Capacity

Mário S. Alvim
(msalvim@dcc.ufmg.br)

Quantitative Information Flow

DCC-UFMG
(2021/1)

Introduction

- The definition of g-leakage compares prior and posterior vulnerabilities for a fixed prior π and gain function g.
- But in our introductory discussion about robustness, we noted that in general we can't be sure of the π's and g's our channels might face in practice.

  So it's worth asking:

  "What is the worst leakage that we, as defenders, might have to deal with if we use this channel?"

  There are two dimensions over which "worst" might vary:

  1. Priors: rather than just some particular π, we might want to consider:

     a) all priors, i.e., the whole set $\mathbb{D}\mathcal{X}$, or b) some subset $\mathcal{D}$ of $\mathbb{D}\mathcal{X}$.

  2. Gain functions: rather than just some particular g, we might consider:

     a) all gain functions $\mathbb{G}\mathcal{X}$, or b) some subset $\mathcal{G}$ of them.

Introduction

- Here we'll see interesting limits to (multiplicative and additive) leakage that are induced by the variations in the priors, or in the gain functions, or both.

  This is a study of channel capacity, i.e., maximum leakage.

- Moreover, we'll see that certain capacities can be computed efficiently, while others cannot.

Multiplicative Bayes capacity

Multiplicative Bayes capacity

- We begin by developing the theory of multiplicative Bayes capacity.

Definition (7.1)

The **multiplicative Bayes capacity** of channel C, denoted $\mathcal{ML}_1^{\times}(\mathbb{D}, C)$, is the maximum multiplicative Bayes leakage over all priors:

$$\mathcal{ML}_1^{\times}(\mathbb{D}, C) \ := \sup_{\pi \in \mathbb{D}\mathcal{X}} \mathcal{L}_1^{\times}(\pi, C) \quad .$$

Note that:

- we add "$\mathcal{M}$" to our usual notation "$\mathcal{L}_1$" to indicate that we are maximizing, and
- write $\mathbb{D}$ instead of π to indicate that the prior varies over all distributions in $\mathbb{D}\mathcal{X}$.

Moreover, both the channel and the gain function are fixed, the latter because it is Bayes capacity we are examining: only the prior varies.

Multiplicative Bayes capacity

- It's not immediately clear how $\mathcal{ML}_1^{\times}(\mathbb{D}, C)$ could be computed, since its definition involves a supremum over the infinite set $\mathbb{D}\mathcal{X}$.
- But multiplicative Bayes capacity is always realized on a uniform prior.

  Hence (by Thm. 5.17) it's easy to compute it: just take any channel matrix C for C, and then sum its column maximums.

Theorem (7.2)

For any channel C, the multiplicative Bayes capacity $\mathcal{ML}_1^{\times}(\mathbb{D}, C)$ is always realized on a uniform prior ϑ.

Moreover, for channel matrices we have that

$$\mathcal{ML}_1^{\times}(\mathbb{D}, \mathsf{C}) = \sum_{y \in \mathcal{Y}} \max_{x \in \mathcal{X}} \mathsf{C}_{x,y} \ ,$$

where C *is a concrete channel corresponding to the abstract channel C .*

Multiplicative Bayes capacity

- **Proof.** Consider an arbitrary channel matrix C such that $[\![\mathsf{C}]\!] = C$.

  Observe that for any prior π, we have

$$
\begin{array}{lll}
 & \mathcal{L}_1^{\times}(\pi, \mathsf{C}) & \\
= & V_1[\pi \triangleright \mathsf{C}] \,/\, V_1(\pi) & \\
= & \sum_{y \in \mathcal{Y}} \max_{x \in \mathcal{X}} (\pi_x \mathsf{C}_{x,y}) \,/\, \max_{x \in \mathcal{X}} \pi_x & \text{“Thm. 5.15”} \\
\leq & \sum_{y \in \mathcal{Y}} \max_{x \in \mathcal{X}} (\max_{x \in \mathcal{X}} \pi_x) \mathsf{C}_{x,y} \,/\, \max_{x \in \mathcal{X}} \pi_x & \\
= & \sum_{y \in \mathcal{Y}} \max_{x \in \mathcal{X}} \mathsf{C}_{x,y} \quad . &
\end{array}
$$

  The upper bound is clearly realized when π is uniform, because in that case all π_x's are equal and the third step above becomes an equality.

  (Note however that the upper bound can also be realized on a non-uniform π, provided that some proper subset of the rows of C includes at least one maximum from each column.) □

Multiplicative Bayes capacity

- The theorem above has two useful corollaries.

Corollary (7.3)

If C *is deterministic, then* $\mathcal{ML}_1^{\times}(\mathbb{D}, \mathsf{C})$ *is the number of possible output values of* C*, where an output value is* possible *just when its corresponding column contains at least one nonzero entry.*

Proof. If C is deterministic, then all of its entries are either 0 or 1.

Hence the sum of its column maximums is simply the number of columns containing at least one nonzero entry, which is the number of possible output values. □

- This corollary helps automated leakage analysis of a deterministic program.

  Multiplicative Bayes capacity can be computed simply by counting the number of distinct outputs that the program can produce!

Multiplicative Bayes capacity

- A second useful corollary is the following.

Corollary (7.4)

$\mathcal{ML}_1^{\times}(\mathbb{D}, \mathsf{C}) = 1$ *iff the rows of* C *are identical.*

Proof. By Thm. 7.2, the multiplicative Bayes capacity of C is 1 iff the sum of its column maximums is 1. But each row of C sums to 1, so after considering the first row of C the sum of the column maximums is already 1. If any subsequent row differs at all from the earlier rows, it increases the sum of the column maximums above 1. On the other hand, if the rows of C are identical, then the sum of the column maximums is 1 exactly. □

- Note that if the rows of C are identical then the output Y of C is completely independent of the input X. Hence C is the non-interferent channel $\mathbb{1}$ which leaks nothing, and the "if" direction of the corollary above is unsurprising.
- But the "only if" direction is more interesting: the only way for a channel to have no Bayes leakage on a uniform prior is for it to be non-interferent.

Multiplicative Bayes capacity

- Multiplicative Bayes capacity $\mathcal{ML}_1^{\times}(\mathbb{D}, C)$ also has very important relationships with other leakage measures.

  These relationships make Bayes capacity a very robust measure of the worst-case leakage of C.

- Most remarkably, it turns out that it's an upper bound on the multiplicative g-leakage of C, regardless of the prior π and non-negative gain function g!

Theorem (7.5 - "Miracle")

For any channel C, prior π, and non-negative gain function $g: \mathbb{G}^{+}\mathcal{X}$, we have

$$\mathcal{L}_g^{\times}(\pi, C) \quad \leq \quad \sum_{y:\mathcal{Y}} \max_{x:\mathcal{X}} C_{x,y} \quad = \quad \mathcal{ML}_1^{\times}(\mathbb{D}, C) \quad .$$

Multiplicative Bayes capacity

- **Proof.** In the following, variables x, y, w range over $\mathcal{X}, \mathcal{Y}, \mathcal{W}$ in the corresponding sums, max's, and sup's; the sets themselves are omitted for brevity.

  We have that

$$
\begin{aligned}
& V_g[\pi \triangleright C] \\
= \quad & \textstyle\sum_y \sup_w \sum_x g(w,x)\pi_x C_{x,y} && \text{“Thm. 5.7, for infinite } \mathcal{W}\text{”} \\
\leq \quad & \textstyle\sum_y \sup_w \sum_x g(w,x)\pi_x \left(\max_x C_{x,y}\right) && \text{“}g(w,x)\pi_x C_{x,y} \leq g(w,x)\pi_x \left(\max_x C_{x,y}\right), \\
& && \text{since } g(w,x)\pi_x \geq 0\text{”} \\
= \quad & \left(\sup_w \textstyle\sum_x \pi_x g(w,x)\right) \left(\textstyle\sum_y \max_x C_{x,y}\right) && \text{“}\textstyle\sum_i cx_i = c\sum_i x_i \\
& && \text{and } \sup_i cx_i = c \sup_i x_i \text{ for } c \geq 0\text{”} \\
= \quad & V_g(\pi)\ \mathcal{ML}_1^{\times}(\mathbb{D}, C) \quad . && \text{“Thm. 7.2”}
\end{aligned}
$$

Multiplicative Bayes capacity

- **Proof (continued).** Note that the "$\leq$" step, above, uses crucially the assumption that $g(w,x) \geq 0$, i.e. that g is a non-negative gain function.

  Finally, we have

$$\mathcal{L}_g^{\times}(\pi, C) \quad = \quad \frac{V_g[\pi \triangleright C]}{V_g(\pi)} \quad \leq \quad \frac{V_g(\pi)\, \mathcal{ML}_1^{\times}(\mathbb{D}, C)}{V_g(\pi)} \quad = \quad \mathcal{ML}_1^{\times}(\mathbb{D}, C) \quad .$$

  □

- Note that a generic gain function $g\colon \mathbb{G}\mathcal{X}$ is allowed to take negative gain values (as long as the induced V_g is always non-negative).

  But in the Theorem above we limit ourselves to non-negative gain-functions in $\mathbb{G}^{+}\mathcal{X}$.

  Indeed, the Miracle theorem can actually fail if g is somewhere negative.

Multiplicative Bayes capacity

- Example 1 Recall the gain function g_{tiger}, in which $\mathcal{W} = \mathcal{X} \cup \{\perp\}$, where the special action $\perp$ is used to opt not to input a guess, and

$$g_{\text{tiger}}(w, x) := \begin{cases} 1 & \text{if } w = x \\ 0 & \text{if } w = \perp \\ -1 & \text{otherwise} \end{cases} .$$

Consider $\mathcal{X} = \{x_1, x_2\}$ and an "almost" uniform prior distribution

$$\pi = (0.5 + \epsilon, 0.5 - \epsilon)$$

for some $\epsilon > 0$.

In this case the best action is to guess x_1, and the expected gain is then

$$(0.5 + \epsilon) \cdot 1 + (0.5 - \epsilon) \cdot (-1) = 2\epsilon ,$$

so that

$$V_{g_{\text{tiger}}}(\pi) = 2\epsilon .$$

Multiplicative Bayes capacity

- Example 1 (Continued)

Now let C be the following channel matrix, which gives rather good information about the secret:

C	y_1	y_2
x_1	0.8	0.2
x_2	0.2	0.8

Using Thm. 5.18 we calculate the posterior g_{tiger}-vulnerability as the sum of the column maximums of the matrix $\mathsf{G} \cdot \mathsf{J}$ as given below:

G	x_1	x_2
x_1	1	-1
$\perp$	0	0
x_2	-1	1

$\cdot$

J	y_1	y_2
x_1	$0.8(0.5+\epsilon)$	$0.2(0.5+\epsilon)$
x_2	$0.2(0.5-\epsilon)$	$0.8(0.5-\epsilon)$

$=$

GJ	y_1	y_2
x_1	$0.3+\epsilon$	$-0.3+\epsilon$
$\perp$	0	0
x_2	$-0.3-\epsilon$	$0.3-\epsilon$

Multiplicative Bayes capacity

- Example 1 (Continued)

G	x_1	x_2
x_1	1	-1
$\perp$	0	0
x_2	-1	1

$\cdot$

J	y_1	y_2
x_1	$0.8(0.5+\epsilon)$	$0.2(0.5+\epsilon)$
x_2	$0.2(0.5-\epsilon)$	$0.8(0.5-\epsilon)$

$=$

GJ	y_1	y_2
x_1	$0.3+\epsilon$	$-0.3+\epsilon$
$\perp$	0	0
x_2	$-0.3-\epsilon$	$0.3-\epsilon$

Hence $$V_{g_{\text{tiger}}}[\pi \triangleright \mathsf{C}] = 0.3+\epsilon+0.3-\epsilon = 0.6$$

and $$\mathcal{L}^{\times}_{g_{\text{tiger}}}(\pi, \mathsf{C}) \quad = \quad \frac{0.6}{2\epsilon} \quad ,$$

which can be arbitrarily large as ϵ approaches 0.

Hence it can be arbitrarily larger than the multiplicative Bayes capacity:

$$\mathcal{ML}^{\times}_1(\mathbb{D}, \mathsf{C}) \quad = \quad 0.8+0.8 \quad = \quad 1.6 \quad .$$

Multiplicative Bayes capacity

- Another interesting aspect of multiplicative Bayes capacity is its relationship with Shannon capacity, i.e., the maximum mutual information over all priors.

Theorem (7.7)

If C is deterministic, then its Shannon capacity is equal to the logarithm of its multiplicative Bayes capacity.

Proof. In the textbook. □

- In the general case of probabilistic channels, the logarithm of multiplicative Bayes capacity gives at least an upper bound on Shannon capacity.

Theorem (7.8)

For any channel C, the logarithm of C's multiplicative Bayes capacity is at least as great as its Shannon capacity.

Proof. In the textbook. □

Additive Bayes capacity

Additive Bayes capacity

- Now we turn our attention to additive Bayes capacity.

Definition

The **additive Bayes capacity** of channel C, denoted $\mathcal{ML}_1^+(\mathbb{D}, C)$, is the maximum additive Bayes leakage over all priors π:

$$\mathcal{ML}_1^+(\mathbb{D}, C) \ := \sup_{\pi \in \mathbb{D}\mathcal{X}} \mathcal{L}_1^+(\pi, C) \quad .$$

- Interestingly, additive Bayes capacity is far harder to compute than multiplicative Bayes capacity.

  The key challenge is that additive Bayes capacity need not be realized on a uniform prior, as the following example shows.

Additive Bayes capacity

- Example 2 On channel matrix

C	y_1	y_2	y_3
x_1	0	$1/2$	$1/2$
x_2	$1/2$	0	$1/2$
x_3	$1/2$	$1/2$	0

the additive Bayes leakage on the uniform prior $\vartheta = (1/3, 1/3, 1/3)$ is

$$\mathcal{L}_1^+(\vartheta, \mathsf{C}) \quad = \quad V_1[\vartheta \triangleright \mathsf{C}] - V_1(\vartheta) \quad = \quad 1/2 - 1/3 \quad = \quad 1/6 \quad .$$

But the additive Bayes leakage on the non-uniform prior $\pi = (1/2, 1/2, 0)$ is higher:

$$\mathcal{L}_1^+(\pi, \mathsf{C}) \quad = \quad V_1[\pi \triangleright \mathsf{C}] - V_1(\pi) \quad = \quad 3/4 - 1/2 \quad = \quad 1/4 \quad .$$

◁

Additive Bayes capacity

- However, it turns out that additive Bayes capacity is always realized on some **sub-uniform distribution**, i.e., a distribution that is uniform on its support.

Theorem (7.11)

For any channel C, the additive Bayes capacity $\mathcal{ML}_1^+(\mathbb{D}, C)$ is always realized on a sub-uniform distribution.

Proof. Given in the textbook. □

Additive Bayes capacity

- Thm. 7.11 above implies that the additive Bayes capacity of C is computable.
  - For if $|\mathcal{X}| = n$, then there are $2^n - 1$ sub-uniform distributions, and we can simply compute the additive Bayes leakage on all of them.
  - But as $2^n - 1$ is exponential in the size of C, that's not an efficient algorithm.
- The following theorem shows that an efficient algorithm probably doesn't exist.

Theorem (7.12)

Given a channel matrix C *and a threshold* t*, it is NP-complete to decide whether* $\mathcal{ML}_1^+(\mathbb{D}, \mathsf{C}) \geq t$.

Proof. Given in the textbook. □

General capacities

General capacities

- The leakage of a channel C depends on both π and g, so we can maximize it:
  - over a set of gain functions in $\mathcal{G} \subseteq \mathbb{G}\mathcal{X}$ and
  - a set of priors in $\mathcal{D} \subseteq \mathbb{D}\mathcal{X}$.

Definition ($(\mathcal{G}, \mathcal{D})$-capacity)

For classes $\mathcal{G} \subseteq \mathbb{G}\mathcal{X}$, $\mathcal{D} \subseteq \mathbb{D}\mathcal{X}$ and channel C, the multiplicative and additive $(\mathcal{G}, \mathcal{D})$-**capacities** of C are given by

$$\mathcal{ML}^{\times}_{\mathcal{G}}(\mathcal{D}, C) := \sup_{g:\mathcal{G}, \pi:\mathcal{D}} \mathcal{L}^{\times}_{g}(\pi, C) \quad \text{and}$$

$$\mathcal{ML}^{+}_{\mathcal{G}}(\mathcal{D}, C) := \sup_{g:\mathcal{G}, \pi:\mathcal{D}} \mathcal{L}^{+}_{g}(\pi, C) \quad .$$

Notation:

- To maximize over π, for a fixed g, we take $\mathcal{G} = \{g\}$ and write $\mathcal{ML}^{\times}_{g}(\mathcal{D}, C)$.
- Similarly, when π is fixed we write $\mathcal{ML}^{\times}_{\mathcal{G}}(\pi, C)$.
- When $\mathcal{G}{=}\mathbb{G}\mathcal{X}$ and $\mathcal{D}{=}\mathbb{D}\mathcal{X}$, we write $\mathcal{ML}^{\times}_{\mathbb{G}}(\mathbb{D}, C)$ instead of $\mathcal{ML}^{\times}_{\mathbb{G}\mathcal{X}}(\mathbb{D}\mathcal{X}, C)$.

General capacities

- Let's now briefly talk about the types of maximization allowed in the capacities we consider.
- $(g, \mathbb{D})$-**capacities: fixed** g**, maximize over** π**.**
  - When g is g_{id}, this scenario corresponds to the Bayes capacities already discussed.
  - We can construct a gain function $g_H \colon \mathbb{G}\mathcal{X}$ such that $\mathcal{ML}^{+}_{g_H}(\mathbb{D}, C)$ is the Shannon capacity,

General capacities

- $(\mathcal{G}, \pi)$-**capacities: fixed π, maximize over g**

  Here an unrestricted maximization over $\mathbb{G}\mathcal{X}$ leads to unbounded leakage.

  Recall from Thm. 5.13 how leakage is affected by shifting and scaling of g.

  - Multiplicative leakage:
    - is invariant under scaling: $\mathcal{L}^{\times}_{g_{\times k}}(\pi, C) = \mathcal{L}^{\times}_{g}(\pi, C)$, but
    - is affected by shifting: $\mathcal{L}^{\times}_{g_{+\kappa}}(\pi, C) = (1-\lambda)\mathcal{L}^{\times}_{g}(\pi, C) + \lambda$, where $\lambda = {}^{\kappa\cdot\pi}/_{V_{g_{+\kappa}}(\pi)}$.

      Note that shifting g down always increases multiplicative leakage. Hence we restrict to the class $\mathbb{G}^{+}\mathcal{X}$ of non-negative gain functions, essentially preventing g from being arbitrarily shifted down.
  - Additive leakage:
    - is invariant under shifting: $\mathcal{L}^{+}_{g_{+\kappa}}(\pi, C) = \mathcal{L}^{+}_{g}(\pi, C)$, but
    - is affected by scaling: $\mathcal{L}^{+}_{g_{\times k}}(\pi, C) = k\ \mathcal{L}^{+}_{g}(\pi, C)$.

      Note that by scaling g up the additive leakage always increases. Hence we restrict to the class $\mathbb{G}^{\updownarrow}\mathcal{X}$ of 1-bounded gain functions, essentially preventing g from being arbitrarily scaled up.

General capacities

- $(\mathcal{G}, \mathbb{D})$-**capacities: maximize over both** g **and** π

  Given the discussion about $(\mathcal{G}, \pi)$-capacities, in this final scenario the leakage clearly also becomes unbounded when maximizing over $\mathbb{G}\mathcal{X}$.

  Hence also here we'll restrict to:

  - $\mathbb{G}^{+}\mathcal{X}$ in the multiplicative case, and
  - $\mathbb{G}^{\updownarrow}\mathcal{X}$ in the additive case.

General capacities

- Given the above remarks, we study 6 instantiations of $(\mathcal{G}, \mathcal{D})$-capacity.

		Maximization over		
		Fixed g, max. over π	**Fixed π, max. over g**	**Max. over both π and g**
Type of leakage	**Multiplicative**	$\mathcal{ML}_g^{\times}(\mathbb{D}, C)$	$\mathcal{ML}_{\mathbb{G}^+}^{\times}(\pi, C)$	$\mathcal{ML}_{\mathbb{G}^+}^{\times}(\mathbb{D}, C)$
	Additive	$\mathcal{ML}_g^{+}(\mathbb{D}, C)$	$\mathcal{ML}_{\mathbb{G}^\updownarrow}^{+}(\pi, C)$	$\mathcal{ML}_{\mathbb{G}^\updownarrow}^{+}(\mathbb{D}, C)$

- Recall that:
  - $\mathbb{G}^+$ denotes $\mathbb{G}^+\mathcal{X}$, the set of **non-negative gain functions** consisting in all $g\colon \mathbb{G}\mathcal{X}$ that are themselves ≥ 0.
  - $\mathbb{G}^\updownarrow$ denotes $\mathbb{G}^\updownarrow\mathcal{X}$, the set of **one-bounded gain functions** consisting in all $g\colon \mathbb{G}\mathcal{X}$ that are themselves ≤ 1. (This class coincides all g st. V_g is in $[0, 1]$.)

General capacities

- In the textbook we carefully derive results regarding the existence of efficient algorithms for all of these scenarios.

  Here we'll just summarize them.

- But, to fully appreciate them, we should recall the following useful definitions.

Definition (3.7)

For $\sigma\colon \mathbb{D}\mathcal{X}$, the **distribution-reciprocal gain function** $g_{\sigma^{-1}}$ is given by

$$\mathcal{W}=\lceil\sigma\rceil$$

and

$$g_{\sigma^{-1}}(w,x) \;:=\; \begin{cases} 1/\sigma_x & \text{if } w = x \\ 0 & \text{otherwise} \end{cases}.$$

- Moreover, for a gain function $g\colon \mathbb{G}^{\updownarrow}$, let g^c be its **complement gain function**, defined as $g^c(w,x) = 1 - g(w,x)$ for all w, x.

General capacities

- Results from the textbook regarding the existence of efficient algorithms.

	Fixed g, max. over π	**Fixed π, max. over g**	**Max. over both π and g**
Mult. leakage	$\mathcal{ML}_g^{\times}(\mathbb{D}, C)$: Algorithm unknown. **(Open problem)**	$\mathcal{ML}_{\mathbb{G}^+}^{\times}(\pi, C) = \mathcal{L}_{g_{\pi^{-1}}}^{\times}(\pi, C) = \sum_{y:\mathcal{Y}} \max_{x:\lceil\pi\rceil} C_{x,y}$ **(Thm. 7.14)**	$\mathcal{ML}_{\mathbb{G}^+}^{\times}(\mathbb{D}, C) = \mathcal{ML}_1^{\times}(\mathbb{D}, C) = \sum_{y:\mathcal{Y}} \max_{x:\mathcal{X}} C_{x,y}$ **(Thm. 7.5)**
Add. leakage	$\mathcal{ML}_g^{+}(\mathbb{D}, C)$: Decision problem is NP-Complete. **(Thm. 7.12)**	$\mathcal{ML}_{\mathbb{G}^{\updownarrow}}^{+}(\pi, C) = \mathcal{L}_{g^c_{\pi^{-1}}}^{+}(\pi, C) = 1 - \sum_{y:\mathcal{Y}} \min_{x:\lceil\pi\rceil} C_{x,y}$ **(Thm. 7.21)**	$\mathcal{ML}_{\mathbb{G}^{\updownarrow}}^{+}(\mathbb{D}, C) = \mathcal{L}_{g^c_{\pi^{-1}}}^{+}(\pi, C),$ for any full support π $= 1 - \sum_{y:\mathcal{Y}} \min_{x:\mathcal{X}} C_{x,y}$ **(Thm. 7.21)**

Obtaining bounds on leakage

Obtaining bounds on leakage

- Capacities are useful when the exact operational scenario is unknown.
- But even when the adversary is fixed (expressed by some gain function g and prior knowledge π), capacities can be useful for obtaining bounds on g-leakage or on g-vulnerability.
  - Of course, for fixed g and π we could compute leakage directly...
  - ... but such a computation might be challenging if, for instance, the size of the channel is big or the gain function g is hard to represent.
- Here we'll:
  1. Formulate an "Additive miracle" theorem, providing an additive leakage bound for the class $\mathbb{G}^{\updownarrow}\mathcal{X}$ (similar to the multiplicative one for $\mathbb{G}^{+}\mathcal{X}$).
  2. Show how we can use gain-function algebra to extend those bounds to (almost) any gain function in $\mathbb{G}\mathcal{X}$.
  3. Provide examples illustrating the bounds.

Obtaining bounds on leakage: The additive miracle theorem

- In the multiplicative case we used the Miracle theorem (Thm. 7.5) which provides an upper bound for g-leakage, to compute the multiplicative $(\mathbb{G}^+,\pi)$- and $(\mathbb{G}^+,\mathbb{D})$-capacities.

  In the additive case, having already a solution for capacity, we can go in the opposite direction and obtain a corresponding "Additive miracle" theorem, giving a tight upper bound for additive g-leakage.

Theorem (7.23 - "Additive Miracle")

For any channel C, prior π, and gain function $g\colon \mathbb{G}^{\updownarrow}\mathcal{X}$, we have

$$\mathcal{L}^+_g(\pi, C) \quad \leq \quad 1 - \sum_{y\colon \mathcal{Y}} \min_{x\colon \mathcal{X}} C_{x,y} \quad .$$

Proof. Given in the textbook. □

Obtaining bounds on leakage: Improved miracle bounds

- If g belongs to either $\mathbb{G}^{+}\mathcal{X}$ or $\mathbb{G}^{\updownarrow}\mathcal{X}$ then the Miracle theorems can be applied.

  For a generic $g\colon \mathbb{G}\mathcal{X}$, however, we can still obtain bounds by:

  1. First constructing a scaled or shifted version g' that falls inside one of those classes, obtaining a bound for g', and then
  2. transforming it into a bound for g using Thm. 5.13 (about leakage invariance).

Obtaining bounds on leakage: Improved miracle bounds

- That technique is employed in the following theorem, providing generic bounds for (almost) any $g\colon \mathbb{G}\mathcal{X}$.

Theorem (7.24)

For any channel C, prior π, and gain function g in $\mathbb{G}\mathcal{X}$, let k be the supremum of g. Then we have

$$\mathcal{L}_g^+(\pi, C) \quad \leq \quad k\big(1 - \sum_{y\colon \mathcal{Y}} \min_{x\colon \mathcal{X}} C_{x,y}\big) \quad .$$

Now assume that g is bounded from below and $V_g(\pi) \neq 0$. Letting $\kappa\colon \mathbb{R}^{|\mathcal{X}|}$ be the vector of per-secret gain infimums (i.e. each κ_x is the infimum of $g(\cdot, x)$), we have

$$\mathcal{L}_g^\times(\pi, C) \quad \leq \quad \big(\sum_{y\colon \mathcal{Y}} \max_{x\colon \mathcal{X}} C_{x,y}\big)(1-c) + c \quad \textit{where} \quad c := \frac{\kappa \cdot \pi}{V_g(\pi)} \quad .$$

Proof. Given in the textbook. □

Obtaining bounds on leakage: Improved miracle bounds

- Note that:
  - All gain functions $g\colon \mathbb{G}\mathcal{X}$ are bounded from above, so the additive bound of the above theorem is always applicable.
  - The multiplicative bound, however, is not applicable when a gain function is not bounded from below, or when the prior g-vulnerability is 0.
  - These improved bounds might be of interest also in the case when g does belongs to either $\mathbb{G}^{+}\mathcal{X}$ or $\mathbb{G}^{\updownarrow}\mathcal{X}$.

    That is because the bounds might be smaller than those provided directly by the Miracle theorems, since the extremums of g are taken into account.

    (a) For instance, consider a gain function with values in $[0, 1/2]$.

    Since its supremum is $k = 1/2$, the bound from Thm. 7.24 will be half the bound from the additive Miracle theorem.

    A similar improvement to the multiplicative−Miracle-theorem bound happens with gain functions $g\colon \mathbb{G}^{+}\mathcal{X}$ that are strictly positive.

- We now illustrate the use of our leakage bounds.

Obtaining bounds on leakage: Examples

- Example 3 **Tiger leakage.** Let's consider:

the channel C

C	y_1	y_2
x_1	0.8	0.2
x_2	0.2	0.8

,

the gain function g_{tiger}

G	x_1	x_2
x_1	1	-1
x_2	-1	1
$\perp$	0	0

,

and a prior that is either:

- the uniform ϑ, or
- a quasi-uniform $\pi = (0.51, 0.49)$.

Then:

$$V_{g_{\text{tiger}}}(\vartheta) = 0 \qquad V_{g_{\text{tiger}}}[\vartheta \triangleright C] = 0.6$$
$$V_{g_{\text{tiger}}}(\pi) = 0.02 \qquad V_{g_{\text{tiger}}}[\pi \triangleright C] = 0.6 \quad .$$

But in this example assume that we cannot compute the posterior vulnerabilities directly, and are therefore interested in bounding them.

Obtaining bounds on leakage: Examples

- Example 3 (Continued)

A vulnerability bound can be obtained from a corresponding leakage bound.

However, since g_{tiger} is not non-negative, the original "Miracle" theorem (Thm. 7.5) doesn't apply.

But for π we can apply the improved multiplicative bound of Thm. 7.24.

The vector of gain infimums is $\kappa = (-1, -1)$, from which we compute

$$c \quad = \quad \kappa \cdot \pi \,/\, V_{g_{\text{tiger}}}(\pi) \quad = \quad -1 \,/\, 0.02 \quad = \quad -50 \quad .$$

Applying the bound, we get

$$\mathcal{L}^{\times}_{g_{\text{tiger}}}(\pi, \mathsf{C}) \quad \leq \quad \Big(\sum_{y:\,\mathcal{Y}} \max_{x:\,\mathcal{X}} \mathsf{C}_{x,y}\Big)(1-c)+c \quad = \quad (0.8+0.8)\cdot 51-50 \quad = \quad 31.6 \quad ,$$

from which we conclude that

$$V_{g_{\text{tiger}}}[\pi \triangleright C] \leq 31.6 \cdot V_{g_{\text{tiger}}}(\pi) = 0.632 \; .$$

Obtaining bounds on leakage: Examples

- Example 3 (Continued)

But for ϑ the bound of Thm. 7.24. is not applicable, since $V_{g_{\text{tiger}}}(\vartheta) = 0$.

We can also employ the additive bounds for our purpose. In this case, the additive Miracle theorem (Thm. 7.23) is directly applicable since $V_{g_{\text{tiger}}}$ is clearly bounded by 1, that is $g_{\text{tiger}} \colon \mathbb{G}^{\updownarrow}\mathcal{X}$.

From the additive Miracle theorem we get that

$$\mathcal{L}^{+}_{g_{\text{tiger}}}(\pi, \mathsf{C}) \quad \leq \quad 1 - \textstyle\sum_y \min_x \mathsf{C}_{x,y} \quad = \quad 1 - 0.2 - 0.2 \quad = \quad 0.6 \quad ,$$

from which we conclude that

$$V_{g_{\text{tiger}}}[\pi \triangleright C] \leq 0.6 + V_{g_{\text{tiger}}}(\pi) = 0.62 \ ,$$

a bound slightly better than the one obtained via multiplicative leakage.

Obtaining bounds on leakage: Examples

- Example 3 (Continued)

  But the additive Miracle theorem (Thm. 7.23) is also applicable for ϑ.

  In fact the leakage bound itself does not depend on the prior, hence we get

$$\mathcal{L}^{+}_{g_{\text{tiger}}}(\vartheta, \mathsf{C}) \leq 0.6 \ ,$$

  from which we conclude that

$$V_{g_{\text{tiger}}}[\vartheta \triangleright C] \leq 0.6 + V_{g_{\text{tiger}}}(\vartheta) = 0.6 \ .$$

  Remarkably, we obtain a tight bound, equal to the actual posterior vulnerability. ◁

Appendix:
Obtaining leakage bounds for Shannon leakage.

Obtaining bounds on leakage: Examples

- Example 4 **Shannon leakage.** Consider the channel

C	y_1	y_2	y_3
x_1	0.5	0.39	0.11
x_2	0.5	0.40	0.10
x_3	0.5	0.41	0.09

,

Assume that we are interested in its Shannon leakage (i.e. its Shannon mutual information) for some unknown prior π.

Notice that the channel is almost noninterfering — its rows differ only slightly.

Because of this, we intuitively expect its Shannon leakage to be small; but it is not zero, so we wish to establish a concrete bound.

Of course Shannon capacity provides such a bound, but computing it requires that we apply the iterative Blahut-Arimoto algorithm.

Is there a way to obtain a simple bound faster?

Obtaining bounds on leakage: Examples

- Example 4 (Continued)

For this purpose, recall that there is a gain function g_H that gives the complement of Shannon entropy; that is, $V_{g_H}(\pi) = \log_2(|\mathcal{X}|) - H(\pi)$.

Note also that $\mathcal{L}^+_{g_H}(\pi, \mathrm{C})$ is exactly equal to the Shannon mutual information of C, so our goal is simply to bound C's additive g_H-leakage.

Since V_{g_H} takes values in $[0, \log_2 3]$, gain function g_H does not belong to $\mathbb{G}^{\updownarrow}\mathcal{X}$ and hence the additive Miracle theorem is not directly applicable.

But we can still apply Thm. 7.24; the supremum of g_H is $k = \log_2 3$, hence

$$\mathcal{L}^+_{g_H}(\pi, \mathrm{C}) \quad \leq \quad \log_2 3 \cdot (1 - \textstyle\sum_y \min_x \mathrm{C}_{x,y}) \quad = \quad \log_2 3 \cdot 0.02 \quad \approx \quad 0.032 \quad .$$

As expected, we have shown that C's Shannon leakage is necessarily small.

(Note that C's actual Shannon capacity is approximately 0.001, so our bound is not tight but still useful.) ◁