KAPITEL 11

Ungleichungen

11.1. Jensen-Ungleichung

Definition 11.1.1. Eine Funktion $g : \mathbb{R} \to \mathbb{R}$ heißt *konvex*, wenn man für jedes $x_0 \in \mathbb{R}$ ein $K_0 = K_0(x_0) \in \mathbb{R}$ finden kann, so dass für alle $x \in \mathbb{R}$ gilt:

$$g(x) \geq g(x_0) + K_0(x - x_0).$$

Bemerkung 11.1.2. Eine Funktion $g(x)$ ist also genau dann konvex, wenn der Graph von g die folgende Eigenschaft besitzt: zu jedem x_0 können wir eine Gerade finden, die durch den Punkt $(x_0, f(x_0))$ geht und die Eigenschaft hat, dass der Graph von g komplett oberhalb dieser Geraden liegt. Die Zahl K_0 heißt das Subdifferential von g an der Stelle x_0 und ist nicht immer eindeutig bestimmt. Falls g differenzierbar ist, dann kann K_0 durch $K_0 = g'(x_0)$ festgelegt werden.

Satz 11.1.3 (Jensen-Ungleichung). *Sei g eine konvexe Funktion und X eine Zufallsvariable mit $X \in L^1$, $g(X) \in L^1$. Dann gilt:*

$$g(\mathbb{E}X) \leq \mathbb{E}[g(X)]. \tag{11.1.1}$$

Beispiel 11.1.4. Wir betrachten eine Zufallsvariable X, die n Werte $x_1, x_2, \ldots, x_n$ mit einer Wahrscheinlichkeit von jeweils $1/n$ annimmt. Der Erwartungswert ist dann das arithmetische Mittel der x_i, also

$$\mathbb{E}X = \frac{x_1 + \ldots + x_n}{n}.$$

Dann besagt die Jensen-Ungleichung für diesen Spezialfall, dass für jede konvexe Funktion g gilt

$$g\left(\frac{x_1 + \ldots + x_n}{n}\right) \leq \frac{g(x_1) + \ldots + g(x_n)}{n}.$$

Beweis von Satz 11.1.3. Wir setzen in die Definition der Konvexität $x_0 = \mathbb{E}X$, $x = X$ ein. Dann existiert laut Definition ein $K_0 \in \mathbb{R}$ mit der Eigenschaft, dass

$$g(X) \geq g(\mathbb{E}X) + K_0(X - \mathbb{E}X).$$

Nun bilden wir den Erwartungswert:

$$\mathbb{E}[g(X)] \geq \mathbb{E}[g(\mathbb{E}X)] + K_0\mathbb{E}[X - \mathbb{E}X] = g(\mathbb{E}X) + K_0\,(\mathbb{E}X - \mathbb{E}X) = g(\mathbb{E}X).$$

□

11.2. Ljapunow-Ungleichung

Ein Spezialfall der Jensen-Ungleichung ist die Ljapunow-Ungleichung.

SATZ 11.2.1 (Ljapunow-Ungleichung). *Seien* $0 < s < t$ *und sei* X *eine Zufallsvariable. Dann gilt*

$$(\mathbb{E}[|X|^s])^{\frac{1}{s}} \leq \left(\mathbb{E}[|X|^t]\right)^{\frac{1}{t}}.$$

DEFINITION 11.2.2. Sei $p \geq 0$. Die L^p*-Norm* einer Zufallsvariable X ist definiert durch

$$\|X\|_p = (\mathbb{E}[|X|^p])^{1/p}.$$

BEMERKUNG 11.2.3. Mit dieser Notation besagt die Ljapunow-Ungleichung, dass

$$\|X\|_s \leq \|X\|_t, \qquad 0 < s < t.$$

BEISPIEL 11.2.4. Wir betrachten eine Zufallsvariable X, die n Werte $x_1, x_2, \ldots, x_n > 0$ mit Wahrscheinlichkeit von jeweils $1/n$ annimmt. Dann besagt die Ljapunow-Ungleichung, dass

$$\left(\frac{x_1^s + \ldots + x_n^s}{n}\right)^{\frac{1}{s}} \leq \left(\frac{x_1^t + \ldots + x_n^t}{n}\right)^{\frac{1}{t}}, \quad 0 < s < t.$$

Dies ist die klassische Ungleichung der verallgemeinerten Mittel. Setzen wir nun $s = 1$ und $t = 2$, so erhalten wir die Ungleichung zwischen dem arithmetischem und dem quadratischem Mittel:

$$\frac{x_1 + \ldots + x_n}{n} \leq \sqrt{\frac{x_1^2 + \ldots + x_n^2}{n}}.$$

BEISPIEL 11.2.5. Wir betrachten den Wahrscheinlichkeitsraum $[0, 1]$ mit der Borel-σ-Algebra und dem Lebesgue-Maß. Sei $f : [0, 1] \to \mathbb{R}$ eine messbare Funktion. Dann besagt die Ungleichung von Ljapunow, dass

$$\left(\int_0^1 |f(z)|^s dz\right)^{\frac{1}{s}} \leq \left(\int_0^1 |f(z)|^t dz\right)^{\frac{1}{t}}, \quad 0 < s < t.$$

BEWEIS VON SATZ 11.2.1. Wir führen den Beweis mit Hilfe der Jensen-Ungleichung. Sei $g(y) = |y|^{\frac{t}{s}}$. Dann ist g konvex, denn $\frac{t}{s} > 1$. Setzen wir nun $Y = |X|^s$ und nutzen die Jensen-Ungleichung, so erhalten wir

$$g(\mathbb{E}Y) \leq \mathbb{E}[g(Y)].$$

Da $g(Y) = (|X|^s)^{\frac{t}{s}} = |X|^t$, ist dies ist äquivalent zu

$$(\mathbb{E}[|X|^s])^{\frac{t}{s}} \leq \mathbb{E}[|X|^t].$$

Daraus ergibt sich die Ljapunow-Ungleichung, wenn man die $1/t$-te Potenz der beiden Seiten betrachtet. $\square$

11.3. Young-Ungleichung

Satz 11.3.1 (Young-Ungleichung). *Seien $p, q > 1$ mit der Eigenschaft, dass $1/p + 1/q = 1$. Dann gilt für alle $a, b > 0$:*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Beispiel 11.3.2. Seien $p = q = 2$. Dann erhalten wir die Ungleichung $ab \leq \frac{a^2+b^2}{2}$. Diese Ungleichung kann man schnell beweisen, indem man alle Terme auf die rechte Seite bringt: $0 \leq \frac{1}{2}(a-b)^2$.

Beweis. Wir betrachten die Funktion $y = x^{p-1}$, $x > 0$. Dabei handelt es sich, da $p > 1$ ist, um eine monoton wachsende Funktion. Die Umkehrfunktion lautet

$$x = y^{\frac{1}{p-1}} = y^{q-1}.$$

Also hat die Umkehrfunktion eine ähnliche Form wie die Ausgangsfunktion. Graphisch ergibt sich

$$ab \leq \int_0^a x^{p-1} dx + \int_0^b y^{q-1} dy.$$

Daraus ergibt sich die Young-Ungleichung. □

11.4. Hölder-Ungleichung

Satz 11.4.1 (Hölder-Ungleichung). *Seien $p, q > 1$ mit $1/p + 1/q = 1$. Seien $X \in L^p$ und $Y \in L^q$ Zufallsvariablen. Dann gilt*

$$\mathbb{E}[|XY|] \leq (\mathbb{E}[|X|^p])^{\frac{1}{p}} \cdot (\mathbb{E}[|Y|^q])^{\frac{1}{q}}.$$

Bemerkung 11.4.2. Äquivalente Schreibweise:

$$\|XY\|_1 \leq \|X\|_p \cdot \|Y\|_q.$$

Insbesondere folgt aus $X \in L^p$ und $Y \in L^q$, dass $X \cdot Y \in L^1$.

Bemerkung 11.4.3. Mit $p = q = 2$ erhalten wir die Cauchy-Schwarzsche Ungleichung

$$\|XY\|_1 \leq \|X\|_2 \cdot \|Y\|_2. \tag{11.4.1}$$

Beweis von Satz 11.4.1. Falls $\|X\|_p = 0$, so gilt $\mathbb{E}[|X|^p] = 0$. Dann ist $X = 0$ fast sicher und daher auch $X \cdot Y = 0$ fast sicher. In diesem Fall gilt die Hölder-Ungleichung, da $0 \leq 0$. Falls $\|Y\|_q = 0$, so gilt die Hölder-Ungleichung, analog zu vorherigem Fall, ebenso.
Seien nun also $\|X\|_p \neq 0$ und $\|Y\|_q \neq 0$. Wir führen die folgenden Zufallsvariablen a und b ein:

$$a = \frac{|X|}{\|X\|_p}, \qquad b = \frac{|Y|}{\|Y\|_q}.$$

Dann gilt:

$$\mathbb{E}[a^p] = 1, \quad \mathbb{E}[b^q] = 1.$$

Mit der Young-Ungleichung erhalten wir

$$\mathbb{E}[ab] \leq \mathbb{E}\left[\frac{a^p}{p}\right] + \mathbb{E}\left[\frac{b^q}{q}\right] = \frac{1}{p} + \frac{1}{q} = 1.$$

Durch Einsetzen von a und b folgt

$$\mathbb{E}\left[\frac{|X||Y|}{\|X\|_p \cdot \|Y\|_q}\right] \leq 1.$$

Da $\|X\|_p \cdot \|Y\|_q$ eine Konstante ist, darf man den Term aus dem Erwartungswert herausziehen, was zu unserer Behauptung führt:

$$\mathbb{E}|X \cdot Y| \leq \|X\|_p \cdot \|Y\|_q.$$

□

Beispiel 11.4.4. Sei $\Omega = \{1, \dots, n\}$ ein diskreter Wahrscheinlichkeitsraum mit $\mathbb{P}[\{\omega\}] = 1/n$ für alle $\omega \in \Omega$. Betrachte zwei Zufallsvariablen $X, Y : \Omega \to \mathbb{R}$ mit

$$X(k) = x_k, \qquad Y(k) = y_k, \qquad x_k, y_k \in \mathbb{R}.$$

Indem wir X und Y in die Hölder-Ungleichung einsetzen, erhalten wir

$$\sum_{i=1}^{n} |x_i y_i| \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^{n} |y_i|^q\right)^{\frac{1}{q}}.$$

Dies ist die klassische Form der Hölder-Ungleichung für Summen.

Beispiel 11.4.5. Betrachte den Wahrscheinlichkeitsraum $\Omega = [0, 1]$ mit der Borel-σ-Algebra und dem Lebesgue-Maß. Seien $f, g : [0, 1] \to \mathbb{R}$ zwei messbare Funktionen. Indem wir $X = f$ und $Y = g$ in die Hölder-Ungleichung einsetzen, erhalten wir

$$\int_0^1 |f(z)g(z)| dz \leq \left(\int_0^1 |f(z)|^p dz\right)^{\frac{1}{p}} \cdot \left(\int_0^1 |g(z)|^q dz\right)^{\frac{1}{q}}.$$

Dies ist die klassische Form der Hölder-Ungleichung für Integrale.

11.5. Minkowski-Ungleichung

Satz 11.5.1 (Minkowski-Ungleichung). *Sei $p \geq 1$ und seien $X, Y \in L^p$ Zufallsvariablen. Dann gilt*

$$\left(\mathbb{E}\left[|X + Y|^p\right]\right)^{\frac{1}{p}} \leq \left(\mathbb{E}\left[|X|^p\right]\right)^{\frac{1}{p}} + \left(\mathbb{E}[|Y|^p]\right)^{\frac{1}{p}}.$$

Bemerkung 11.5.2. Äquivalente Schreibweise:

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p.$$

Die Minkowski-Ungleichung ist also eine Dreiecksungleichung für die L^p-Norm.

Beispiel 11.5.3. Sei $\Omega = \{1, \dots, n\}$ ein diskreter Wahrscheinlichkeitsraum mit $\mathbb{P}[\{\omega\}] = 1/n$ für alle $\omega \in \Omega$. Betrachte zwei Zufallsvariablen $X, Y : \Omega \to \mathbb{R}$ mit

$$X(k) = x_k, \qquad Y(k) = y_k, \qquad x_k, y_k \in \mathbb{R}.$$

Indem wir X und Y in die Minkowski-Ungleichung einsetzen, erhalten wir

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}}.$$

Dies ist die klassische Minkowski-Ungleichung für Summen. Für $p = 2$ ist dies genau die klassische Dreiecksungleichung, die besagt, dass die Länge einer Summe von zwei Vektoren nicht größer sein kann, als die Summe der Längen der beiden Vektoren:

$$\left(\sum_{i=1}^{n}(x_i + y_i)^2\right)^{\frac{1}{2}} \le \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}} + \left(\sum_{i=1}^{n} y_i^2\right)^{\frac{1}{2}}.$$

Beispiel 11.5.4. Betrachte den Wahrscheinlichkeitsraum $\Omega = [0,1]$ mit der Borel-σ-Algebra und dem Lebesgue-Maß. Seien $f, g : [0,1] \to \mathbb{R}$ zwei messbare Funktionen. Indem wir $X = f$ und $Y = g$ in die Minkowski-Ungleichung einsetzen, erhalten wir

$$\left(\int_0^1 |f(z) + g(z)|^p dz\right)^{1/p} \le \left(\int_0^1 |f(z)|^p dz\right)^{\frac{1}{p}} + \left(\int_0^1 |g(z)|^p dz\right)^{\frac{1}{p}}.$$

Dies ist die klassische Form der Minkowski-Ungleichung für Integrale.

Beweis von Satz 11.5.1. Sei zunächst $p = 1$, dann gilt mit der Ungleichung $|X + Y| \le |X| + |Y|$:

$$\mathbb{E}|X + Y| \le \mathbb{E}(|X| + |Y|) = \mathbb{E}|X| + \mathbb{E}|Y|.$$

Also gilt die Minkowski-Ungleichung.
Nun betrachten wir den Fall $p > 1$. Wir definieren $q = \frac{p}{p-1} > 1$. Mit dieser Wahl von q gilt die Relation $\frac{1}{p} + \frac{1}{q} = 1$. Wir wenden nun die Ungleichung $|X + Y| \le |X| + |Y|$ und danach die Hölder-Ungleichung an:

$$\begin{aligned}
\mathbb{E}[|X + Y|^p] &= \mathbb{E}[|X + Y|^{p-1} \cdot |X + Y|] \\
&\le \mathbb{E}[|X + Y|^{p-1} \cdot |X|] + \mathbb{E}[|X + Y|^{p-1} \cdot |Y|] \\
&\le \left(\mathbb{E}[|X + Y|^{(p-1)q}]\right)^{\frac{1}{q}} \cdot \|X\|_p + \left(\mathbb{E}[|X + Y|^{(p-1)q}]\right)^{\frac{1}{q}} \cdot \|Y\|_p \\
&= (\mathbb{E}[|X + Y|^p])^{\frac{1}{q}} \cdot (\|X\|_p + \|Y\|_p).
\end{aligned}$$

Mit $1 - \frac{1}{q} = \frac{1}{p}$ erhalten wir

$$(\mathbb{E}[|X + Y|^p])^{\frac{1}{p}} \le \|X\|_p + \|Y\|_p.$$

□

11.6. L^p-Räume und L^p-Konvergenz

Definition 11.6.1. Sei $p > 0$. Sei X eine Zufallsvariable. Wir schreiben $X \in L^p$, wenn $\mathbb{E}[|X|^p] < \infty$. Die *$L^p$-Norm* von $X \in L^p$ ist definiert durch

$$\|X\|_p = (\mathbb{E}[|X|^p])^{\frac{1}{p}}.$$

Bemerkung 11.6.2. Die Ljapunow-Ungleichung beasgt, dass

$$\|X\|_s \le \|X\|_t, \text{ wenn } 0 < s < t.$$

Somit sind die L^p-Räume ineinander geschachtelt: $L^s \supset L^t$, wenn $s < t$. Insbesondere gilt

$$L^1 \supset L^2 \supset L^3 \supset \ldots.$$

Wir haben früher bereits gezeigt und sehr oft benutzt, dass $L^1 \supset L^2$. (Wenn eine Zufallsvariable eine Varianz besitzt, dann besitzt sie auch einen Erwartungswert).

Bemerkung 11.6.3. Für $p \geq 1$ ist L^p ein Vektorraum, denn

(1) Für $X \in L^p$ und $\lambda \in \mathbb{R}$ ist auch $\lambda X \in L^p$.
(2) Für $X \in L^p$ und $Y \in L^p$ ist auch $X + Y \in L^p$.

Die zweite Eigenschaft folgt aus der Minkowski-Ungleichung, denn

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p < \infty.$$

Für $p < 1$ ist L^p im Allgemeinen kein Vektorraum.

Definition 11.6.4. Sei $p > 1$. Der L^p-*Abstand* zwischen zwei Zufallsvariablen $X \in L^p$ und $Y \in L^p$ ist

$$d_p(X, Y) = \|X - Y\|_p = (\mathbb{E}[|X - Y|^p])^{\frac{1}{p}}.$$

Bemerkung 11.6.5. Für $X, Y \in L^p$ ist der L^p-Abstand $\|X - Y\|_p$ endlich. Das folgt aus der Minkowski-Ungleichung:

$$\|X - Y\|_p \leq \|X\|_p + \|-Y\|_p = \|X\|_p + \|Y\|_p < \infty.$$

Bemerkung 11.6.6. Für $p \geq 1$ ist der L^p-Abstand eine Metrik, denn es gilt

(1) $d_p(X, Y) = 0$ genau dann, wenn $X = Y$ *fast sicher*.
(2) $d_p(X, Y) = d_p(Y, X)$.
(3) $d_p(X, Z) \leq d_p(X, Y) + d_p(Y, Z)$.

Die letzte Eigenschaft folgt aus der Minkowski-Ungleichung, wobei hier $p \geq 1$ benutzt wird. Somit erfüllt der L^p-Raum die drei Axiome eines metrischen Raumes bis auf eine Kleinigkeit: Es kann sein, dass $d_p(X, Y) = 0$ und dennoch $X \neq Y$. Deshalb macht man eine Konvention: man betrachtet zwei Zufallsvariablen als identisch, wenn sie *fast überall* gleich sind. Dann gelten alle drei Eigenschaften eines metrischen Raumes. Ist aber $p < 1$, so gilt die Dreiecksungleichung nicht.

Definition 11.6.7. Eine Folge $X_1, X_2, \ldots$ von Zufallsvariablen *konvergiert* in L^p gegen eine Zufallsvariable X, wenn $X, X_1, X_2, \ldots \in L^p$ und

$$\lim_{n \to \infty} \|X - X_n\|_p = 0.$$

Bemerkung 11.6.8. Äquivalente Schreibweise: $\lim_{n \to \infty} \mathbb{E}[|X - X_n|^p] = 0$.

Bezeichnung. $X_n \xrightarrow[n \to \infty]{L^p} X$.

Bemerkung 11.6.9. Aus der Ljapunow-Ungleichung folgt, dass

$$X_n \xrightarrow[n \to \infty]{L^t} X \Rightarrow X_n \xrightarrow[n \to \infty]{L^s} X, \text{ wenn } s < t.$$

Daher gelten folgende Implikationen zwischen den L^p-Konvergenzen:

$$L^1 \Leftarrow L^2 \Leftarrow L^3 \Leftarrow \ldots.$$

Wenn man die Konvergenz in Wahrscheinlichkeit als L^0-Konvergenz und die gleichmäßige Konvergenz als L^∞-Konvergenz betrachtet, dann kan man dieses Schema vervollständigen:

$$L^0 \Leftarrow L^1 \Leftarrow L^2 \Leftarrow L^3 \Leftarrow \ldots \Leftarrow L^\infty.$$