Theorem 1.6 (Jensen's Inequality). *Let μ be a positive measure on a σ-algebra $\mathfrak{M}$ with $\mu(\Omega) = 1$. Let $\varphi\colon (a, b) \to \mathbb{R}$ be convex and $f \in L^1(\mu)$ with $a < f(x) < b$ for all $x \in \Omega$. Then*

$$\varphi\Big(\int_\Omega f \,\mathrm{d}\mu\Big) \leq \int_\Omega \varphi \circ f \,\mathrm{d}\mu$$

Proof. Set $t = \int_\Omega f \,\mathrm{d}\mu$. Let β be the supremum of the left hand side of (2) for this choice of t and with the supremum taken over all values of x, $a < x < t$. Then we have

$$\frac{\varphi(t) - \varphi(x)}{t - x} \leq \beta \leq \frac{\varphi(y) - \varphi(t)}{y - t}, \qquad a < x < t < y < b,$$

or, more compactly written,

$$\varphi(s) \geq \varphi(t) + \beta(s - t), \qquad a < s < b.$$

Putting $f(x)$ in place of s and integrating both sides of the inequality with respect to x yields

$$\int_\Omega \varphi \circ f \,\mathrm{d}\mu \geq \varphi(t) + \beta\Big(\int_\Omega f \,\mathrm{d}\mu - t\Big),$$

which, when recalling the definition of t, is seen to be the wanted inequality. □

Jensen's inequality basically says that if you evaluate a convex function at the "average value" of another function, you get something which is smaller than the "average value" of the convex function composed with the other function, something which hopefully is relatively obvious from an intuitive point of view.

Example 1.7. The exponential function $\exp\colon (-\infty, \infty) \to \mathbb{R}$ is clearly convex as can be seen from the definition or by using Proposition 1.3. This means that for a real $f \in L^1(\mu)$ and μ as in Theorem 1.6, we have

$$\exp\Big(\int_\Omega f \,\mathrm{d}\mu\Big) \leq \int_\Omega e^f \,\mathrm{d}\mu\,. \tag{3}$$

Assume now that $\Omega = \{p_1, \ldots, p_n\}$ is a finite set and that μ gives each point p_i in Ω equal weight $\mu(p_i) = \frac{1}{n}$. Then (3) reduces to

$$\exp\Big(\frac{1}{n}\big(f(p_1) + \cdots + f(p_n)\big)\Big) \leq \frac{1}{n}\big(e^{f(p_1)} + \cdots + e^{f(p_n)}\big)\,. \tag{4}$$

Setting $x_i = e^{f(p_i)}$ and rewriting (4) yields the familiar inequality between the geometric and arithmetic average:

$$(x_1 \cdots x_n)^{\frac{1}{n}} \leq \frac{1}{n}(x_1 + \cdots + x_n)\,. \tag{5}$$

Analogously, the left and right-hand side of

$$\exp\left(\int_\Omega \log(g)\,\mathrm{d}\mu\right) \leq \int_\Omega g\,\mathrm{d}\mu$$

are often referred to as the geometric and arithmetic means, respectively, of a positive function g.

Example 1.8. A more general result than (5) can be obtained if the points p_i are not given equal weight, i.e. if $\mu(p_i) = \alpha_i > 0$ with $\sum_{i=1}^n \alpha_i = 1$ and α_i not necessarily equal to $\frac{1}{n}$. The same arguments as before now give

$$x_1^{\alpha_i} \cdots x_n^{\alpha_n} \leq \alpha_1 x_1 + \cdots + \alpha_n x_n\,. \tag{6}$$

Definition 1.9 (Hölder conjugates). If $p = 1$ and $q = \infty$ or if p and q are both real and positive and satisfy

$$\frac{1}{p} + \frac{1}{q} = 1\,, \tag{7}$$

which can also be written $p + q = pq$, then p and q are said to be *Hölder conjugates* or a pair of *conjugate exponents* or *dual exponents*.

We note that the pair 1 and ∞ also satisfies (7), appropriately interpreted, and that it can be seen as a limit case for a pair of real numbers.

Theorem 1.10 (The Hölder and Minkowski inequalities). *Let p and q be Hölder conjugates satisfying $1 < p, q < \infty$. Let f and g be be measurable functions on X with range in $[0, \infty]$. Then Hölder's inequality,*

$$\|fg\|_1 = \int_X fg\,\mathrm{d}\mu \leq \left(\int_X f^p\,\mathrm{d}\mu\right)^{\frac{1}{p}} \left(\int_X g^q\,\mathrm{d}\mu\right)^{\frac{1}{q}} = \|f\|_p \|g\|_q$$

and Minkowski's inequality

$$\|f+g\|_p = \left(\int_X (f+g)^p\,\mathrm{d}\mu\right)^{\frac{1}{p}} \leq \left(\int_X f^p\,\mathrm{d}\mu\right)^{\frac{1}{p}} + \left(\int_X g^p\,\mathrm{d}\mu\right)^{\frac{1}{p}} = \|f\|_p + \|g\|_p$$

holds true, where, for any complex, measurable function f on X and $0 < p < \infty$,

$$\|f\|_p = \left(\int_X |f|^p\,\mathrm{d}\mu\right)^{\frac{1}{p}}.$$

Remark 1.11. As indicated by the notation, $\|\cdot\|_p$ turns out to be a norm (when properly interpreted). However, as norms are, by definition, finite, $\|\cdot\|_p$ is usually (but not in the present context) restricted to the space of functions f for which $\|f\|_p$ is finite. We will return to this topic in the next lecture.

Proof of Hölder's inequality. If $\|f\|_p = 0$ then $f = 0$ a.e. and hence $\|fg\|_1 = 0$ and vice versa for $\|g\|_q$. If $\|fg\|_1 = 0$ or if $\|f\|_p\|g\|_g = \infty$, we're done before we even started, so assume now that $\|fg\|_1 > 0$ and $\|f\|_p\|g\|_q < \infty$. Let

$$F = \frac{f}{\|f\|_p} \qquad \text{and} \qquad G = \frac{g}{\|g\|_q}\,.$$

Then clearly

$$\int_X F^p \,\mathrm{d}\mu = \int_X G^q \,\mathrm{d}\mu = 1\,.$$

If $x \in X$ is such that $0 < F(x), G(x) < \infty$, then, since $\exp$ is onto $(0, \infty)$, we can find $s, t \in \mathbb{R}$ such that $F(x) = e^{s/p}$ and $G(x) = e^{t/q}$. Putting $n = 2$, $\alpha_1 = \frac{1}{p}$ and $\alpha_2 = \frac{1}{q}$ and writing $e^{s/p} = (e^s)^{1/p}$ and $e^{t/q} = (e^t)^{1/q}$ now gives something suspiciously close to what goes on in Example 1.8 (note that Ω and the p_i's are irrelevant for the conclusion), so (6) becomes

$$F(x)G(x) = (e^s)^{\frac{1}{p}}(e^t)^{\frac{1}{q}} \leq \frac{1}{p}e^s + \frac{1}{q}e^t = \frac{1}{p}F(x)^p + \frac{1}{q}G(x)^q\,,$$

which holds for all $x \in X$ such that $0 < F(x), G(x) < \infty$ and suitably chosen s and t. Moreover, the inequality between the leftmost and the rightmost quantities obviously remains true for $F(x)$ or $G(x)$ equal to 0 or ∞. Integrating this inequality yields

$$\frac{1}{\|f\|_p\|g\|_q}\int_X fg \,\mathrm{d}\mu = \int_X FG \,\mathrm{d}\mu \leq \frac{1}{p}\int_X F^p \,\mathrm{d}\mu + \frac{1}{q}\int_X G^q \,\mathrm{d}\mu = \frac{1}{p} + \frac{1}{q} = 1.$$

The desired inequality now follows from basic algebra. □

Proof of Minkowski's inequality. Minkowski's inequality can be seen as a corollary of Hölder's. As before, we begin by noting that the inequality is trivial if either $\|f+g\|_p = 0$ or $\|f\|_p + \|g\|_p = \infty$, so assume $\|f+g\|_p > 0$ and $\|f\|_p + \|g\|_p < \infty$. Write

$$(f+g)^p = (f+g)(f+g)^{p-1} = f\cdot(f+g)^{p-1} + g\cdot(f+g)^{p-1} \tag{8}$$

and apply Hölder's inequality to both terms on the right-hand side:

$$\int_X f\cdot(f+g)^{p-1}\,\mathrm{d}\mu \leq \|f\|_p\|(f+g)^{p-1}\|_q \qquad \text{and} \tag{9}$$

$$\int_X g\cdot(f+g)^{p-1}\,\mathrm{d}\mu \leq \|g\|_p\|(f+g)^{p-1}\|_q\,. \tag{10}$$

Now, since $(p-1)q = p$,

$$\|(f+g)^{p-1}\|_q = \left(\int_X \left((f+g)^{p-1}\right)^q \mathrm{d}\mu\right)^{\frac{1}{q}} = \left(\int_X (f+g)^p \,\mathrm{d}\mu\right)^{\frac{1}{q}} = \|f+g\|_p^{\frac{p}{q}} \tag{11}$$

and hence, putting (8), (9), (10) and (11) together, we get that

$$\|f+g\|_p^p \leq (\|f\|_p + \|g\|_p)\|f+g\|_p^{\frac{p}{q}}\,. \tag{12}$$

As $p - \frac{p}{q} = p(1-\frac{1}{q}) = p\cdot\frac{1}{p} = 1$, we're done if we can show that $\|f+g\|_p < \infty$ (because then we can divide the inequality (12) with $\|f+g\|_p^{p/q}$ — we have already assumed that $\|f+g\|_p > 0$). But this follows from the convexity of $(0,\infty) \ni t \mapsto t^p \in \mathbb{R}$, as

$$\left(\frac{f+g}{2}\right)^p \leq \frac{1}{2}(f^p + g^p)$$

shows that when $\|f\|_p + \|g\|_p$ is finite, so is $\|f+g\|_p$. □