Total Variation of Finite Signed Measures

Let (X, Σ) be a measurable space, and let $\mathcal{M}_s(X)$ denote the vector space of finite signed measures on (X, Σ).

1 Total variation

Theorem 1.1 (Hahn decomposition)**.** *Let $\mu \in \mathcal{M}_s(X)$. Then there exist measurable sets $P, N \in \Sigma$ such that*

$$P \cap N = \varnothing, \qquad P \cup N = X,$$

and

$$\mu(E) \geq 0 \quad \text{for every } E \in \Sigma \text{ with } E \subseteq P,$$

$$\mu(E) \leq 0 \quad \text{for every } E \in \Sigma \text{ with } E \subseteq N.$$

Such a pair (P, N) is called a Hahn decomposition for μ.

Given a Hahn decomposition, one defines

$$\mu^+(E) := \mu(E \cap P), \qquad \mu^-(E) := -\mu(E \cap N), \qquad E \in \Sigma.$$

Then μ^+ and μ^- are finite positive measures, mutually singular, and

$$\mu = \mu^+ - \mu^-.$$

This is the *Jordan decomposition* of μ.

Definition 1.2 (Total variation norm)**.** Let $\mu \in \mathcal{M}_s(X)$ with Jordan decomposition $\mu = \mu^+ - \mu^-$. The *variation measure* of μ is the finite positive measure

$$|\mu| := \mu^+ + \mu^-.$$

Its total mass

$$\|\mu\|_{\mathrm{TV}} := |\mu|(X) = \mu^+(X) + \mu^-(X)$$

is called the *total variation norm* of μ.

2 The total variation norm is a norm

$\|\mu\|_{\mathrm{TV}}$ is a norm on the vector space of finite signed measures.

2.1 Positive definiteness

Since $|\mu|$ is a positive measure,

$$\|\mu\|_{\mathrm{TV}} = |\mu|(X) \geq 0.$$

If $\|\mu\|_{\mathrm{TV}} = 0$, then

$$\mu^+(X) + \mu^-(X) = 0.$$

Since μ^+ and μ^- are positive measures, this implies

$$\mu^+(X) = 0, \qquad \mu^-(X) = 0.$$

A positive measure with total mass 0 is identically zero: indeed, for every $E \in \Sigma$,

$$0 \leq \mu^+(E) \leq \mu^+(X) = 0,$$

so $\mu^+(E) = 0$, and similarly $\mu^-(E) = 0$. Thus $\mu^+ = \mu^- = 0$, hence $\mu = 0$.

Conversely, if $\mu = 0$, then clearly $\mu^+ = \mu^- = 0$, so $\|\mu\|_{\mathrm{TV}} = 0$.

2.2 Absolute homogeneity.

Let $a \in \mathbb{R}$.

If $a \geq 0$, then

$$a\mu = a\mu^+ - a\mu^-,$$

so $(a\mu)^+ = a\mu^+$ and $(a\mu)^- = a\mu^-$. Therefore

$$\|a\mu\|_{\mathrm{TV}} = (a\mu)^+(X) + (a\mu)^-(X) = a\mu^+(X) + a\mu^-(X) = a\,\|\mu\|_{\mathrm{TV}}\,.$$

If $a < 0$, write $a = -|a|$. Then

$$a\mu = |a|(\mu^- - \mu^+),$$

hence the positive and negative parts are exchanged:

$$(a\mu)^+ = |a|\mu^-, \qquad (a\mu)^- = |a|\mu^+.$$

Thus

$$\|a\mu\|_{\mathrm{TV}} = |a|\mu^-(X) + |a|\mu^+(X) = |a|\ \|\mu\|_{\mathrm{TV}}\,.$$

2.3 Triangle inequality.

$$\begin{aligned}
|\mu + \nu|(X) &= (\mu + \nu)^+(X) - (\mu + \nu)^-(X) \\
&= (\mu + \nu)(X \cap P) - (\mu + \nu)(X \cap N) \\
&\leq |\mu|(P) + |\nu|(P) + |\mu|(N) + |\nu|(N) \\
&= (|\mu| + |\nu|)(X)
\end{aligned}$$

3 A partition formula for the variation measure

Proposition 3.1 (Partition formula)**.** *For every $E \in \Sigma$,*

$$|\mu|(E) = \sup \left\{ \sum_{i=1}^{n} |\mu(E_i)| : E = \bigsqcup_{i=1}^{n} E_i,\ E_i \in \Sigma,\ n \in \mathbb{N} \right\}.$$

In particular,

$$\|\mu\|_{\mathrm{TV}} = \sup \left\{ \sum_{i=1}^{n} |\mu(E_i)| : X = \bigsqcup_{i=1}^{n} E_i,\ E_i \in \Sigma,\ n \in \mathbb{N} \right\}.$$

Proof. Fix $E \in \Sigma$, and let

$$E = \bigsqcup_{i=1}^{n} E_i$$

be a finite measurable partition.

Since $\mu = \mu^+ - \mu^-$, we have for each i,

$$\mu(E_i) = \mu^+(E_i) - \mu^-(E_i),$$

hence

$$|\mu(E_i)| \le \mu^+(E_i) + \mu^-(E_i) = |\mu|(E_i).$$

Summing over i gives

$$\sum_{i=1}^{n} |\mu(E_i)| \le \sum_{i=1}^{n} |\mu|(E_i) = |\mu|(E),$$

because $|\mu|$ is a positive measure and the sets E_i are pairwise disjoint.

Therefore

$$\sup \left\{ \sum_{i=1}^{n} |\mu(E_i)| : E = \bigsqcup_{i=1}^{n} E_i \right\} \le |\mu|(E).$$

To prove the reverse inequality, let (P, N) be a Hahn decomposition for μ. Then

$$E = (E \cap P) \sqcup (E \cap N)$$

is a measurable partition of E. Hence

$$|\mu(E \cap P)| + |\mu(E \cap N)| = \mu^+(E) + \mu^-(E) = |\mu|(E).$$

Thus the supremum is at least $|\mu|(E)$, and equality follows. □

4 Pushforward is a contraction for total variation

Let (Y, Σ_Y) be another measurable space, and let $T : X \to Y$ be measurable.

Definition 4.1 (Pushforward)**.** If $\mu \in \mathcal{M}_s(X)$, the pushforward $T_\#\mu$ is the signed measure on (Y, Σ_Y) defined by

$$T_\#\mu(B) := \mu(T^{-1}(B)), \qquad B \in \Sigma_Y$$

Proposition 4.2. *For every measurable map $T : X \to Y$ and every $\mu \in \mathcal{M}_s(X)$,*

$$\|T_\#\mu\|_{\mathrm{TV}} \leq \|\mu\|_{\mathrm{TV}}\,.$$

Thus pushforward is a contraction in the total variation norm.

Proof. Let

$$Y = \bigsqcup_{j=1}^{m} B_j$$

be a finite measurable partition of Y. Then the inverse images $T^{-1}(B_j)$ are measurable, pairwise disjoint, and satisfy

$$X = \bigsqcup_{j=1}^{m} T^{-1}(B_j).$$

Hence, by the partition formula,

$$\sum_{j=1}^{m} |T_\#\mu(B_j)| = \sum_{j=1}^{m} \left|\mu(T^{-1}(B_j))\right| \leq \|\mu\|_{\mathrm{TV}}\,.$$

Taking the supremum over all finite measurable partitions of Y gives

$$\|T_\#\mu\|_{\mathrm{TV}} \leq \|\mu\|_{\mathrm{TV}}\,.$$

□

5 Characterization by integration against bounded measurable test functions

Let $B_b(X)$ denote the space of bounded measurable real-valued functions on X, equipped with the supremum norm

$$\|f\|_\infty := \sup_{x \in X} |f(x)| .$$

Proposition 5.1. *For every $\mu \in \mathcal{M}_s(X)$,*

$$\|\mu\|_{\mathrm{TV}} = \sup\left\{ \left| \int_X f \, d\mu \right| : f \in B_b(X),\ \|f\|_\infty \le 1 \right\}.$$

Proof. Let $f \in B_b(X)$ with $\|f\|_\infty \le 1$. Since $\mu = \mu^+ - \mu^-$,

$$\int_X f \, d\mu = \int_X f \, d\mu^+ - \int_X f \, d\mu^-.$$

Hence

$$\left| \int_X f \, d\mu \right| \le \int_X |f| \, d\mu^+ + \int_X |f| \, d\mu^- = \int_X |f| \, d|\mu|.$$

Because $\|f\|_\infty \le 1$, we have $|f| \le 1$ pointwise, so

$$\int_X |f| \, d|\mu| \le \int_X 1 \, d|\mu| = |\mu|(X) = \|\mu\|_{\mathrm{TV}} .$$

Therefore

$$\left| \int_X f \, d\mu \right| \le \|\mu\|_{\mathrm{TV}} \qquad \text{for every } f \in B_b(X) \text{ with } \|f\|_\infty \le 1.$$

Taking the supremum yields

$$\sup_{\|f\|_\infty \le 1} \left| \int_X f \, d\mu \right| \le \|\mu\|_{\mathrm{TV}} .$$

To obtain the reverse inequality, let (P, N) be a Hahn decomposition for μ, and define

$$h := \mathbf{1}_P - \mathbf{1}_N.$$

Then h is measurable and satisfies $\|h\|_\infty \le 1$. Moreover,

$$\int_X h \, d\mu = \int_P 1 \, d\mu + \int_N (-1) \, d\mu = \mu(P) - \mu(N).$$

Since

$$\mu(P) = \mu^+(X), \qquad -\mu(N) = \mu^-(X),$$

we get

$$\int_X h \, d\mu = \mu^+(X) + \mu^-(X) = \|\mu\|_{\mathrm{TV}} .$$

Thus

$$\sup_{\|f\|_\infty \le 1} \int_X f \, d\mu \ge \|\mu\|_{\mathrm{TV}} .$$

Combining this with the previous upper bound, we conclude that

$$\|\mu\|_{\mathrm{TV}} = \sup_{\|f\|_\infty \le 1} \int_X f \, d\mu.$$

□

6 Equivalent formulas in terms of suprema over measurable sets

Proposition 6.1. *For every finite signed measure* μ*,*

$$\|\mu\|_{\mathrm{TV}} = \sup_{A\in\Sigma}\big(\mu(A) - \mu(A^c)\big).$$

Equivalently,

$$\|\mu\|_{\mathrm{TV}} = \sup_{A\in\Sigma}\big(2\mu(A) - \mu(X)\big) = \sup_{A\in\Sigma}\left|2\mu(A) - \mu(X)\right|.$$

If moreover $\mu(X) = 0$*, then*

$$\|\mu\|_{\mathrm{TV}} = 2\sup_{A\in\Sigma}\left|\mu(A)\right|.$$

Proof. Write $\mu = \mu^+ - \mu^-$. Fix $A \in \Sigma$. Since

$$\mu(A) = \mu^+(A) - \mu^-(A) \le \mu^+(A)$$

and

$$-\mu(A^c) = \mu^-(A^c) - \mu^+(A^c) \le \mu^-(A^c),$$

we get

$$\mu(A) - \mu(A^c) \le \mu^+(A) + \mu^-(A^c) \le \mu^+(X) + \mu^-(X) = \|\mu\|_{\mathrm{TV}}.$$

Therefore

$$\sup_{A\in\Sigma}\big(\mu(A) - \mu(A^c)\big) \le \|\mu\|_{\mathrm{TV}}.$$

Now let (P, N) be a Hahn decomposition for μ. Take $A = P$. Then $A^c = N$, and

$$\mu(P) = \mu^+(X), \qquad \mu(N) = -\mu^-(X).$$

Hence

$$\mu(P) - \mu(N) = \mu^+(X) + \mu^-(X) = \|\mu\|_{\mathrm{TV}}.$$

So equality holds:

$$\|\mu\|_{\mathrm{TV}} = \sup_{A\in\Sigma}\big(\mu(A) - \mu(A^c)\big).$$

Since

$$\mu(A^c) = \mu(X) - \mu(A),$$

we obtain

$$\mu(A) - \mu(A^c) = 2\mu(A) - \mu(X),$$

and therefore

$$\|\mu\|_{\mathrm{TV}} = \sup_{A\in\Sigma}\big(2\mu(A) - \mu(X)\big).$$

Next observe that replacing A by A^c flips the sign:

$$2\mu(A^c) - \mu(X) = 2(\mu(X) - \mu(A)) - \mu(X) = -(2\mu(A) - \mu(X)).$$

Hence the set

$$\{2\mu(A) - \mu(X) : A \in \Sigma\}$$

is symmetric about 0, and therefore

$$\sup_{A\in\Sigma}\big(2\mu(A) - \mu(X)\big) = \sup_{A\in\Sigma}\left|2\mu(A) - \mu(X)\right|.$$

Finally, if $\mu(X) = 0$, then this simplifies to

$$\|\mu\|_{\mathrm{TV}} = \sup_{A\in\Sigma}\left|2\mu(A)\right| = 2\sup_{A\in\Sigma}\left|\mu(A)\right|.$$

□

7 Total variation distance on probability measures

Now suppose P and Q are probability measures on (X, Σ).

Definition 7.1 (Total variation distance)**.** The *total variation distance* between P and Q is defined by

$$d_{\mathrm{TV}}(P, Q) := \sup_{A \in \Sigma} |P(A) - Q(A)| .$$

Proposition 7.2. *For probability measures P, Q,*

$$d_{\mathrm{TV}}(P, Q) = \frac{1}{2} \|P - Q\|_{\mathrm{TV}} .$$

Proof. Let

$$\nu := P - Q.$$

Then ν is a finite signed measure and

$$\nu(X) = P(X) - Q(X) = 1 - 1 = 0.$$

By the formula proved above for signed measures with total mass zero,

$$\|\nu\|_{\mathrm{TV}} = 2 \sup_{A \in \Sigma} |\nu(A)| .$$

Since $\nu(A) = P(A) - Q(A)$, this becomes

$$\|P - Q\|_{\mathrm{TV}} = 2 \sup_{A \in \Sigma} |P(A) - Q(A)| = 2\, d_{\mathrm{TV}}(P, Q).$$

Therefore

$$d_{\mathrm{TV}}(P, Q) = \frac{1}{2} \|P - Q\|_{\mathrm{TV}} .$$

□