Parameter-Invariant Monitor Design for Cyber-Physical Systems
Part 1: Fundamentals of Parameter-Invariance

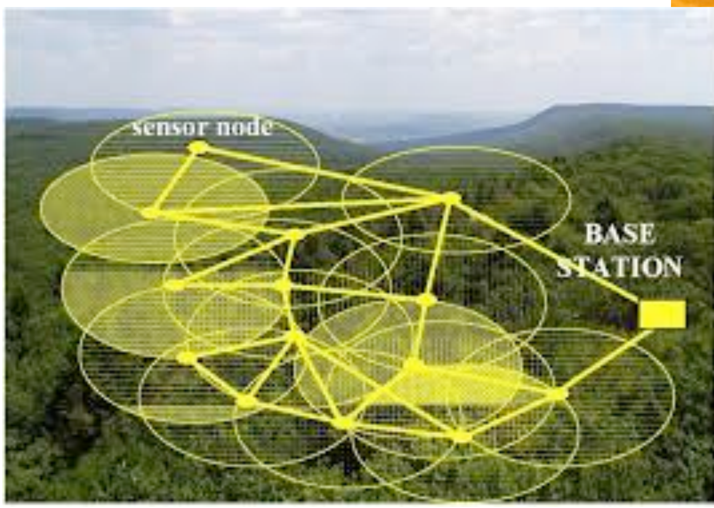
James Weimer, Oleg Sokolsky, Insup Lee

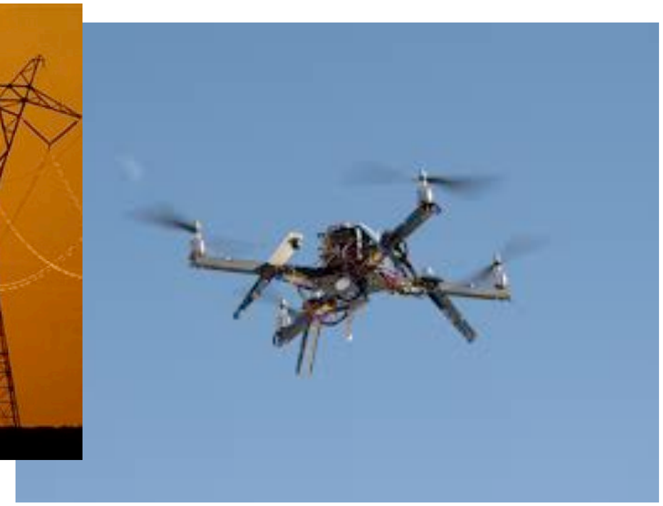
Penn Engineering

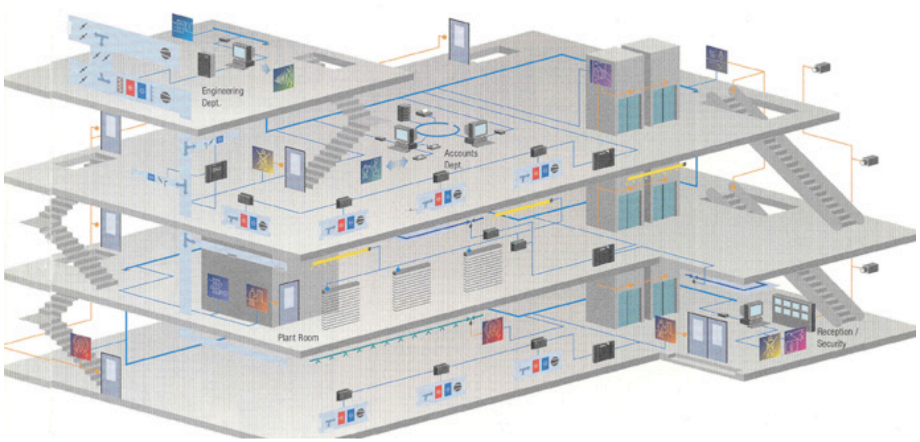
PRECISE

PENN RESEARCH IN EMBEDDED COMPUTING AND INTEGRATED SYSTEMS ENGINEERING

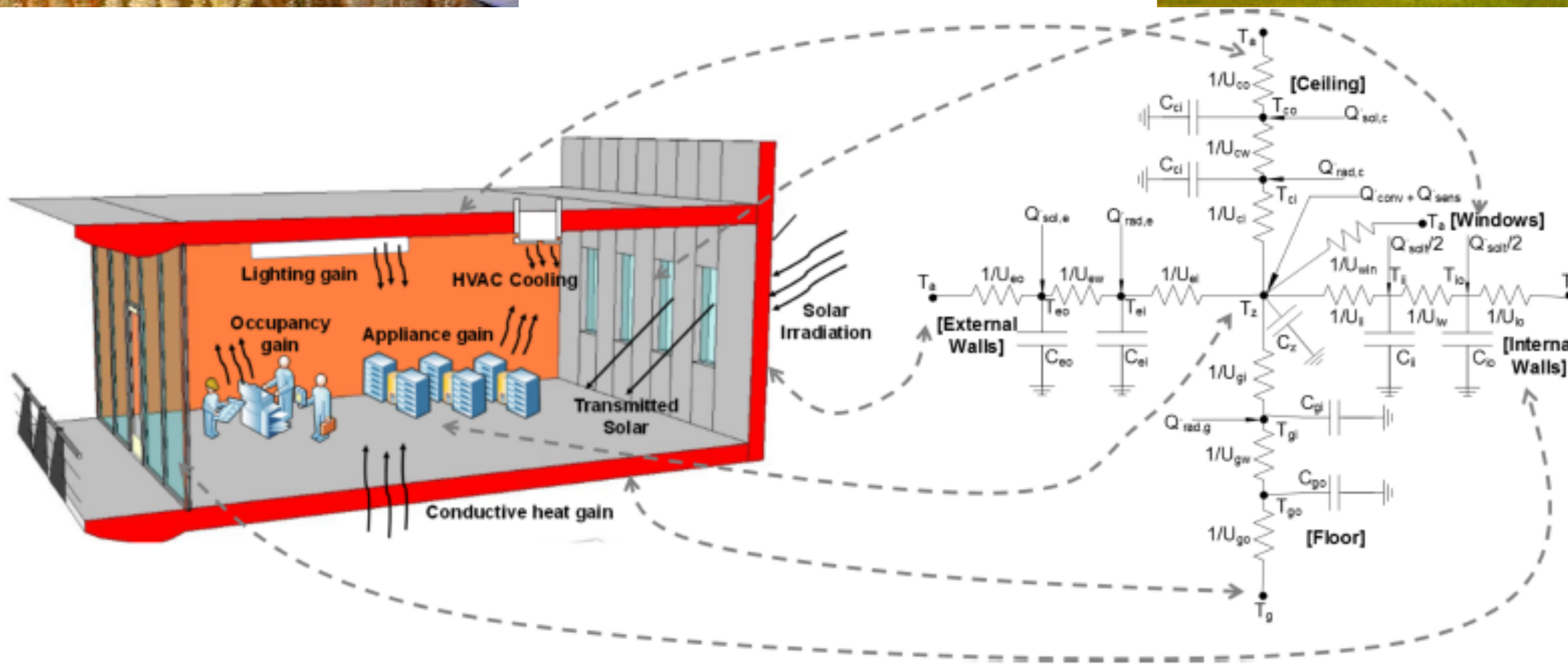
Cyber Physical Systems are Everywhere







Building Energy Management

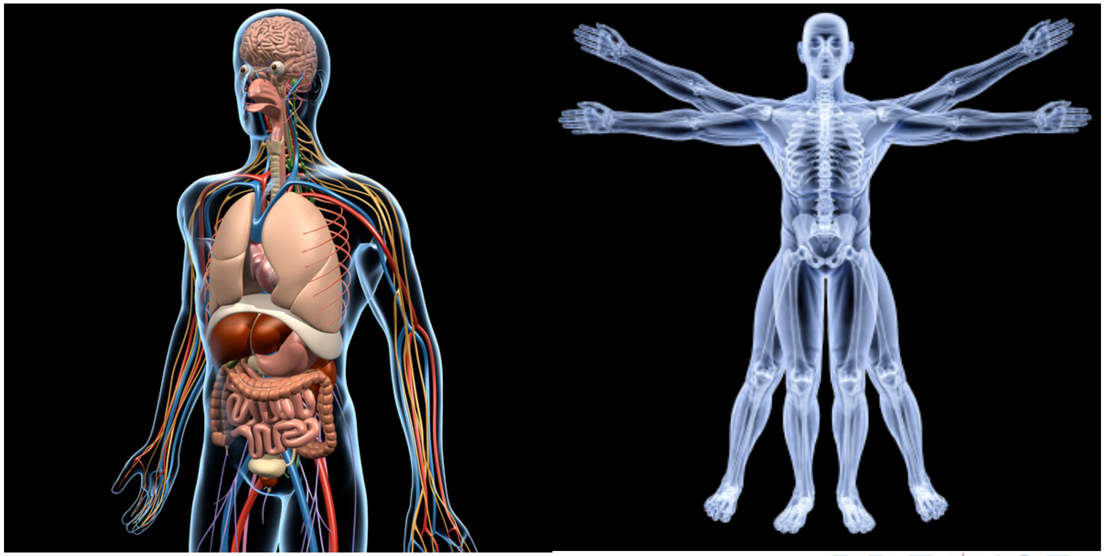


Smart Grid Monitoring

Penn Engineering

PRECISE
PENN RESEARCH IN EMBEDDED COMPUTING AND INTEGRATED SYSTEMS ENGINEERING

Health Monitoring



Outline of Tutorial

- Part 1 : Foundations of Parameter Invariance
  - underlying mathematics
  - motivation
  - **supplemental slides : https://rtg.cis.upenn.edu/parameter-invariant.html**
- Part 2 : Design of Parameter Invariant Monitors
  - parameter invariant testing for linear time invariant dynamics
  - design short cuts
- Part 3 : Implementation of Parameter Invariant Monitors
  - real-world applications
  - design trade-off insight

Outline of Tutorial

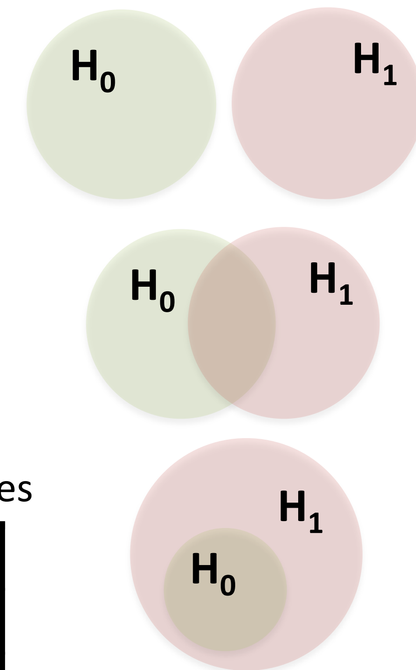
- Part 1 : Foundations of Parameter Invariance
  - underlying mathematics
  - motivation
  - **supplemental slides : https://rtg.cis.upenn.edu/parameter-invariant.html**
- Part 2 : Design of Parameter Invariant Monitors
  - parameter invariant testing for linear time invariant dynamics
  - design short cuts
- Part 3 : Implementation of Parameter Invariant Monitors
  - real-world applications
  - design trade-off insight

Binary Hypothesis Testing Basics

- Two hypotheses: null hypothesis (H_0) and event hypothesis (H_1)
  - H_0 = normal/safe, H_1 = abnormal/unsafe
  - H_0 = absence of X , H_1 = presence of X
  - H_0 = X happens at T_0, H_1 = X happens at T_1

- Hypotheses are classified as “simple” or “composite”
  - simple : hypothesis contains 1 scenario
    - e.g. H_0 = X happens at time 1
  - composite : hypothesis contains multiple (possibly infinite) scenarios
    - e.g. H_0 = X happens at time 1 or time 2

composite hypotheses



- Performance of a test classified “false positive” and “true positive” rates

	H_0 is true	H_1 is true
test claims H_0	correct non-detection	missed detection
test claims H_1	false positive	true positive

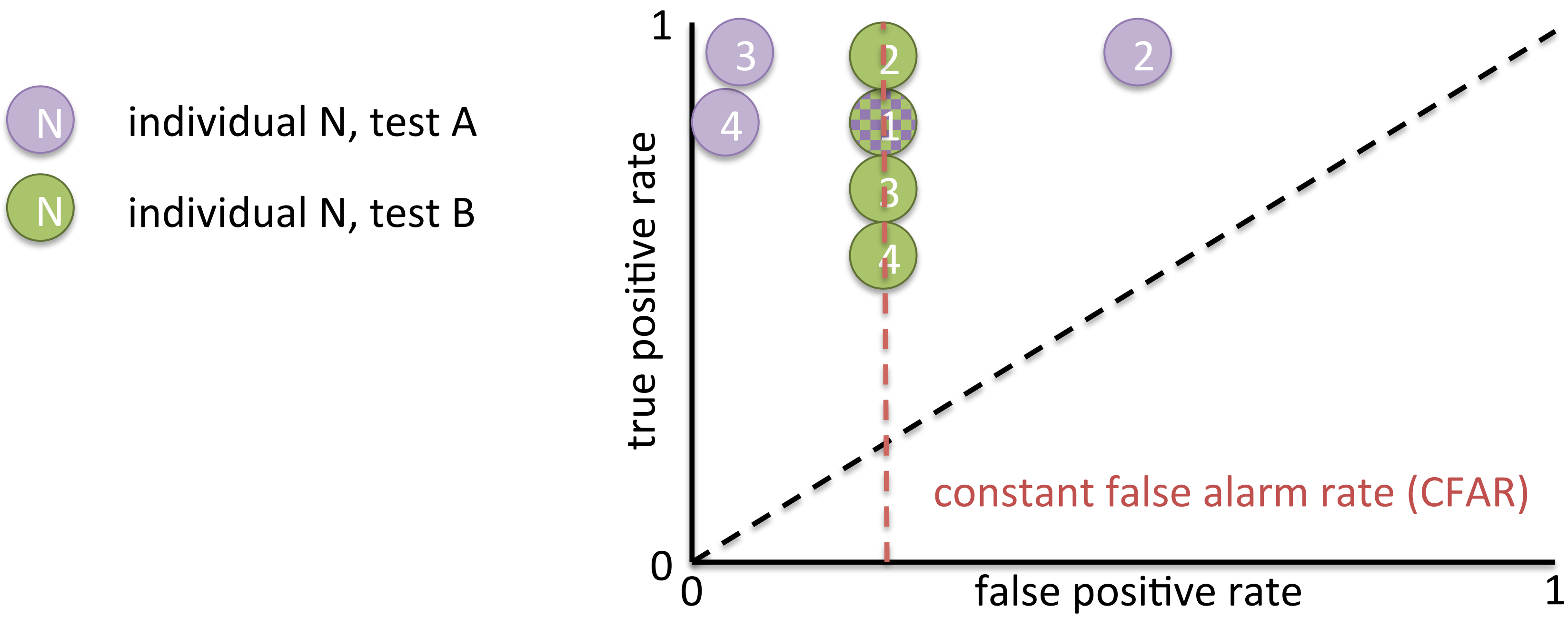
- Two Classical Approaches to test design
  - **Bayesian**: assumes prior knowledge on probability of H_0 and H_1 being true
    - easier to design, “absolute” performance
  - **Frequentist**: does not assume prior knowledge on probability **(many CPS applications)**
    - harder to design, “relative” performance

Robust Monitoring Problem

“All models are wrong, but some are useful” – George E.P. Box

Known
- event we want to monitor
- basic physical interactions

Unknown
- probability of event
- exact physical models



Monitor Design: *maximize worst-case true positive rate, bound worst-case false positive rate*

Robust Monitoring Problem … More Formally

- $\mathcal{V} = \{1, \ldots, V\}$ is a population of V individuals.
- For each individual, $v \in \mathcal{V}$, we gather N measurements, $\mathbf{y}_v \in \mathbb{R}^N$.
- The PDF of $\mathbf{y}_v$ is parameterized by $\theta_v \in \Theta_0 \cup \Theta_1$, $\mathbf{y}_v \sim f(y \mid \theta_v)$

- We wish to design a test, $\hat{\phi}(\mathbf{y}_v) \in \{0, 1\}$, to evaluate:

$$\mathcal{H}_0 : \theta_v \in \Theta_0 \quad \text{vs.} \quad \mathcal{H}_1 : \theta_v \in \Theta_1$$

  – maximum probability of false positive is bounded:

$$\forall \theta_v \in \Theta_0,\ P_{\theta_v}[\hat{\phi}(Y_v) = 1] \le \sup_{v \in \mathcal{V},\ \theta \in \Theta_0} P_\theta[\hat{\phi}(Y_v) = 1] = \sup_{v \in \mathcal{V},\ \theta \in \Theta_0} E_\theta[\hat{\phi}(Y_v)] = P_{FP}(\hat{\phi})$$

  – minimum probability of true positive is maximized:

$$\forall \theta_v \in \Theta_1,\ P_{\theta_v}[\hat{\phi}(Y_v) = 1] \ge \inf_{v \in \mathcal{V},\ \theta \in \Theta_1} P_\theta[\hat{\phi}(Y_v) = 1] = \inf_{v \in \mathcal{V},\ \theta \in \Theta_1} E_\theta[\hat{\phi}(Y_v)] = P_{TP}(\hat{\phi})$$

- Monitor design problem:

$$\hat{\phi} = \arg\max_{\phi \in \Phi_\alpha} P_{TP}(\phi) \qquad \Phi_\alpha = \left\{ \hat{\phi} \mid P_{FP}(\hat{\phi}) \le \alpha \right\}$$

maximize true positive **bounded false positive**

Overview of Testing Approaches

- Neyman-Pearson (NP)
  - likelihood ratio test (LRT)
  - uniformly most powerful (UMP) Test
- Maximum Likelihood (ML)
  - generalized likelihood ratio test (GLRT)
- Maximal Invariance (MI)
  - maximally invariant statistic
  - uniformly most powerful invariant (UMPI) test
- Parameter-Invariance (PAIN)
  - near-maximally invariant statistic

Form of Illustrative Examples

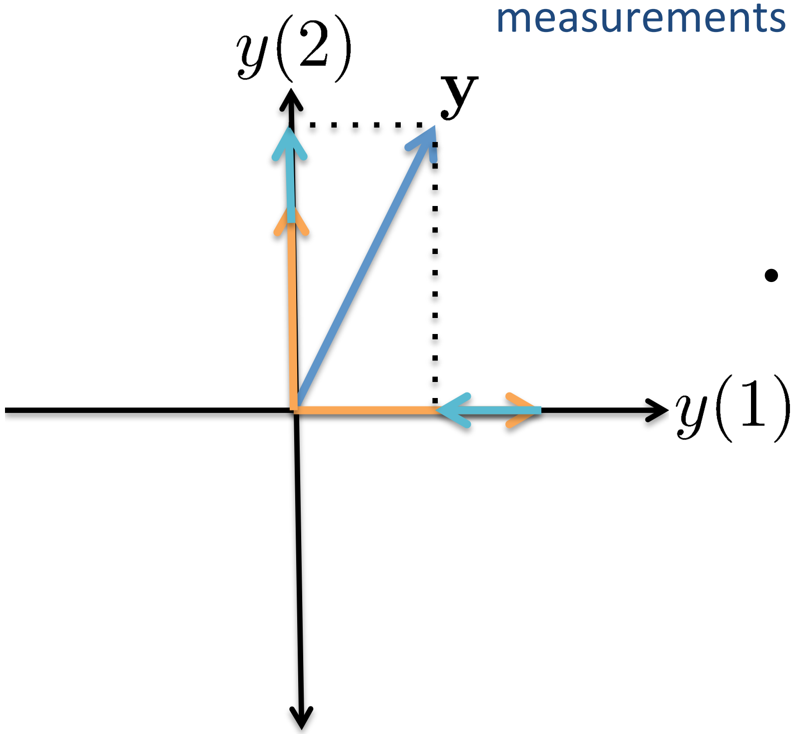
- 2 measurements, 2 parameters, and noise

$$\begin{bmatrix} y(1) \\ y(2) \end{bmatrix} = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} + \begin{bmatrix} n(1) \\ n(2) \end{bmatrix}$$

measurements = parameters + noise

$$n(k) \sim \mathrm{N}[0, 1]$$

i.i.d. Gaussian



- Binary hypothesis testing problem on parameters

$$\mathcal{H}_0 : (\theta_1, \theta_2) \in \Theta_{0,1} \times \Theta_{0,2}$$
$$\mathcal{H}_1 : (\theta_1, \theta_2) \in \Theta_{1,1} \times \Theta_{1,2}$$

3 Examples will be considered

Penn Engineering

PRECISE

Illustrative Examples Considered

$$\begin{bmatrix} y(1) \\ y(2) \end{bmatrix} = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} + \begin{bmatrix} n(1) \\ n(2) \end{bmatrix}$$

example 1

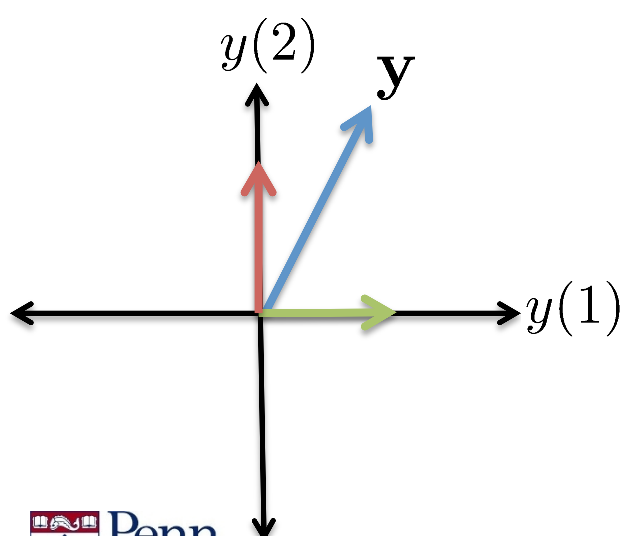
$$\mathcal{H}_0 : (\theta_1, \theta_2) \in \{1\} \times \{0\}$$
$$\mathcal{H}_1 : (\theta_1, \theta_2) \in \{0\} \times \{1\}$$

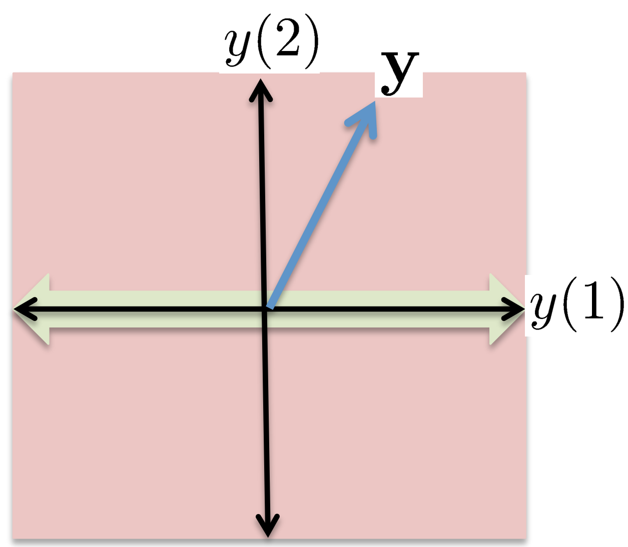
example 2

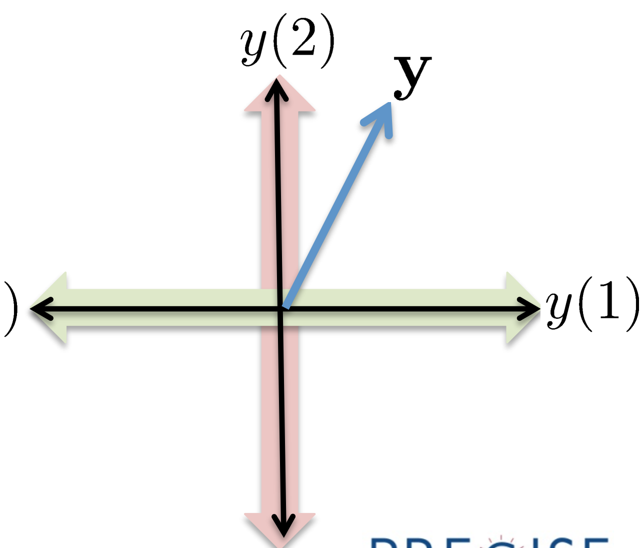
$$\mathcal{H}_0 : (\theta_1, \theta_2) \in \mathbb{R} \times \{0\}$$
$$\mathcal{H}_1 : (\theta_1, \theta_2) \in \mathbb{R} \times \mathbb{R}$$

example 3

$$\mathcal{H}_0 : (\theta_1, \theta_2) \in \mathbb{R} \times \{0\}$$
$$\mathcal{H}_1 : (\theta_1, \theta_2) \in \{0\} \times \mathbb{R}$$

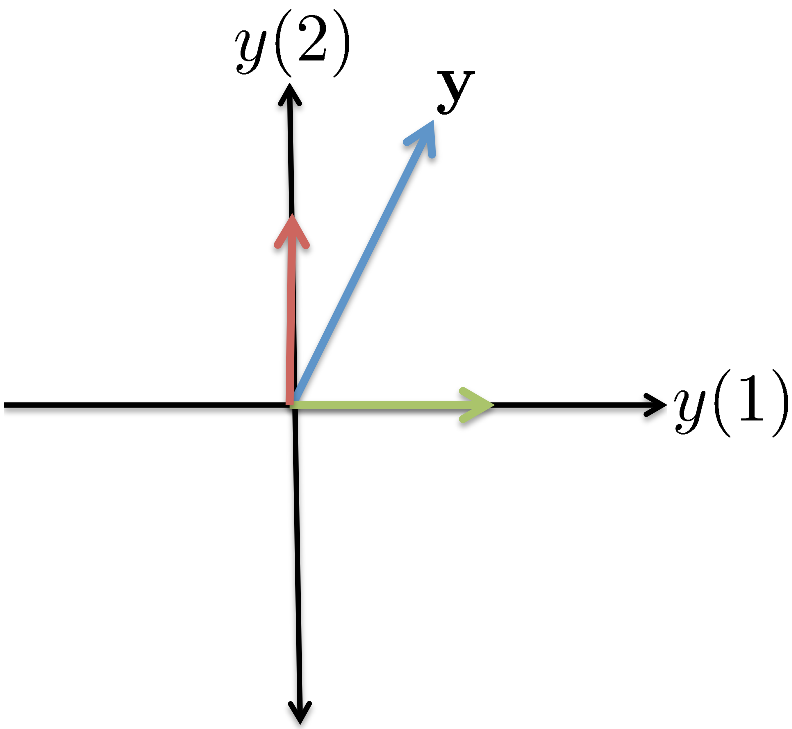






Example 1

$$\mathcal{H}_0 : (\theta_1, \theta_2) \in \{1\} \times \{0\}$$
$$\mathcal{H}_1 : (\theta_1, \theta_2) \in \{0\} \times \{1\}$$



What is the "best" monitor?

Likelihood Ratio Tests

- Example 1 has *simple* hypotheses
  - $\mathcal{H}_0 : (\theta_1, \theta_2) \in \{1\} \times \{0\} \;\; \text{vs.} \;\; \mathcal{H}_1 : (\theta_1, \theta_2) \in \{0\} \times \{1\}$

- Neyman-Pearson Lemma [1933]: (for simple hypotheses)
  - Likelihood ratio test (LRT) maximizes true positive rate at all false positive rates

- Likelihood Ratio Test (LRT)
  - likelihood of parameters, given measurements: $L(\theta; \mathbf{y}_v)$
  - likelihood ratio: $L_R(\mathbf{y}_v) = \frac{L(\theta_1; \mathbf{y}_v)}{L(\theta_0; \mathbf{y}_v)}$
  - $LRT(\phi) \longleftrightarrow \phi(\mathbf{y}_v) = \begin{cases} 1, & L_R(\mathbf{y}_v) > \eta \\ 0, & L_R(\mathbf{y}_v) \leq \eta \end{cases}$
    - where $P[L_R(\mathbf{y}_v) \leq \eta] = \alpha$

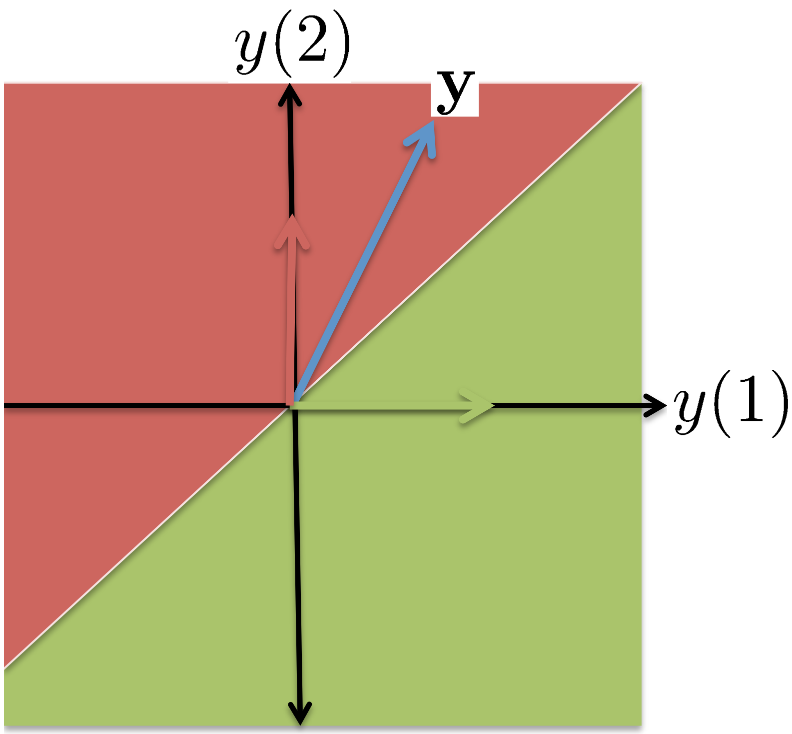
Penn Engineering

PRECISE

Example 1 : LRT

$\mathcal{H}_0 : (\theta_1, \theta_2) \in \{1\} \times \{0\}$

$\mathcal{H}_1 : (\theta_1, \theta_2) \in \{0\} \times \{1\}$



apply the likelihood ratio test

log-likelihood ratio

$$\ln L_R(\mathbf{y}) = \ln L(\Theta_1; \mathbf{y}) - \ln L(\Theta_0; \mathbf{y})$$
$$= \mathbf{y}(2) - \mathbf{y}(1)$$ **(supplemental slides)**

likelihood ratio test (LRT)

$$\phi(\mathbf{y}) = \begin{cases} 1, & \mathbf{y}(2) > \mathbf{y}(1) + \eta \\ 0, & \mathbf{y}(2) \leq \mathbf{y}(1) + \eta \end{cases}$$

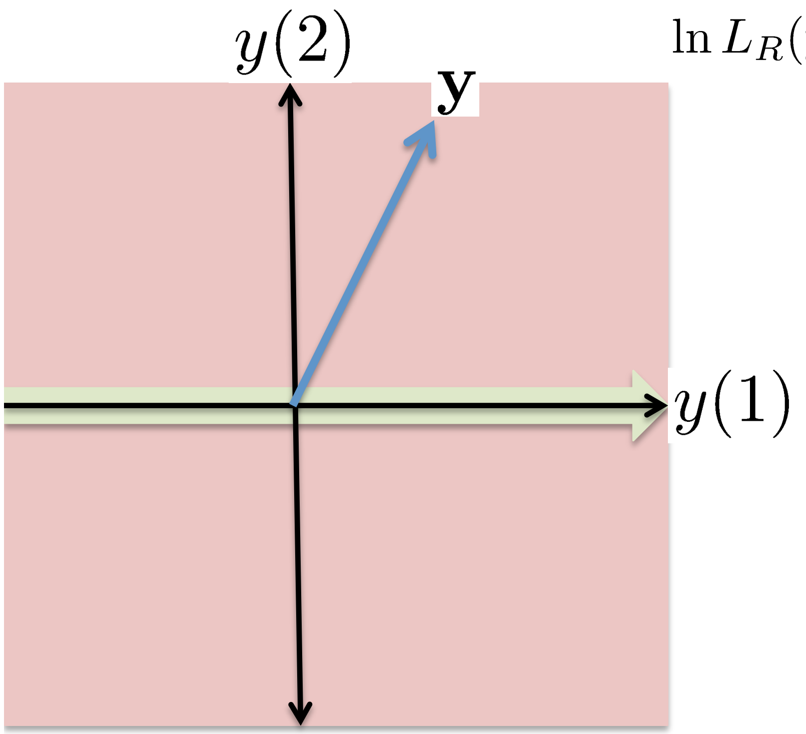
η chosen to achieve desired false positive rate

- **LRT is optimal for simple binary hypotheses.**
- **What about composite hypotheses?**

Example 2 : LRT

$$\mathcal{H}_0 : (\theta_1, \theta_2) \in \mathbb{R} \times \{0\}$$
$$\mathcal{H}_1 : (\theta_1, \theta_2) \in \mathbb{R} \times \mathbb{R}$$

~~**apply the likelihood ratio test**~~



log-likelihood ratio

$$\begin{aligned}\ln L_R(\mathbf{y}) &= \ln L\left(\Theta_1; \mathbf{y}\right) - \ln L\left(\Theta_0; \mathbf{y}\right) \\ &= \frac{1}{2}\left(\sum_{k=1}^{2}(y(k) - \Theta_{0,k})^2 - \sum_{k=1}^{2}(y(k) - \Theta_{1,k})^2\right) \\ &= ???\end{aligned}$$

LRT is not applicable, what is the best monitor?

Generalized Likelihood Ratio Test

- Example 2: H_0 is *composite*, H_1 is *composite*
  - $\mathcal{H}_0 : (\theta_1, \theta_2) \in \mathbb{R} \times \{0\} \text{ vs. } \mathcal{H}_1 : (\theta_1, \theta_2) \in \mathbb{R} \times \mathbb{R}$

- Neyman-Pearson Lemma [1933]: (H_0 is *simple*, H_1 is *composite)*
  - The LRT can be the uniformly most powerful (UMP) test
    - *"no other test with the same false positive rate has greater true positive rate"*
    - results in an optimal robust monitor – if UMP test exists

- Concept: estimate unknown parameters in LRT
  - use ratio of maximum likelihood under each hypothesis

- Generalized Likelihood Ratio Test (GLRT)
  - generalized likelihood ratio: $\hat{L}_R(\mathbf{y}) = \frac{\max_{\theta_1 \in \Theta_1} L(\theta_1; \mathbf{y})}{\max_{\theta_0 \in \Theta_0} L(\theta_0; \mathbf{y})}$
  - $GLRT(\phi) \longleftrightarrow \phi(\mathbf{y}) = \begin{cases} 1, & \hat{L}_R(\mathbf{y}) > \eta \\ 0, & \hat{L}_R(\mathbf{y}) \leq \eta \end{cases}$

Penn Engineering

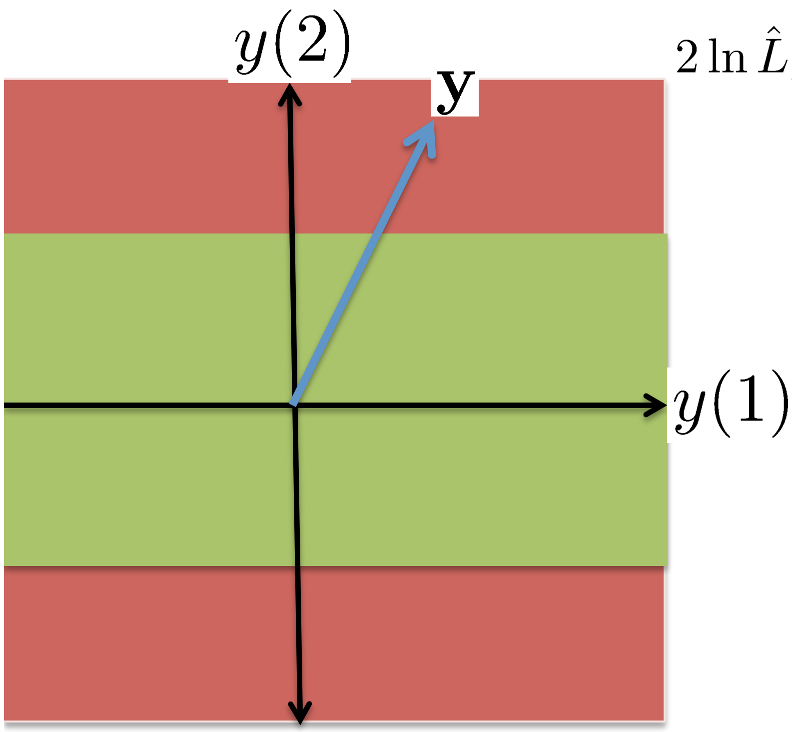
PRECISE

Example 2 : GLRT

$\mathcal{H}_0 : (\theta_1, \theta_2) \in \mathbb{R} \times \{0\}$

$\mathcal{H}_1 : (\theta_1, \theta_2) \in \mathbb{R} \times \mathbb{R}$

apply the generalized likelihood ratio test



generalized log-likelihood ratio

$$2 \ln \hat{L}_R(\mathbf{y}) = \max_{\hat{\theta}_1 \in \Theta_1} 2 \ln L\left(\hat{\theta}_1; \mathbf{y}\right) - \max_{\hat{\theta}_0 \in \Theta_0} 2 \ln L\left(\hat{\theta}_0; \mathbf{y}\right)$$

$$= y(2)^2$$

(supplemental slides)

generalized likelihood ratio test (GLRT)

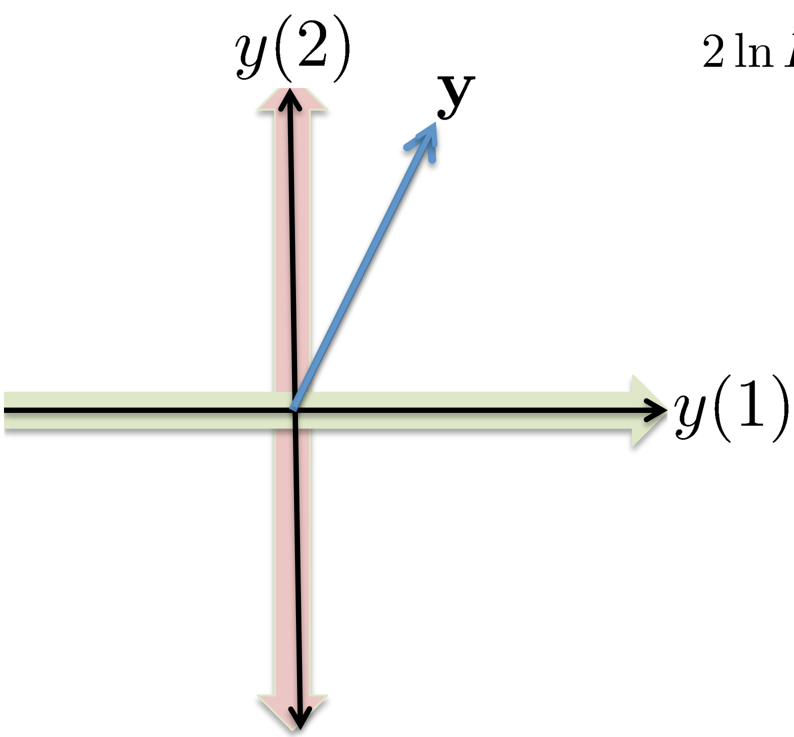
$$\phi(\mathbf{y}) = \begin{cases} 1, & y(2)^2 > \eta \\ 0, & y(2)^2 \leq \eta \end{cases}$$

GLRT is optimal in this example.
Is the GLRT always the solution?

Example 3 : GLRT

$$\mathcal{H}_0 : (\theta_1, \theta_2) \in \mathbb{R} \times \{0\}$$
$$\mathcal{H}_1 : (\theta_1, \theta_2) \in \{0\} \times \mathbb{R}$$

apply the generalized likelihood ratio test



generalized log-likelihood ratio

$$2 \ln \hat{L}_R(\mathbf{y}) = \max_{\hat{\theta}_1 \in \Theta_1} 2 \ln L\left(\hat{\theta}_1; \mathbf{y}\right) - \max_{\hat{\theta}_0 \in \Theta_0} 2 \ln L\left(\hat{\theta}_0; \mathbf{y}\right)$$
$$= \mathbf{y}(2)^2 - \mathbf{y}(1)^2$$

(supplemental slides)

distribution of generalized log-likelihood ratio

$$\mathcal{H}_0 : 2 \ln \hat{L}_R(\mathbf{y}_v) \sim \text{???}$$

"difference of unknown non-central chi-squared and central chi-squared"

No bounds on false positive rate

Penn Engineering

PRECISE

The Problem with the GLRT

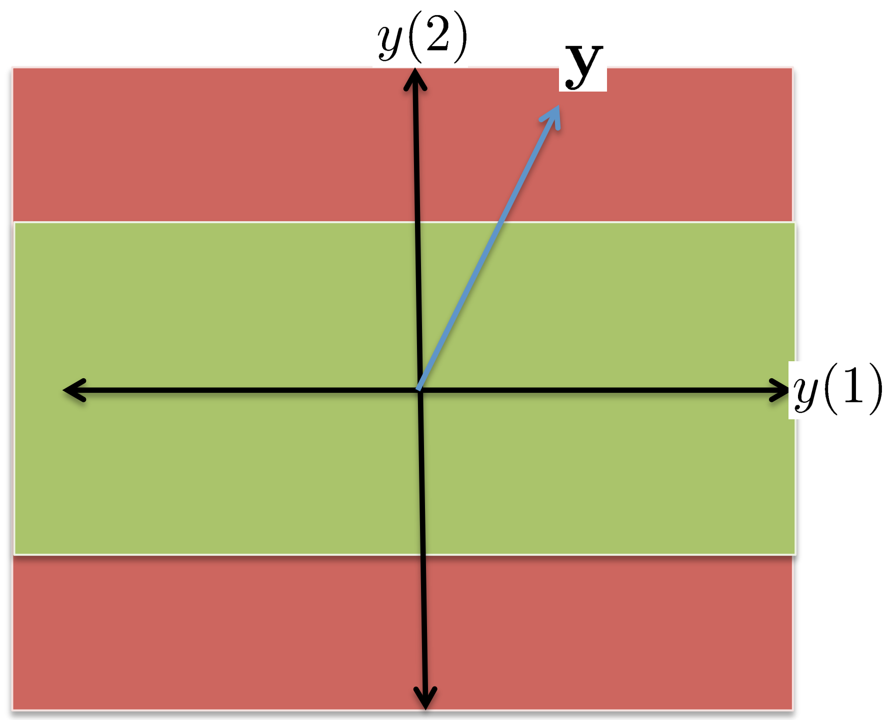
- The GLRT does not (in general) bound the false positive rate
  - Form of Example 2 is a special case [Kraut & Scharf 1999]
  - Example 3 illustrates when it fails.

- Q. What was special about Example 2?

$$\phi(\mathbf{y}) = \begin{cases} 1, & y(2)^2 > \eta \\ 0, & y(2)^2 \leq \eta \end{cases}$$

- A. Invariant to $y(1)$
  - results in known null distribution

- Can this "invariance" be generalized?



Maximally Invariant Test

- Tests over Maximally Invariant (MI) Statistics
  - adaptive approach to signal detection (from radar signal processing)

- Goal: test is maximally (optimally) invariant to unknown *nuisance parameters*
  - nuisance parameters -- do not discriminate b/w hypotheses **(nuisance parameter set)**
  - test parameters = discriminate b/w hypotheses **(test parameter set)**

example 2:

$$\mathcal{H}_0 : (\theta_1, \theta_2) \in \mathbb{R} \times \{0\} \qquad \mathbb{R} \times \{0\} \quad \textbf{nuisance set}$$

$$\mathcal{H}_1 : (\theta_1, \theta_2) \in \mathbb{R} \times \mathbb{R} \qquad \emptyset \times \mathbb{R} \backslash \{0\} \quad \textbf{test set}$$

- Maximally invariant testing is a 2-step approach
  1) design a statistic maximally invariant to nuisance parameters
  2) apply GLRT to MI statistic

- Can yield a uniformly most powerful invariant (UMPI) test
  - *“no other test, **also invariant to the nuisance parameters**, with the same false positive rate has greater true positive rate”*

Penn Engineering

PRECISE

Maximally Invariant Testing … More Formally

- Transformation groups: $\mathcal{G}_\Theta$
  - exactly captures all possible transformations induced by parameters
  - contains inverse transformation:
    - $\forall g \in \mathcal{G}_\Theta, \exists g' \in \mathcal{G}_\Theta,\ g(g'(\mathbf{y})) = \mathbf{y}$
  - group theory mathematics, Sophus Lie 1888 – 1896

- Maximally Invariant statistics to a transformation group:

$$INV(t;\mathcal{G}) \longleftrightarrow \forall g \in \mathcal{G}, t(g(\mathbf{y})) = t(\mathbf{y})$$

$$MI(t;\mathcal{G}) \longleftrightarrow INV(t;\mathcal{G})\ \wedge\ t(\mathbf{y}) = t(\mathbf{y}') \longrightarrow \exists g \in \mathcal{G},\ \mathbf{y}' = g(\mathbf{y})$$

- Consider hypotheses: $\mathcal{H}_k : \theta \in \Theta_k,\ k \in \{0,1\}$
  - nuisance parameter sets: $\hat{\Theta}_N = \Theta_0 \cap \Theta_1$
  - null test parameter sets: $\hat{\Theta}_0 = \Theta_0 \backslash \Theta_1$
  - event test parameter sets: $\hat{\Theta}_1 = \Theta_1 \backslash \Theta_0$

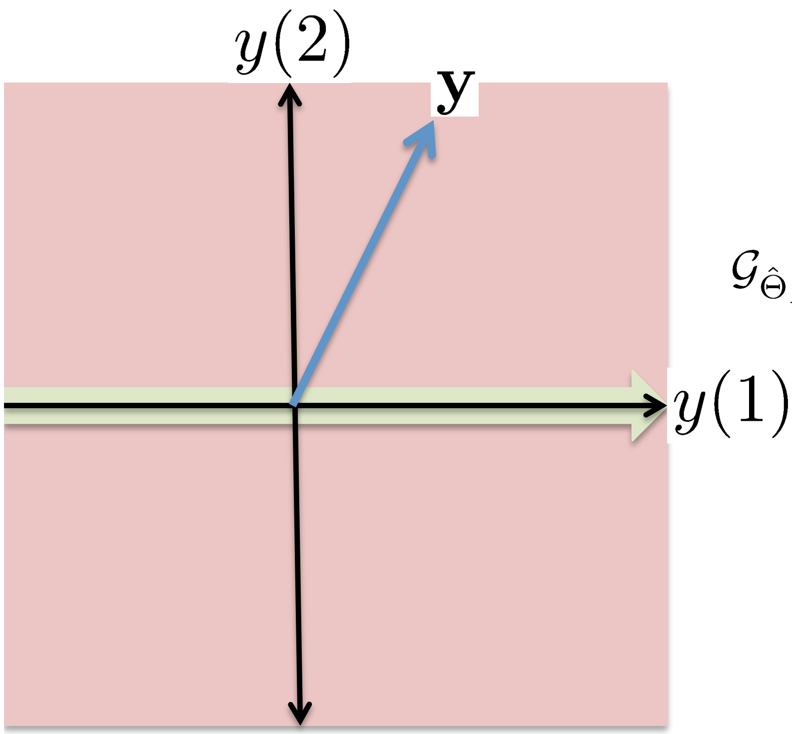
Penn Engineering

PRECISE

PENN RESEARCH IN EMBEDDED COMPUTING AND INTEGRATED SYSTEMS ENGINEERING

Example 2 : MI test

$$\mathcal{H}_0 : (\theta_1, \theta_2) \in \mathbb{R} \times \{0\}$$
$$\mathcal{H}_1 : (\theta_1, \theta_2) \in \mathbb{R} \times \mathbb{R}$$



apply maximally invariant test

nuisance parameters

$$\hat{\Theta}_{N,1} = \mathbb{R}, \hat{\Theta}_{N,2} = 0$$

group of nuisance transformations

$$\mathcal{G}_{\hat{\Theta}_N} = \left\{ g \middle| g\left(\begin{bmatrix} y(1) \\ y(2) \end{bmatrix} \right) = \begin{bmatrix} y(1) \\ y(2) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} c, \ \forall c \in \mathbb{R} \right\}$$

candidate MI statistic

$$t\left(\begin{bmatrix} y(1) \\ y(2) \end{bmatrix} \right) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y(1) \\ y(2) \end{bmatrix} = \begin{bmatrix} 0 \\ y(2) \end{bmatrix}$$

Penn Engineering

PRECISE
PENN RESEARCH IN EMBEDDED COMPUTING AND INTEGRATED SYSTEMS ENGINEERING

Example 2 : MI test

$$\mathcal{H}_0 : (\theta_1, \theta_2) \in \mathbb{R} \times \{0\}$$
$$\mathcal{H}_1 : (\theta_1, \theta_2) \in \mathbb{R} \times \mathbb{R}$$

group of nuisance transformations

$$\mathcal{G}_{\hat{\Theta}_N} = \left\{ g \middle| g\left(\begin{bmatrix} y(1) \\ y(2) \end{bmatrix} \right) = \begin{bmatrix} y(1) \\ y(2) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} c, \ \forall c \in \mathbb{R} \right\}$$

candidate MI statistic

$$t\left(\begin{bmatrix} y(1) \\ y(2) \end{bmatrix} \right) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y(1) \\ y(2) \end{bmatrix} = \begin{bmatrix} 0 \\ y(2) \end{bmatrix}$$

to be invariant

$$INV(t; \mathcal{G}) \longleftrightarrow \forall g \in \mathcal{G}, t(g(\mathbf{y})) = t(\mathbf{y})$$

$$\begin{aligned} t\left(g\left(\begin{bmatrix} y(1) \\ y(2) \end{bmatrix} \right) \right) &= \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} y(1) \\ y(2) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} c \right), \quad c \in \mathbb{R} \\ &= \begin{bmatrix} 0 \\ y(2) \end{bmatrix} \\ &= t\left(\begin{bmatrix} y(1) \\ y(2) \end{bmatrix} \right) \end{aligned}$$

TRUE

Example 2 : MI test

$$\mathcal{H}_0 : (\theta_1, \theta_2) \in \mathbb{R} \times \{0\}$$
$$\mathcal{H}_1 : (\theta_1, \theta_2) \in \mathbb{R} \times \mathbb{R}$$

group of nuisance transformations

$$\mathcal{G}_{\hat{\Theta}_N} = \left\{ g \middle| g\left(\begin{bmatrix} y(1) \\ y(2) \end{bmatrix} \right) = \begin{bmatrix} y(1) \\ y(2) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} c,\ \forall c \in \mathbb{R} \right\}$$

candidate MI statistic

$$t\left(\begin{bmatrix} y(1) \\ y(2) \end{bmatrix} \right) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y(1) \\ y(2) \end{bmatrix} = \begin{bmatrix} 0 \\ y(2) \end{bmatrix}$$

to be maximally invariant

$$MI(t;\mathcal{G}) \longleftrightarrow INV(t;\mathcal{G}) \ \wedge \ t(\mathbf{y}) = t(\mathbf{y}') \longrightarrow \exists g \in \mathcal{G},\ \mathbf{y}' = g(\mathbf{y})$$

TRUE

$$t\left(\begin{bmatrix} y_i(1) \\ y_i(2) \end{bmatrix} \right) = t\left(\begin{bmatrix} y_j(1) \\ y_j(2) \end{bmatrix} \right)$$

$$\longrightarrow \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_i(1) \\ y_j(2) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_j(1) \\ y_j(2) \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} y_i(1) \\ y_j(2) \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} y_i(1) = \begin{bmatrix} y_j(1) \\ y_j(2) \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} y_j(1)$$

$$\longrightarrow \begin{bmatrix} y_i(1) \\ y_j(2) \end{bmatrix} = \begin{bmatrix} y_j(1) \\ y_j(2) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} c, \quad c = y_i(1) - y_j(1)$$

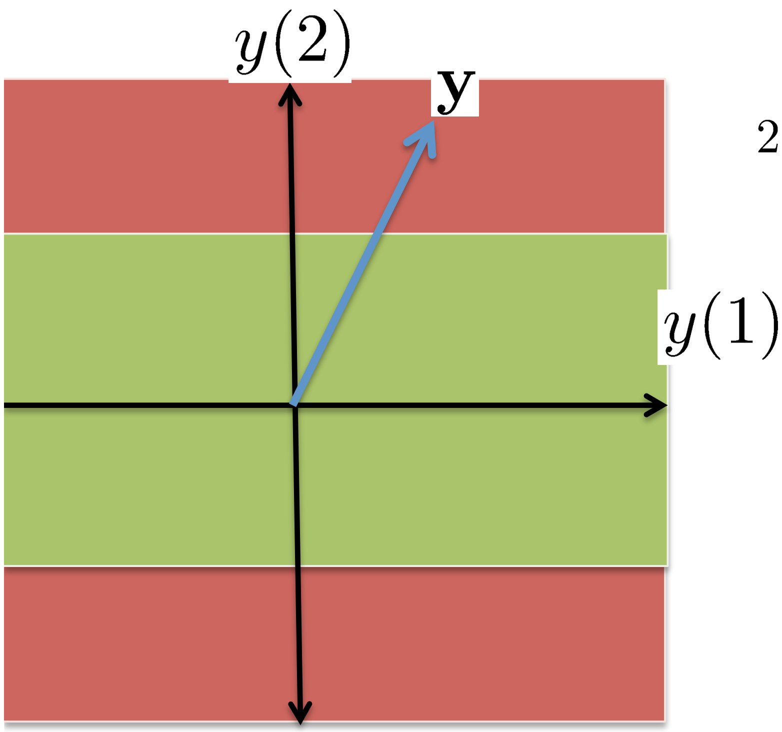
$$\longrightarrow \begin{bmatrix} y_i(1) \\ y_j(2) \end{bmatrix} = g\left(\begin{bmatrix} y_j(1) \\ y_j(2) \end{bmatrix} \right), \quad \exists g \in \mathcal{G}$$

also TRUE

Example 2 : MI test

$$\mathcal{H}_0 : (\theta_1, \theta_2) \in \mathbb{R} \times \{0\}$$

$$\mathcal{H}_1 : (\theta_1, \theta_2) \in \mathbb{R} \times \mathbb{R}$$



apply maximally invariant test

maximally invariant (MI) statistic

$$t\left(\begin{bmatrix} y(1) \\ y(2) \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y(1) \\ y(2) \end{bmatrix} = \begin{bmatrix} 0 \\ y(2) \end{bmatrix}$$

GLRT over MI statistic

$$2 \ln \hat{L}_R(t(\mathbf{y})) = \max_{c \in \mathbb{R}} 2 \ln L\,(c; t(\mathbf{y})) - 2 \ln L\,(0; t(\mathbf{y}))$$

$$= y(2)^2$$ **(supplemental slides)**

maximally invariant test

$$\phi(\mathbf{y}) = \begin{cases} 1, & y(2)^2 > \eta \\ 0, & y(2)^2 \le \eta \end{cases}$$

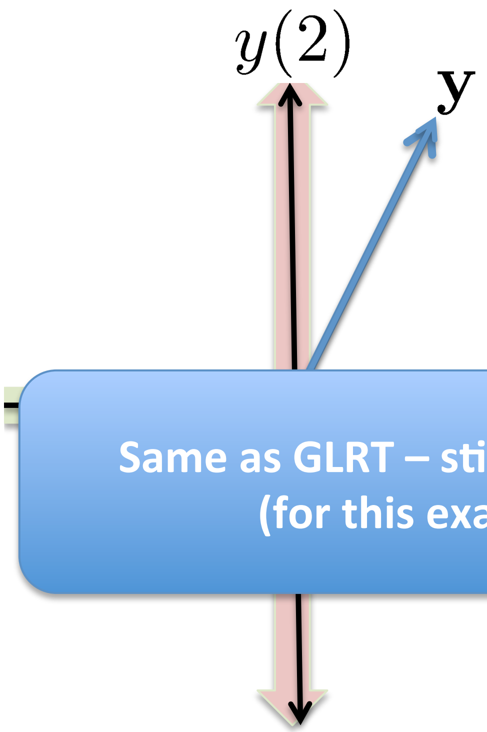
Same as the GLRT … and optimal (for this example)

Penn Engineering

PRECISE

Example 3 : MI test

$$\mathcal{H}_0 : (\theta_1, \theta_2) \in \mathbb{R} \times \{0\}$$
$$\mathcal{H}_1 : (\theta_1, \theta_2) \in \{0\} \times \mathbb{R}$$



Same as GLRT – still won't work (for this example)

apply maximally invariant test

nuisance parameters

$$\hat{\Theta}_N = \Theta_1 \cap \Theta_0 \rightarrow \hat{\Theta}_{N,1} = 0, \hat{\Theta}_{N,2} = 0$$

maximally invariant statistic

$$t(\mathbf{y}) = \mathbf{y}$$

GLRT over MI statistic

$$2 \ln \hat{L}_R(t(\mathbf{y})) = \max_{\hat{\theta}_1 \in \Theta_1} 2 \ln L\left(\hat{\theta}_1; \mathbf{y}\right) - \max_{\hat{\theta}_0 \in \Theta_0} 2 \ln L\left(\hat{\theta}_0; \mathbf{y}\right)$$
$$= y(2)^2 - y(1)^2$$

distribution of GLRT over MI statistic

$$\mathcal{H}_0 : 2 \ln \hat{L}_R(\mathbf{y}) \sim \text{ ???}$$

"difference of unknown non-central chi-squared and central chi-squared"

No bounds on false positive rate

Parameter-Invariant Test

- Tests over Parameter-Invariant (PAIN) statistics
  - extends principles of MI statistics
  - guarantees test has constant false positive rates
- Goal: test is MI to *null* parameters *(i.e. nuisance & null test parameters)*
  - discriminates based only *on the event test parameters*
- Parameter invariant testing is a 2-step approach
  1) design a PAIN statistic that is MI to nuisance and null test parameters
     - possibly "expanding" nuisance parameters for each hypothesis
  2) apply GLRT to PAIN statistic
- Performance Tradeoff
  - constant false positive rate
  - *loss of discriminatory information*
    - *null test parameters*
  - *if there are no null test parameters → PAIN = MI*

Parameter Invariant Testing … More Formally

- Transformation groups:

$$\hat{\mathcal{G}}_{N,0} = \left\{ g(g_0(\mathbf{y})) \ \mid g \in \hat{\mathcal{G}}_N, \ g_0 \in \hat{\mathcal{G}}_0 \right\}$$

- PAIN statistic is MI statistics to a transformation group:

$$MI(t; \hat{\mathcal{G}}_{N,0})$$

- Properties:

$$MI(t; \hat{\mathcal{G}}_{N,0}) \longrightarrow INV(t; \hat{\mathcal{G}}_0) \ \wedge \ INV(t; \hat{\mathcal{G}}_N)$$

$$\Theta_0 = \emptyset \ \wedge \ MI(t; \hat{\mathcal{G}}_{N,0}) \longrightarrow MI(t; \mathcal{G}_N)$$

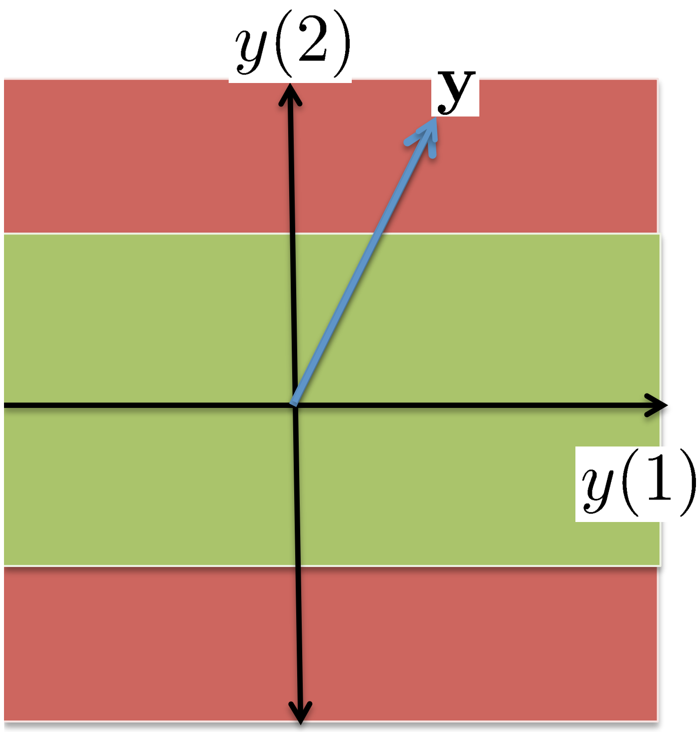
Penn Engineering

PRECISE

Example 3 : PAIN test

$\mathcal{H}_0 : (\theta_1, \theta_2) \in \mathbb{R} \times \{0\}$

$\mathcal{H}_1 : (\theta_1, \theta_2) \in \{0\} \times \mathbb{R}$



apply parameter invariant test

original transformation groups

$$\mathcal{G}_{\hat{\Theta}_N} = \left\{ g \middle| g\left(\begin{bmatrix} y(1) \\ y(2) \end{bmatrix}\right) = \begin{bmatrix} y(1) \\ y(2) \end{bmatrix} \right\}$$

$$\mathcal{G}_{\hat{\Theta}_0} = \left\{ g \middle| g\left(\begin{bmatrix} y(1) \\ y(2) \end{bmatrix}\right) = \begin{bmatrix} y(1) \\ y(2) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} c, \ \forall c \in \mathbb{R} \right\}$$

$$\mathcal{G}_{\hat{\Theta}_1} = \left\{ g \middle| g\left(\begin{bmatrix} y(1) \\ y(2) \end{bmatrix}\right) = \begin{bmatrix} y(1) \\ y(2) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} c, \ \forall c \in \mathbb{R} \right\}$$

parameter-invariant transformation group

$$\mathcal{G}_{\hat{\Theta}_{N,0}} = \left\{ g \middle| g\left(\begin{bmatrix} y(1) \\ y(2) \end{bmatrix}\right) = \begin{bmatrix} y(1) \\ y(2) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} c, \ \forall c \in \mathbb{R} \right\}$$

same transformation group as in example 2

parameter invariant test

$$\phi(\mathbf{y}) = \begin{cases} 1, & y(2)^2 > \eta \\ 0, & y(2)^2 \leq \eta \end{cases}$$

Penn Engineering

PRECISE

Parameter-Invariant Test (in 2 directions)

- Most real-world scenarios are not binary
- Goal: ensure test maintains:
  - upper bounds on false positive rate
  - lower bounds on true positive rate
- Parameter invariant test in both directions

  1) design a test with constant false positive rate (test 1)

  2) design a test with constant true positive rate (test 2)

  3) only accept hypothesis when both tests agree

	test 1 accepts H_0	test 1 rejects H_0
test 2 accepts H_1	low power	accept H_1
test 2 rejects H_1	accept H_0	bad model

Penn Engineering

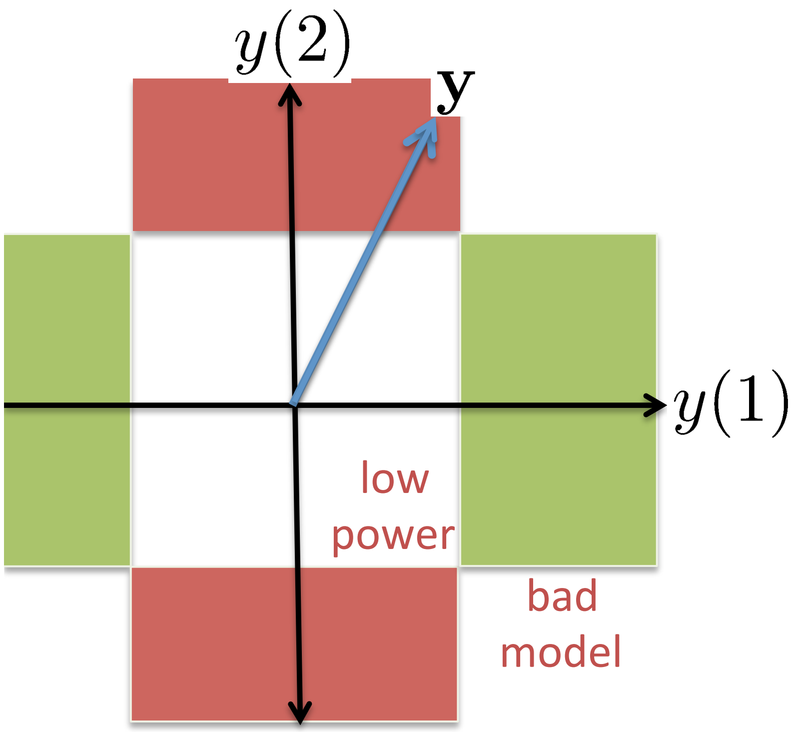
PRECISE

Example 3 : 2-sided PAIN test

$$\mathcal{H}_0 : (\theta_1, \theta_2) \in \mathbb{R} \times \{0\}$$

$$\mathcal{H}_1 : (\theta_1, \theta_2) \in \{0\} \times \mathbb{R}$$

apply 2-sided parameter invariant test



PAIN test 1

$$\phi_1(\mathbf{y}) = \begin{cases} 1, & y(2)^2 > \eta_1 \\ 0, & y(2)^2 \leq \eta_1 \end{cases}$$

PAIN test 2

$$\phi_2(\mathbf{y}) = \begin{cases} 1, & y(1)^2 \leq \eta_2 \\ 0, & y(1)^2 > \eta_2 \end{cases}$$

2-sided PAIN test

$$\phi(\mathbf{y}) = \begin{cases} 1, & y(2)^2 > \eta_1 \ \wedge \ y(1)^2 \leq \eta_2 \\ 0, & y(2)^2 \leq \eta_1 \ \wedge \ y(1)^2 > \eta_2 \\ ?, & \text{else} \end{cases}$$

Penn Engineering

PRECISE

PAIN Foundations : Summary

- Classical tests do not always bound performance
  - likelihood ratio tests
  - generalized likelihood ratio tests
  - maximally invariant tests
- Parameter Invariant tests:
  - constant false alarm rate
  - optimal (when an optimal test exists)
  - 2-sided test addresses errors in hypothesis design
- The following tutorial module addresses how to design a PAIN monitor
  - a common transformation group