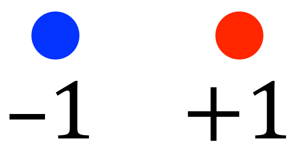
On
Consistent Surrogate Risk Minimization
and
Property Elicitation

Arpit Agarwal and Shivani Agarwal

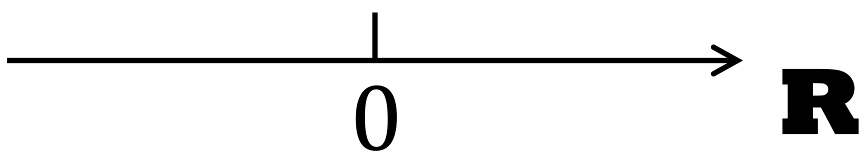
Department of Computer Science & Automation
Indian Institute of Science

COLT 2015

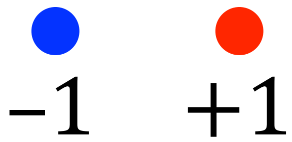
Surrogate Risk Minimization



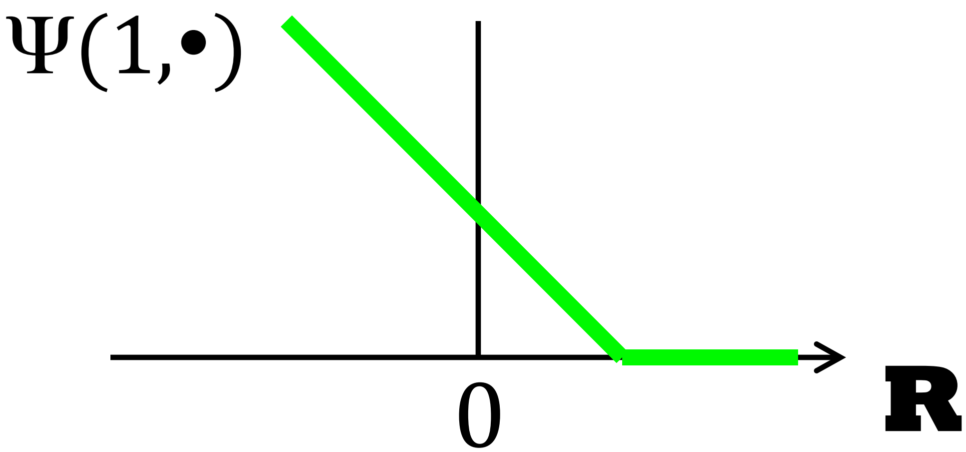
	−1	+1
−1	0	1
+1	1	0



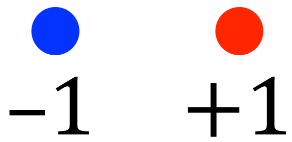
Surrogate Risk Minimization



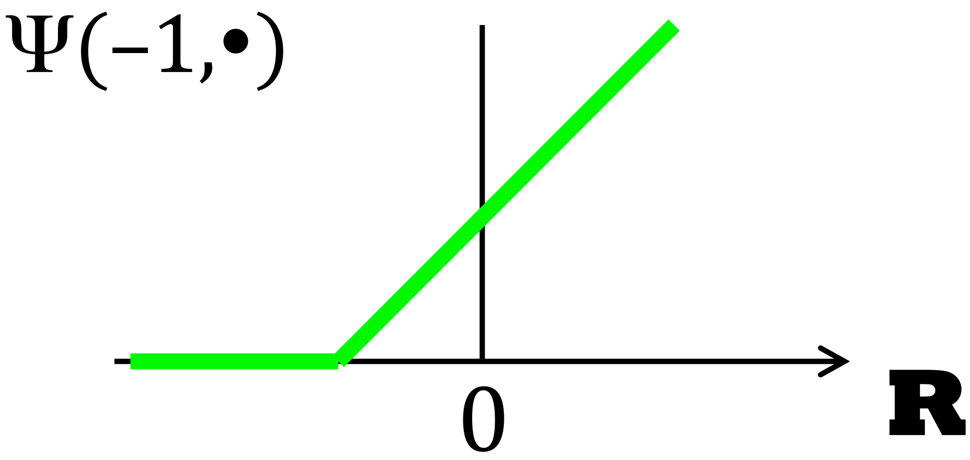
	−1	+1
−1	0	1
+1	1	0



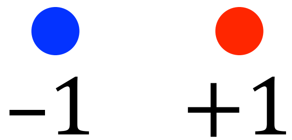
Surrogate Risk Minimization



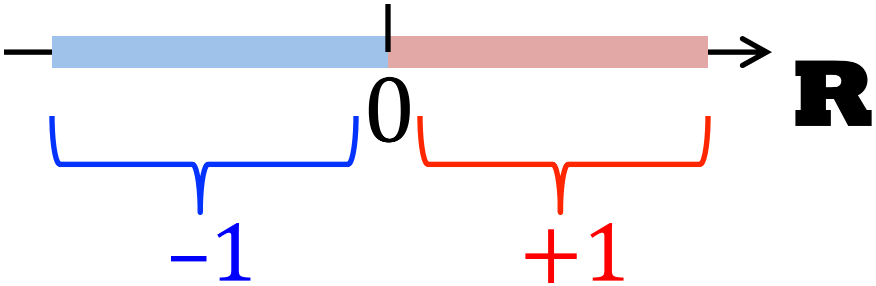
	–1	+1
–1	0	1
+1	1	0



Surrogate Risk Minimization



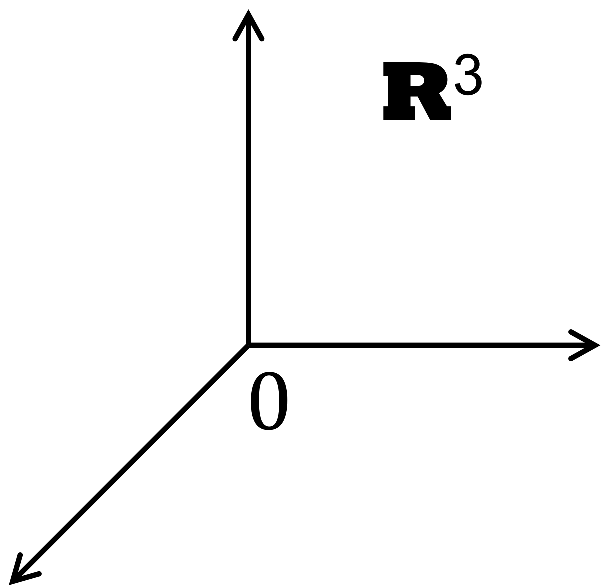
	–1	+1
–1	0	1
+1	1	0



predict according to sign

Surrogate Risk Minimization

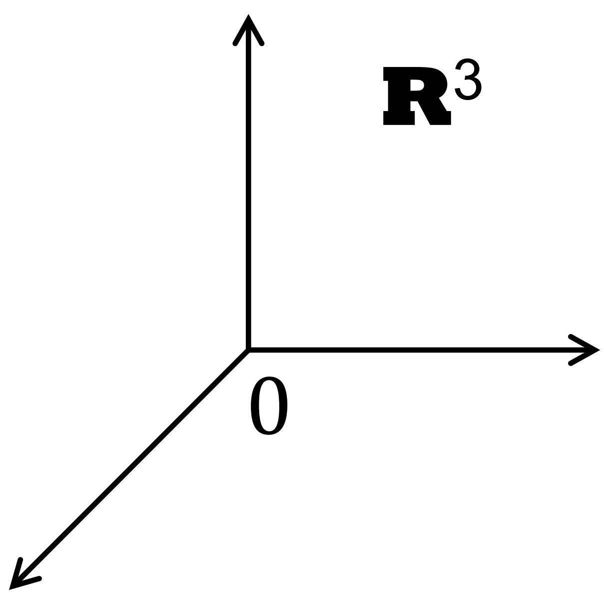
	1	2	3
1	0	1	1
2	1	0	1
3	1	1	0



Surrogate Risk Minimization

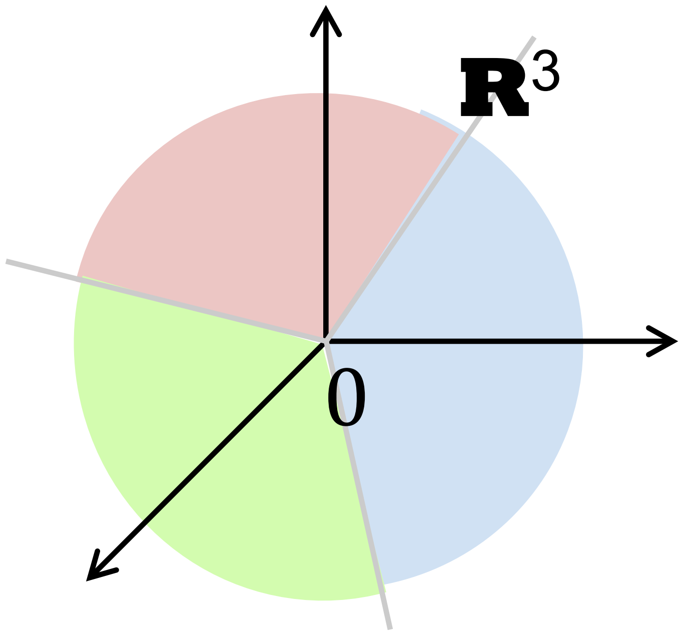
	1	2	3
1	0	1	1
2	1	0	1
3	1	1	0

$\Psi(1,\bullet)$
$\Psi(2,\bullet)$
$\Psi(3,\bullet)$



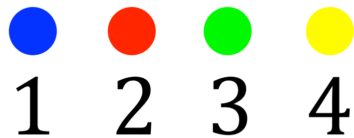
Surrogate Risk Minimization

	1	2	3
1	0	1	1
2	1	0	1
3	1	1	0



predict according to argmax

Surrogate Risk Minimization



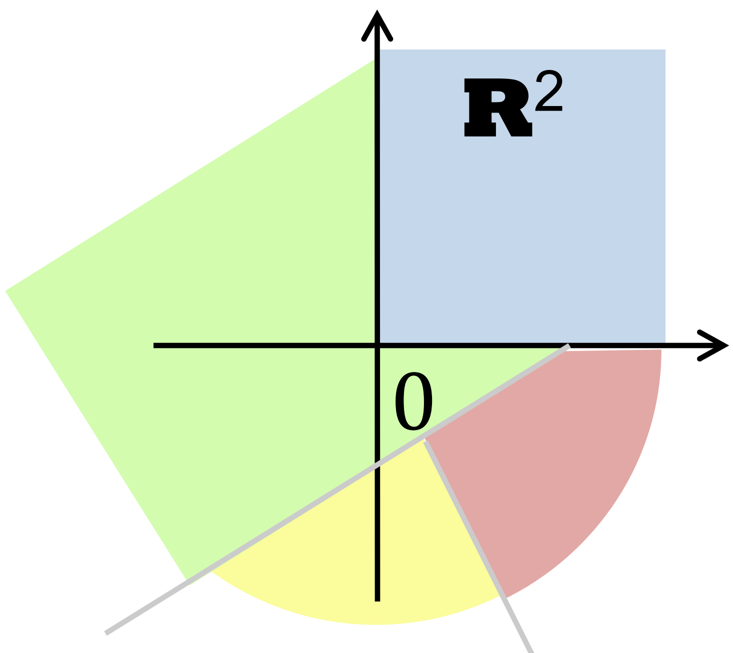
	1	2	3	4
1	0	1	4	2
2	2	0	5	2
3	1	7	0	3
4	2	2	2	0

$\Psi(1,\bullet)$

$\Psi(2,\bullet)$

$\Psi(3,\bullet)$

$\Psi(4,\bullet)$



Surrogate Risk Minimization

	1	2	3	4
1	0	1	4	2

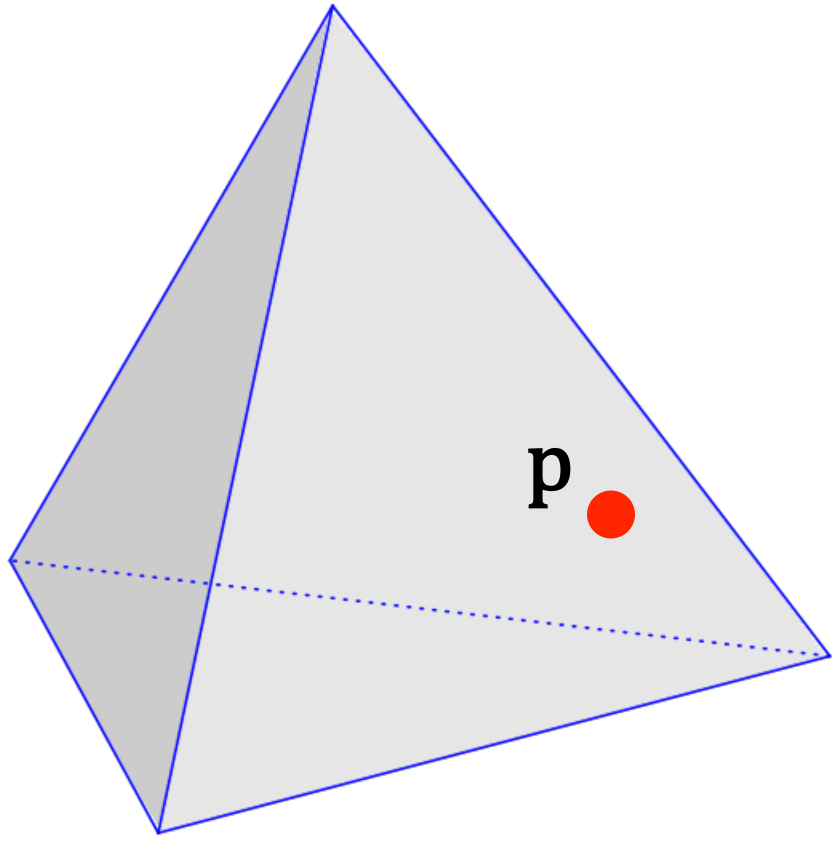
Surrogate Ψ is calibrated for loss **L** if $\exists$ pred s.t.

$$\forall \mathbf{p} \in \Delta_n : \inf_{\mathbf{u}:\mathrm{pred}(\mathbf{u}) \notin \arg\min_t p'\ell_t} E_{Y\sim\mathbf{p}}[\Psi(Y,\mathbf{u})] > \inf_{\mathbf{u}} E_{Y\sim\mathbf{p}}[\Psi(Y,\mathbf{u})]$$

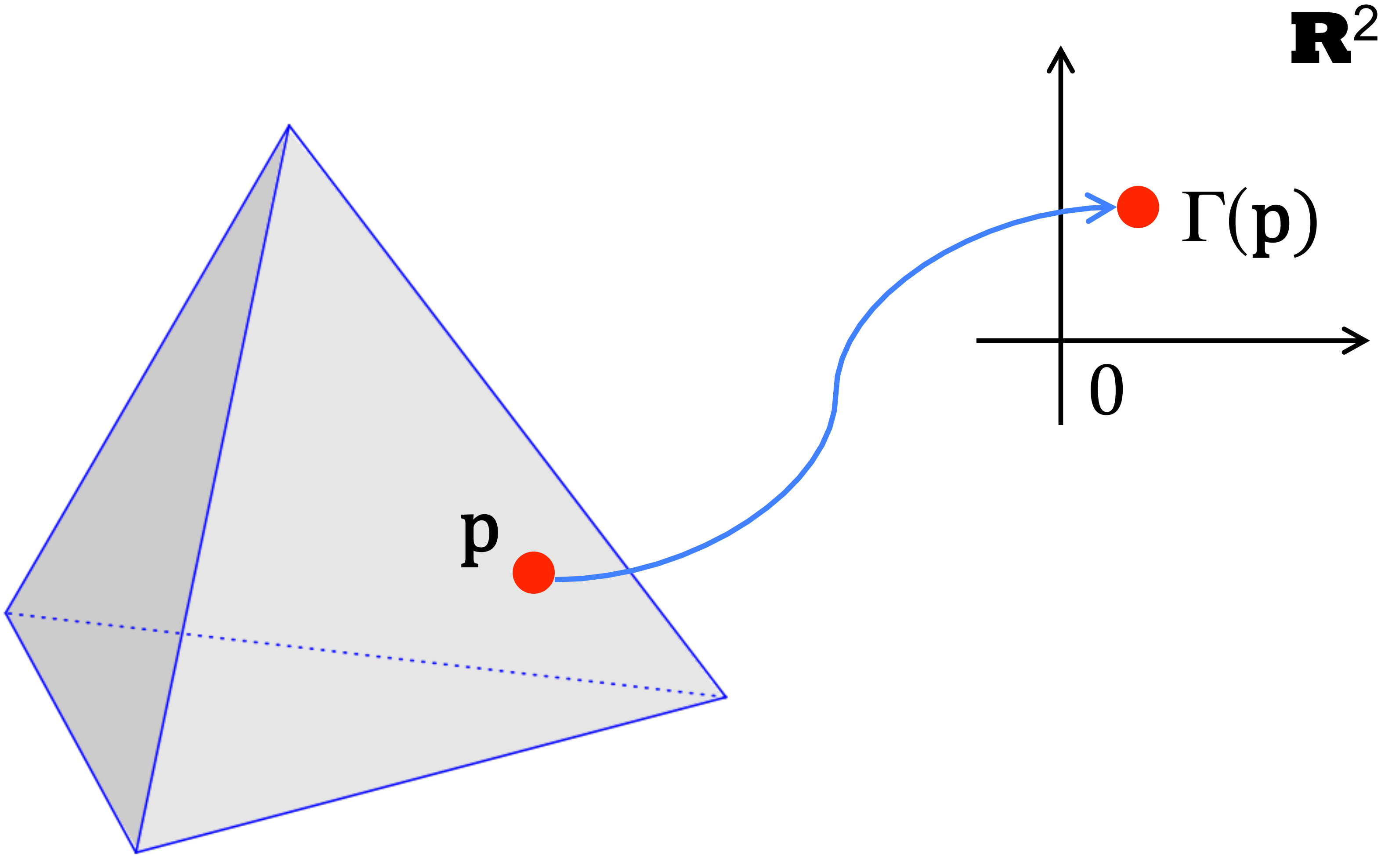
Calibrated Surrogates

- Bartlett et al, 2003, 2006
- Zhang, 2004
- Tewari & Bartlett, 2005, 2007
- Steinwart, 2007
- Duchi et al, 2010
- Ramaswamy & Agarwal, 2012, 2014

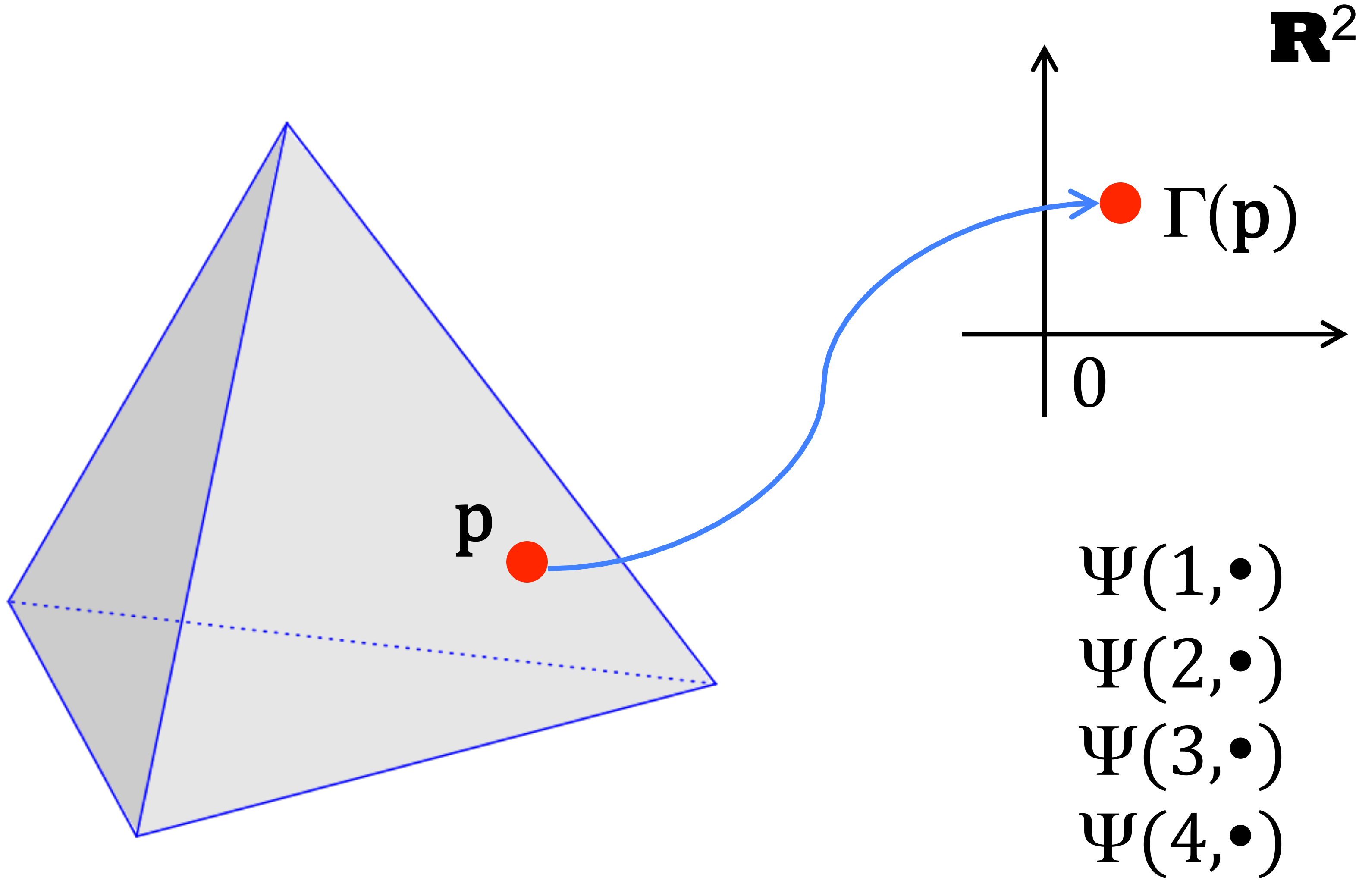
Property Elicitation



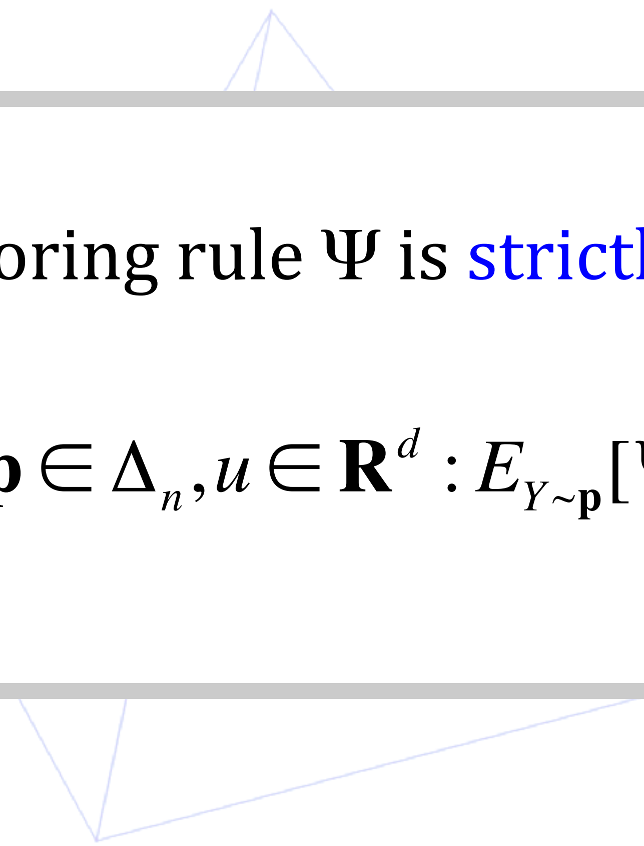
Property Elicitation

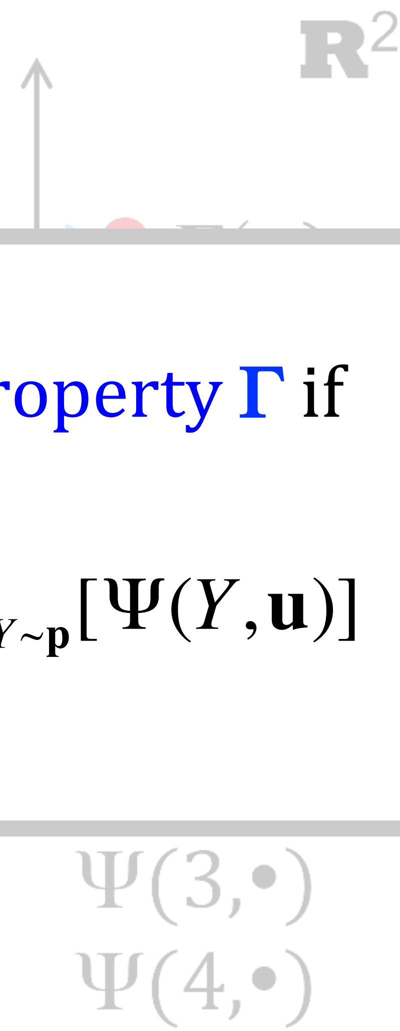


Property Elicitation



Property Elicitation





Scoring rule Ψ is strictly proper for property Γ if

$$\forall \mathbf{p} \in \Delta_n, u \in \mathbf{R}^d : E_{Y\sim\mathbf{p}}[\Psi(Y,\Gamma(\mathbf{p}))] < E_{Y\sim\mathbf{p}}[\Psi(Y,\mathbf{u})]$$

Strictly Proper Scoring Rules for Property Elicitation

- Lambert et al, 2008
- Lambert & Shoham, 2009
- Abernethy & Frongillo, 2012
- Steinwart et al, 2014
- Frongillo & Kash, 2015

Our Paper

‘L-Calibrated’ properties

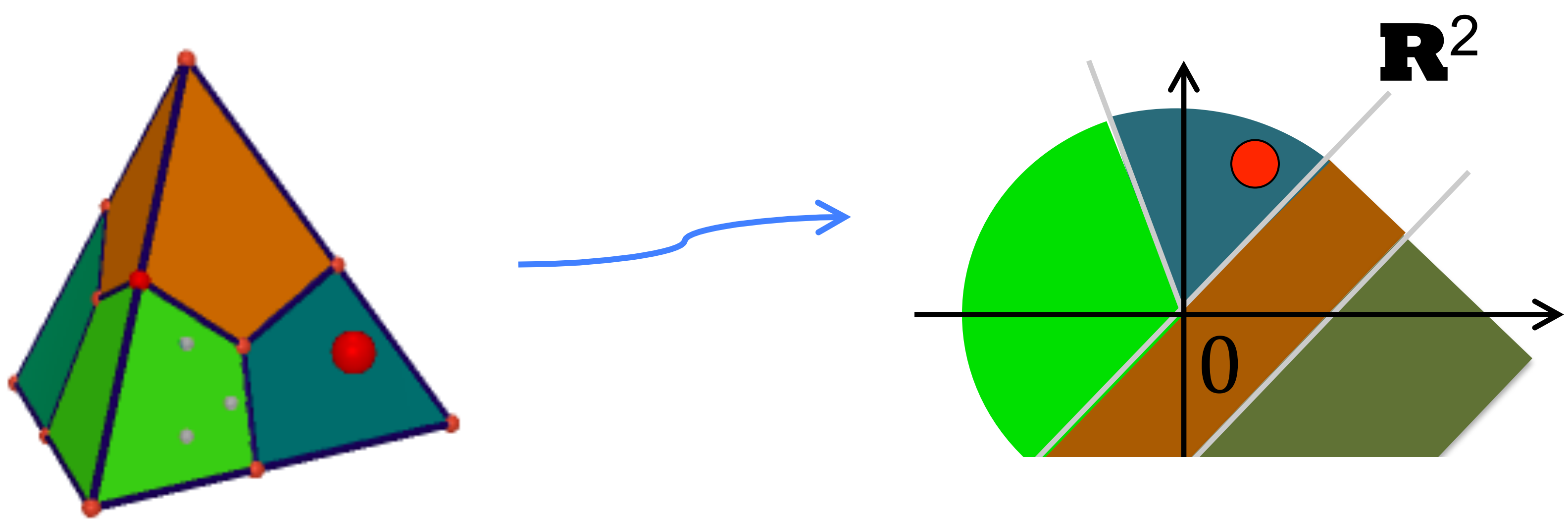
Our Paper

'L-Calibrated' properties

Trigger probability sets of **L**
(Ramaswamy & Agarwal, 2012, 2014)

Our Paper

'L-Calibrated' properties



Trigger probability sets of **L**
(Ramaswamy & Agarwal, 2012, 2014)

Our Paper

Strictly proper scoring rules for
'L-Calibrated' properties

⇩

L-Calibrated surrogates

Calibrated Surrogates via Linear Properties

- ‘Standardization function’ based ranking-calibrated surrogates of Buffoni et al (2010) elicit linear properties!

Calibrated Surrogates via Linear Properties

- ‘Standardization function’ based ranking-calibrated surrogates of Buffoni et al (2010) elicit linear properties!

- Generic calibrated surrogates of Ramaswamy et al (2013) elicit linear properties!

Calibrated Surrogates via Linear Properties

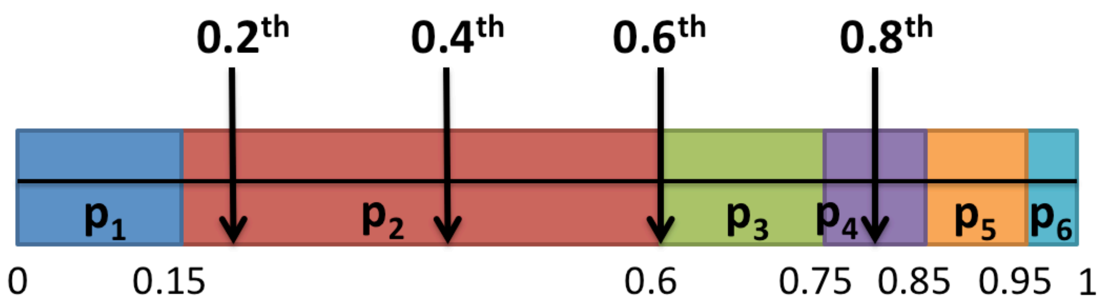
- ‘Standardization function’ based ranking-calibrated surrogates of Buffoni et al (2010) elicit linear properties!

- Generic calibrated surrogates of Ramaswamy et al (2013) elicit linear properties!

- Lower bound on dimension of any **L**-calibrated linear property

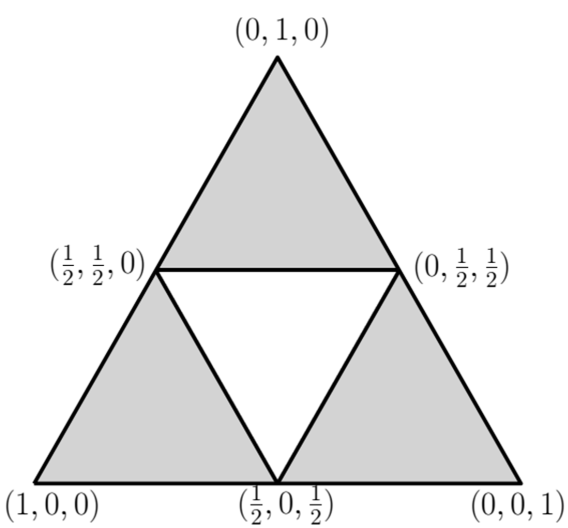
Calibrated Surrogates via Nonlinear Properties

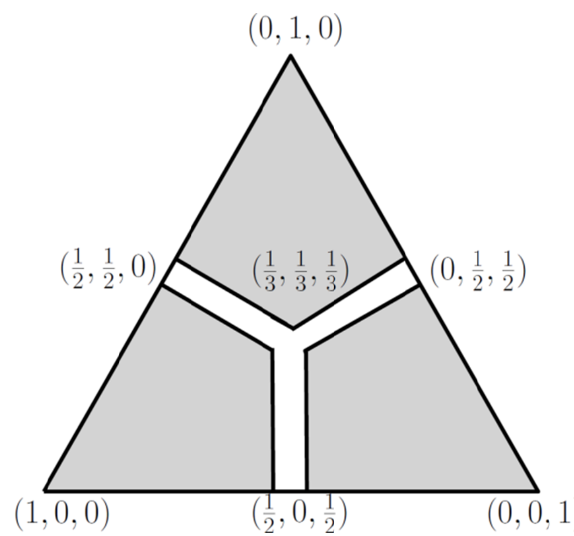
- Vectors of quantiles give 'coarse' estimates of conditional label probability; yield surrogates that are calibrated under low-noise conditions



Calibrated Surrogates via Nonlinear Properties

- Vectors of quantiles give 'coarse' estimates of conditional label probability; yield surrogates that are calibrated under low-noise conditions





Poster Session: 4:00-5:45 PM

Arpit Agarwal