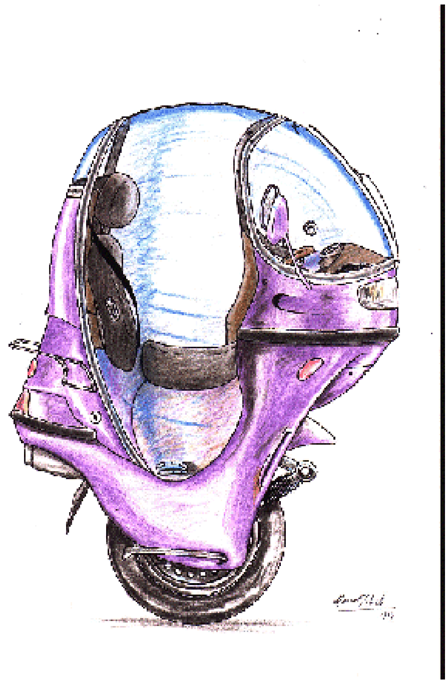
Bayesian estimation using interval analysis

Luc Jaulin,
Laboratoire E^3I^2
(Extraction et Exploitation de l'Information en Environnements Incertains)
ENSIETA, 2 rue François Verny, 29806 Brest Cédex 09

Franco-Japanese Workshop on Constraint Programming
Tokyo, October 27, 2004

What is control theory ?



1 Constraints for control

1.1 Set inversion

When the CSP can be written as

$$\mathbf{f}(\mathbf{x}) \in [\mathbf{y}], \mathbf{x} \in [\mathbf{x}] \subset \mathbb{R}^n, [\mathbf{y}] \subset \mathbb{R}^m$$

One can get an inner and an outer set for the solution set

$$\mathbb{S} = \mathbf{f}^{-1}([\mathbf{y}]) \cap [\mathbf{x}].$$

(Set demo, Proj2d)

1.2 Equilibrium points

Sailing boat

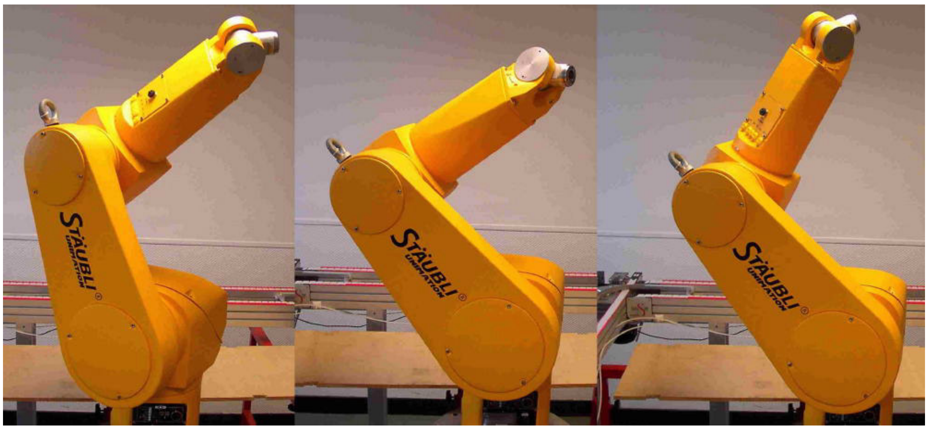
$$\begin{cases} \dot{x} & = & v \cos\theta, \\ \dot{y} & = & v \sin\theta - \beta V, \\ \dot{\theta} & = & \omega, \\ \dot{\delta}_v & = & u_1, \\ \dot{\delta}_g & = & u_2, \\ \dot{v} & = & \frac{f_v \sin\delta_v - f_g \sin\delta_g - \alpha_f v}{m}, \\ \dot{\omega} & = & \frac{(\ell - r_v \cos\delta_v) f_v - r_g \cos\delta_g f_g - \alpha_\theta \omega}{J}, \\ f_v & = & \alpha_v \left(V \cos(\theta + \delta_v) - v \sin\delta_v \right), \\ f_g & = & \alpha_g v \sin\delta_g. \end{cases}$$

Equilibrium points should satisfy

$$
\left\{
\begin{array}{rcl}
0 & = & v \cos\theta, \\
0 & = & v \sin\theta - \beta V, \\
0 & = & \omega, \\
0 & = & u_1, \\
0 & = & u_2, \\
0 & = & \frac{f_v \sin\delta_v - f_g \sin\delta_g - \alpha_f v}{m}, \\
0 & = & \frac{(\ell - r_v \cos\delta_v) f_v - r_g \cos\delta_g f_g - \alpha_\theta \omega}{J}, \\
f_v & = & \alpha_v \left(V \cos(\theta + \delta_v) - v \sin\delta_v \right), \\
f_g & = & \alpha_g v \sin\delta_g.
\end{array}
\right.
$$

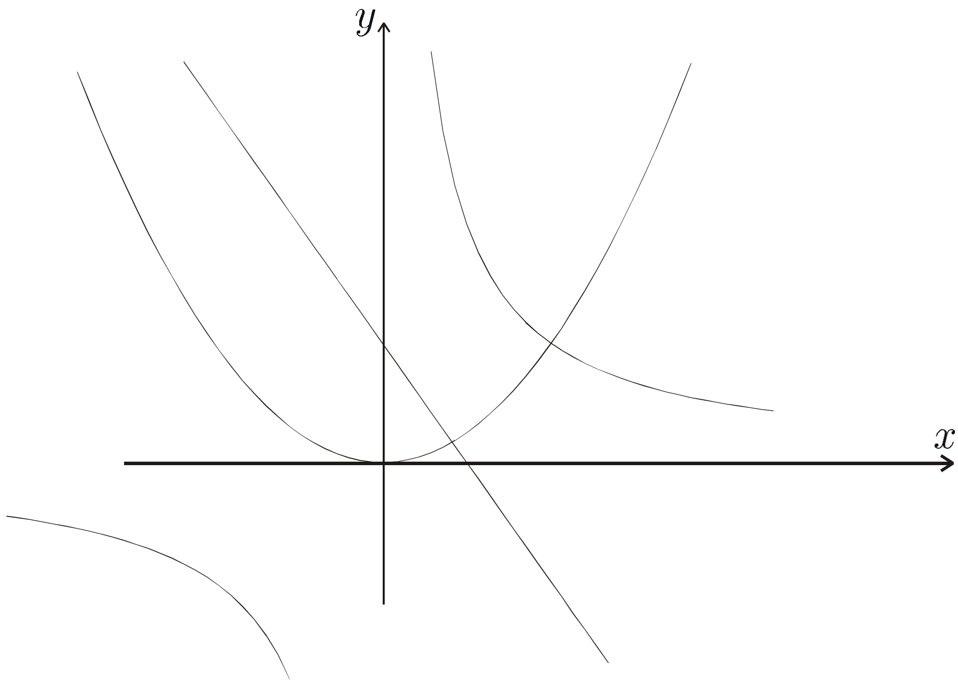
This CSP has no solution.

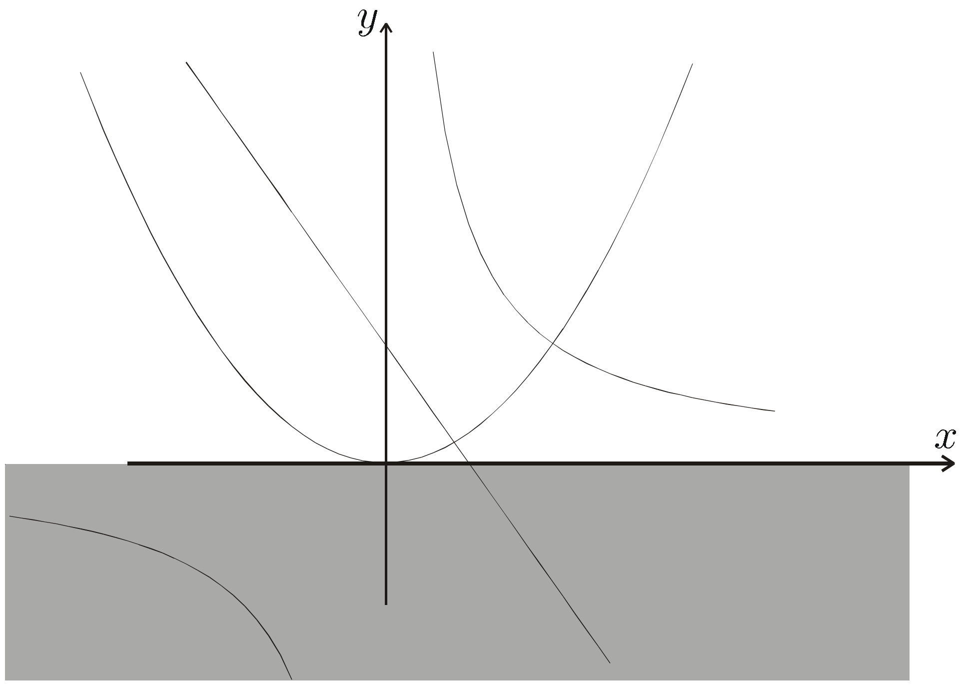
1.3 Calibration and state estimation

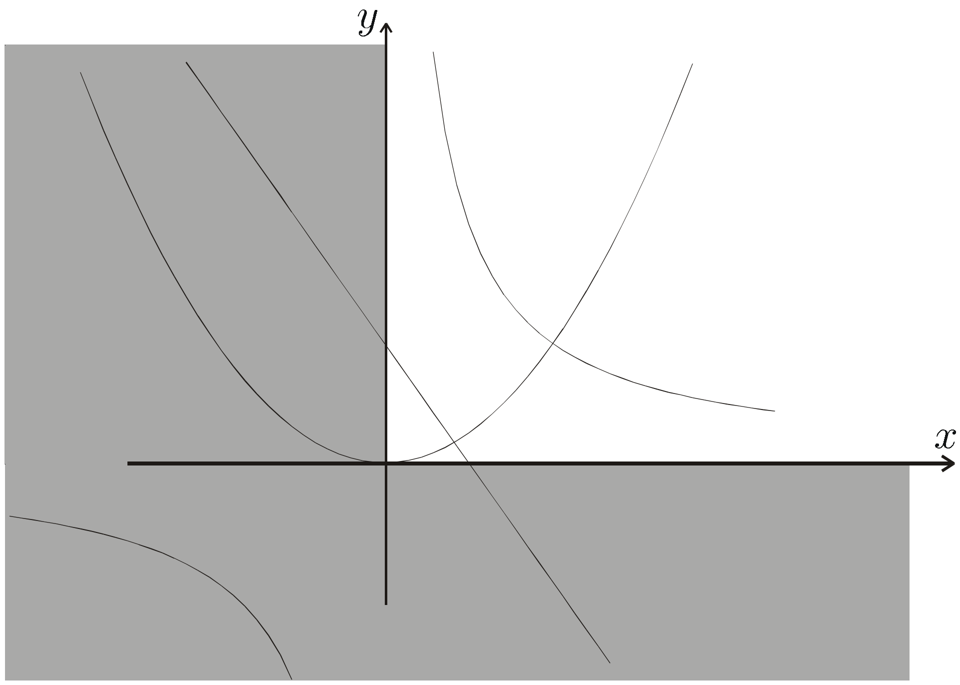


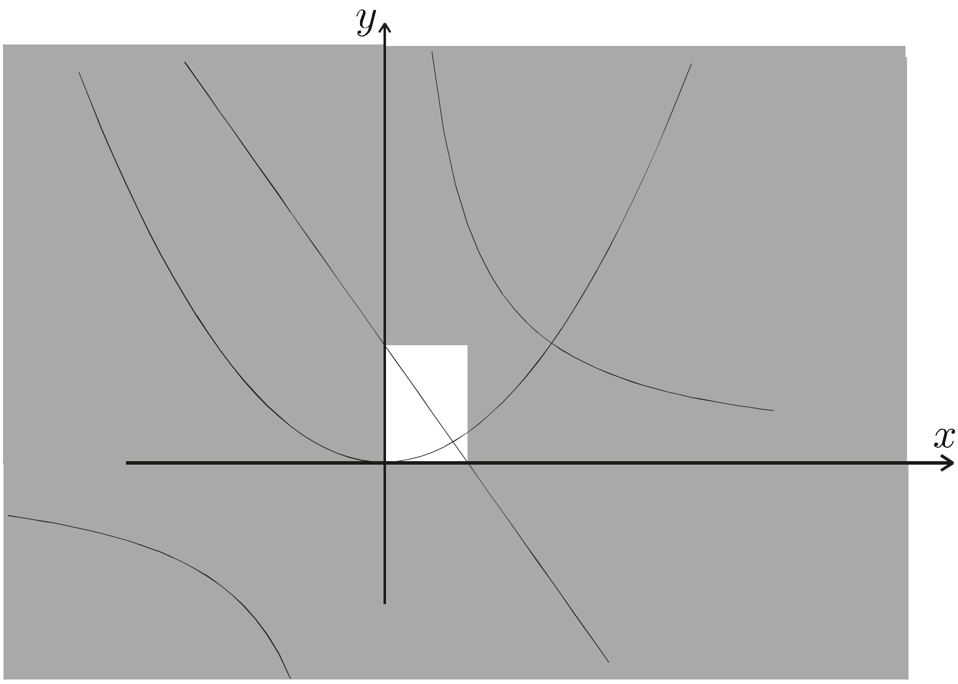
1.4 Topology for robotics

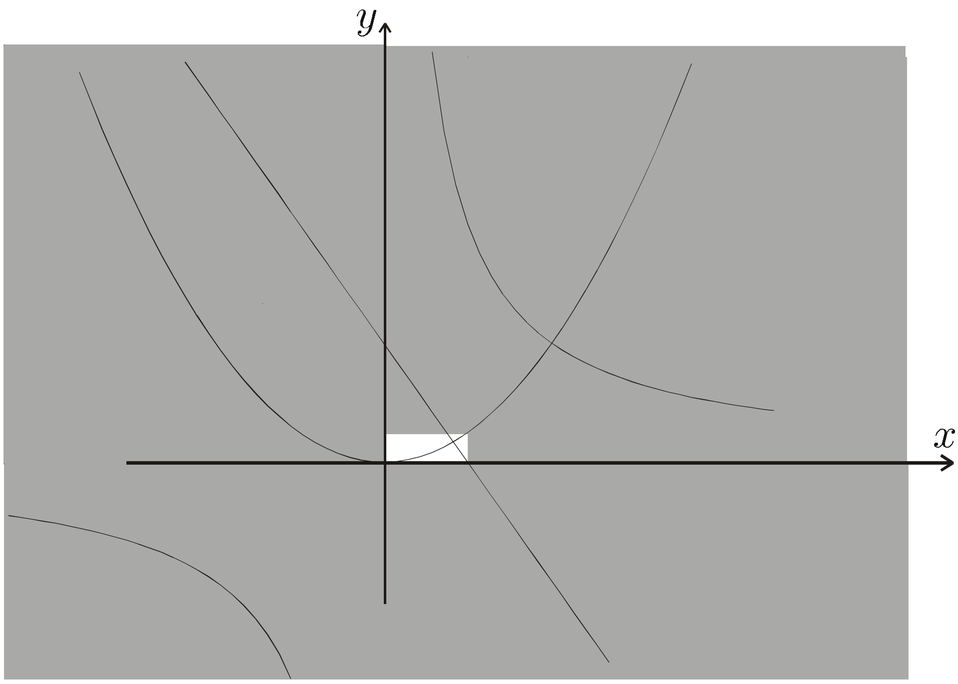
2 Constraint propagation

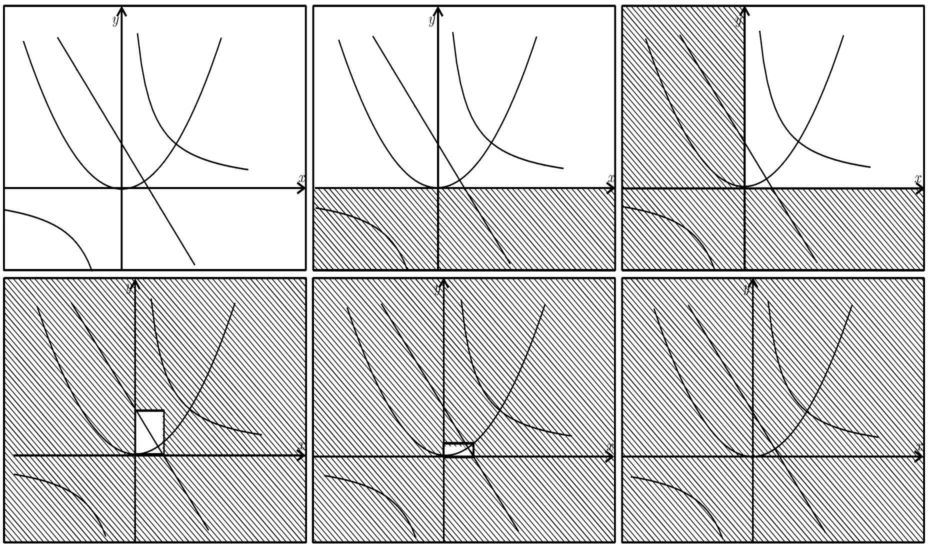



y
x


y
x


y
x


y
x



$$
\begin{array}{lll}
(C_1) \Rightarrow & y \in &]-\infty,\infty[^2 = \ [0,\infty[\\
(C_2) \Rightarrow & x \in & 1/[0,\infty[= [0,\infty[\\
(C_3) \Rightarrow & y \in & [0,\infty[\cap \ ((-2)\,.[0,\infty[+1) \\
& & = [0,\infty[\cap \ (]-\infty,1]) = [0,1] \\
& x \in & [0,\infty[\cap \ (-[0,1]/2+1/2) \\
& & = [0,1/2] \\
(C_1) \Rightarrow & y \in & [0,1] \cap [0,1/2]^2 = [0,1/4] \\
(C_2) \Rightarrow & x \in & [0,1/2] \ \cap \ 1/[0,1/4] = \emptyset \\
& y \in & [0,1/4] \cap 1/\emptyset = \emptyset
\end{array}
$$

Extend the class of constraints that can be projected in a polynomial time (i.e., global constraints ?). For instance, the constraint

$$\mathbf{A} \geq \mathbf{0} \text{ where } \mathbf{A} \in \mathcal{M}_{n,n}$$

can be projected in $o(n^{8.5})$. The constraint

$$C(a_n, \ldots, a_0) \ : \ (a_n s^n + \cdots + a_1 s + a_0 \text{ unstable})$$

can be projected in $o(n^2)$.

What about

$$\text{Rot}(\mathbf{A}), \quad \mathbf{A} = \exp(\mathbf{B}), \quad \mathbf{A} = \mathbf{B}.\mathbf{C},$$

where $\mathbf{A} \in \mathcal{M}_{n,n}, \mathbf{B} \in \mathcal{M}_{n,n}, \mathbf{C} \in \mathcal{M}_{n,n}$?

What about the constraint

$$a_n s^n + \cdots + a_1 s + a_0 \text{ stable ?}$$

3 Confidence regions

Consider a function $f(\mathbf{p})$ positive for all $\mathbf{p} \in \mathbb{R}^n$, such as $\int_{\mathbb{R}^n} f(\mathbf{p}) d\mathbf{p}$ is finite and a real number $\alpha \in [0, 1]$. Characterize the set $\mathbb{S}_\alpha$ defined by

$$\begin{array}{ll} \text{(i)} & \mathbb{S}_\alpha = f^{-1}([s_\alpha, +\infty[), \\ \text{(ii)} & \dfrac{\int_{\mathbb{S}_\alpha} f(\mathbf{p}) d\mathbf{p}}{\int_{\mathbb{R}^n} f(\mathbf{p}) d\mathbf{p}} = \alpha. \end{array}$$

The set $\mathbb{S}_\alpha$ is the confidence region associated with the unnormalized pdf f.

It corresponds to the smallest set which contains $\mathbf{p}$ with a probability equal to α.

Example : Consider a random variable p, described by the unnormalized pdf:

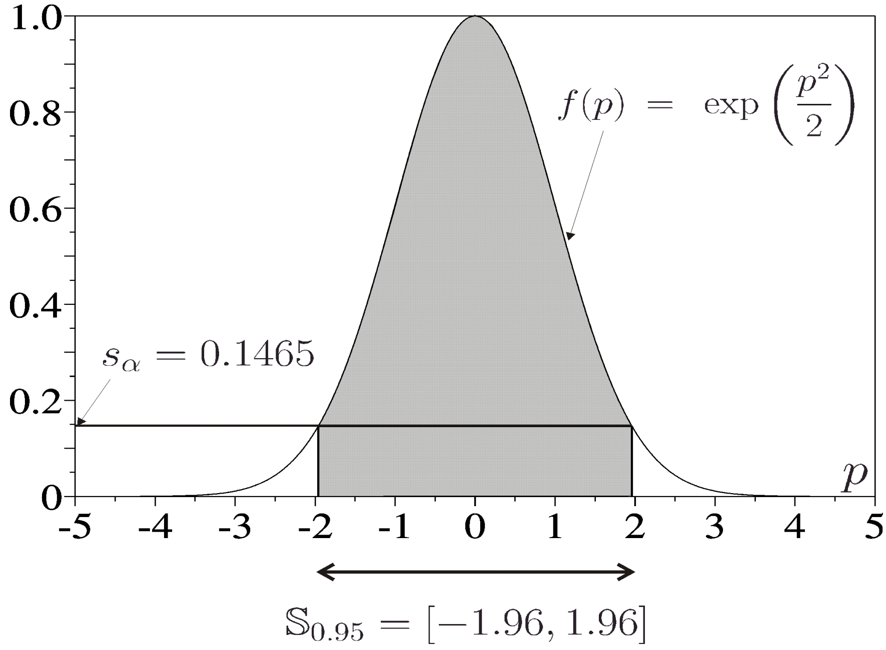
$$f(p) = \text{exp}\left(-\frac{p^2}{2}\right).$$

Since,

$$\int_{-\infty}^{\infty} \text{exp}\left(-\frac{p^2}{2}\right) dp = \sqrt{2\pi},$$

Finding its confidence region $\mathbb{S}_{0.95}$ amounts to solving

$$\begin{array}{ll} \text{(i)} & \mathbb{S}_{0.95} = f^{-1}([s_\alpha, +\infty[), \\ \text{(ii)} & \frac{1}{\sqrt{2\pi}} \int_{\mathbb{S}_\alpha} f(\mathbf{p}) d\mathbf{p} = 0.95. \end{array}$$


1.0
0.8
0.6
0.4
0.2
0
-5 -4 -3 -2 -1 0 1 2 3 4 5
p
$f(p) = \exp\left(\frac{p^2}{2}\right)$
$s_\alpha = 0.1465$
$\mathbb{S}_{0.95} = [-1.96, 1.96]$

4 Intervals

A *lattice* $(\mathcal{E}, \leq)$ is a partially ordered set, closed under least upper and greatest lower bounds.

The least upper bound (*join*) of x and y is written $x \vee y$.

The greatest lower bound (*meet*) is written $x \wedge y$.

A lattice $\mathcal{E}$ is *complete* if for all subsets $\mathcal{A}$ of $\mathcal{E}$, $\vee \mathcal{A}$ and $\wedge \mathcal{A}$ belong to $\mathcal{E}$.

An *interval* $[x]$ of a complete lattice $\mathcal{E}$ is a subset of $\mathcal{E}$ which satisfies

$$[x] = \{x \in \mathcal{E} \mid \wedge [x] \leq x \leq \vee[x]\} \,.$$

Both $\emptyset$ and $\mathcal{E}$ are intervals of $\mathcal{E}$.

The sets $[0,1]_{\bar{\mathbb{R}}}$ and $[0,\infty]_{\bar{\mathbb{R}}}$ are intervals of $\bar{\mathbb{R}}$.

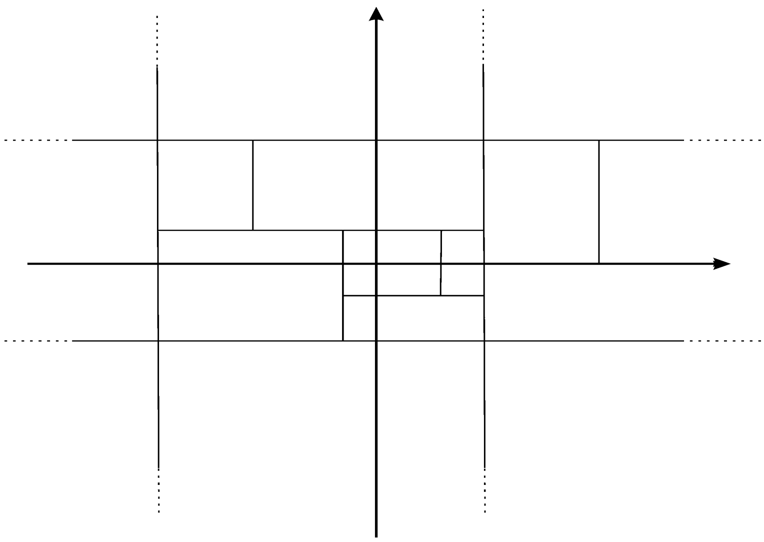
The set $\{2,3,4,5\} = [2,5]_{\bar{\mathbb{N}}}$ is an interval of $\bar{\mathbb{N}}$.

The set $\{4,6,8,10\} = [4,10]_{2\bar{\mathbb{N}}}$ is an interval of $2\bar{\mathbb{N}}$.

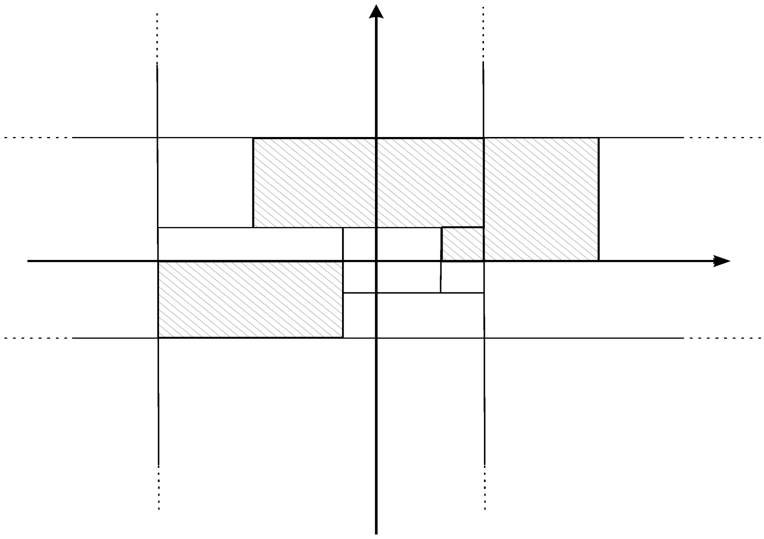
The set $[1,2] \times [3,4]) = [(1,3),(2,4)]_{\bar{\mathbb{R}}^2}$ is an interval of $\bar{\mathbb{R}}^2$.

5 Interval subpavings

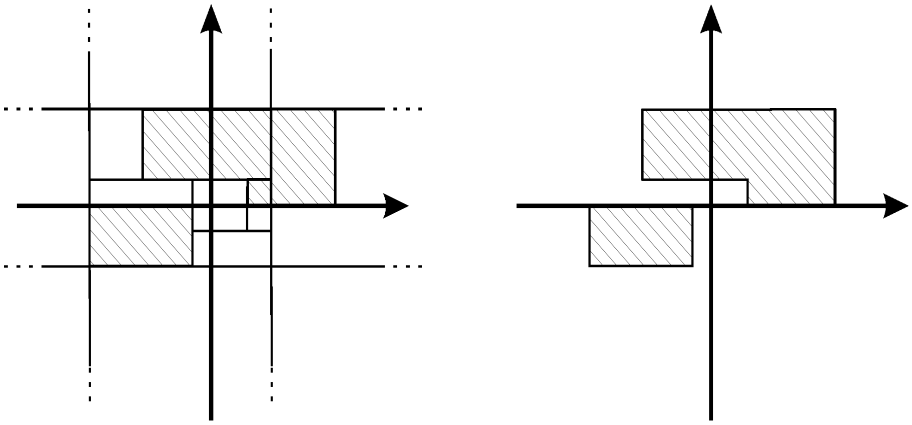
A *paving* $\mathcal{Q}$ of $\mathbb{R}^n$ is a set of nonoverlapping boxes covering $\mathbb{R}^n$.



A *subpaving* of $\mathcal{Q}$ is a subset of $\mathcal{Q}$.



The *support* $\{\mathcal{K}\} \subset \mathbb{R}^n$ of a subpaving $\mathcal{K}$ is the union of all boxes of $\mathcal{K}$.



If $\mathcal{P}(\mathcal{Q})$ denotes the set of all subpavings of $\mathcal{Q}$ then $(\mathcal{P}(\mathcal{Q}), \subset)$ is a complete lattice.

- The least upper bound is the union:

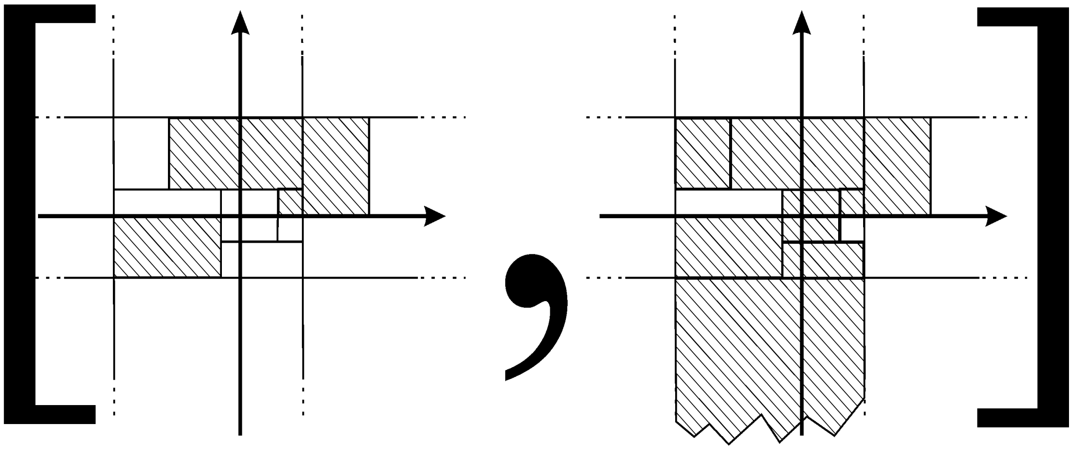
$$\mathcal{K}_1 \vee \mathcal{K}_2 = \mathcal{K}_1 \cup \mathcal{K}_2.$$

- The greatest lower bound is the intersection

$$\mathcal{K}_1 \wedge \mathcal{K}_2 = \mathcal{K}_1 \cap \mathcal{K}_2.$$

As a consequence intervals of $(\mathcal{P}(\mathcal{Q}), \subset)$ can be defined.

An interval subpaving $[\mathcal{K}^-, \mathcal{K}^+]$ of $\mathcal{Q}$ can be represented by pair of subpavings of $\mathcal{Q}$ such that $\mathcal{K}^- \subset \mathcal{K}^+$.

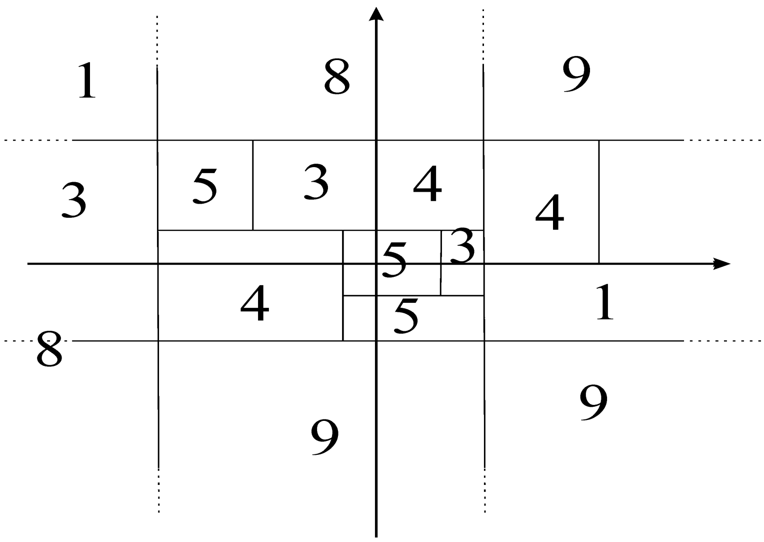


Definition:

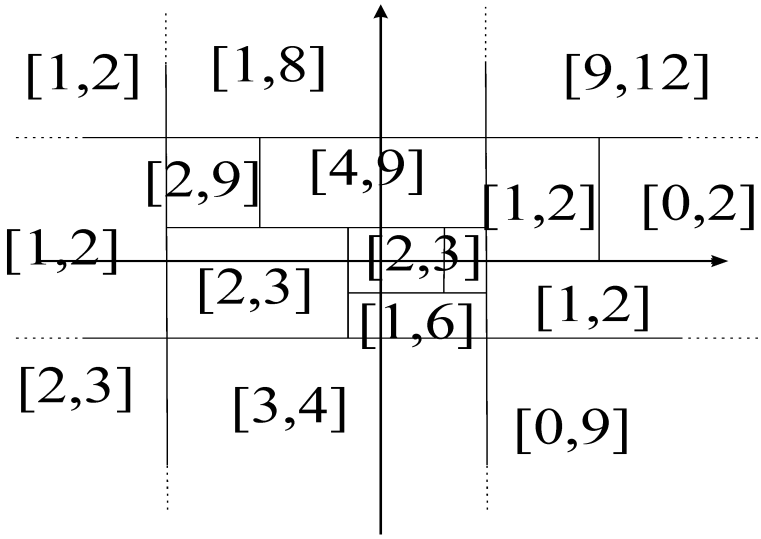
$$\mathbb{S} \in [\mathcal{K}^-, \mathcal{K}^+] \Leftrightarrow \left\{\mathcal{K}^-\right\} \subset \mathbb{S} \subset \left\{\mathcal{K}^+\right\}.$$

6 Interval staircase functions

A *staircase function* $\hat{f}$ associated with a paving $\mathcal{Q}$ is a function from $\mathcal{Q}$ to $\bar{\mathbb{R}}$.



The set of all staircase functions $(\hat{\mathcal{F}}, \leq)$ is a complete lattice. Interval staircase functions can thus be defined



A function f from $\mathbb{R}^n \longrightarrow \mathbb{R}$ is said to belong to the interval staircase function $[\hat{f}]$ if

$$\forall [\mathbf{p}] \in \mathcal{Q}, \forall \mathbf{p} \in [\mathbf{p}], f(\mathbf{p}) \in [\hat{f}^-([\mathbf{p}]), \hat{f}^+([\mathbf{p}])].$$

An interval staircase function for $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ can be obtained by using interval techniques.

The *reciprocal image* of the interval $[s^-, s^+] \in \mathbb{IR}$ by the interval staircase function $[\hat{f}] = [\hat{f}^-, \hat{f}^+]$ is the interval subpaving of $\mathcal{Q}$ defined by

$$[\hat{f}]^{-1}([s^-, s^+]) \triangleq \left[\left\{[\mathbf{p}] \in \mathcal{Q} \mid [\hat{f}]([\mathbf{p}]) \subset [s^-, s^+]\right\}, \left\{[\mathbf{p}] \in \mathcal{Q} \mid [\hat{f}]([\mathbf{p}]) \cap [s^-, s^+] \neq \emptyset\right\}\right]$$

Theorem

If f belongs to $[\hat{f}]$, then for all $[s^-, s^+] \in \mathbb{IR}$,

$$f^{-1}([s^-, s^+]) \in [\hat{f}]^{-1}([s^-, s^+]).$$

If $\left[\mathcal{K}^{-}, \mathcal{K}^{+}\right]$ is an interval subpaving of $\mathcal{Q}$ and if $[\hat{f}]$ is a positive interval staircase function, the *integral* of $[\hat{f}]$ over $\left[\mathcal{K}^{-}, \mathcal{K}^{+}\right]$ is

$$\int_{[\mathcal{K}^{-},\mathcal{K}^{+}]} [\hat{f}](\mathbf{p})d\mathbf{p} \triangleq \left[\sum_{[\mathbf{p}]\in\mathcal{K}^{-}} \hat{f}^{-}([\mathbf{p}]).\text{volume}([\mathbf{p}]) \; . \right.$$

$$\left. \sum_{[\mathbf{p}]\in\mathcal{K}^{+}} \hat{f}^{+}([\mathbf{p}]).\text{volume}([\mathbf{p}]) \right]$$

Theorem

If $f \in [\hat{f}]$ and if $\mathbb{S} \in \left[\mathcal{K}^-, \mathcal{K}^+\right]$, then

$$\int_{\mathbb{S}} f(\mathbf{p}) d\mathbf{p} \in \int_{[\mathcal{K}^-, \mathcal{K}^+]} [\hat{f}](\mathbf{p}) d\mathbf{p}.$$

7 Algorithm

Equation in s_α to be solved

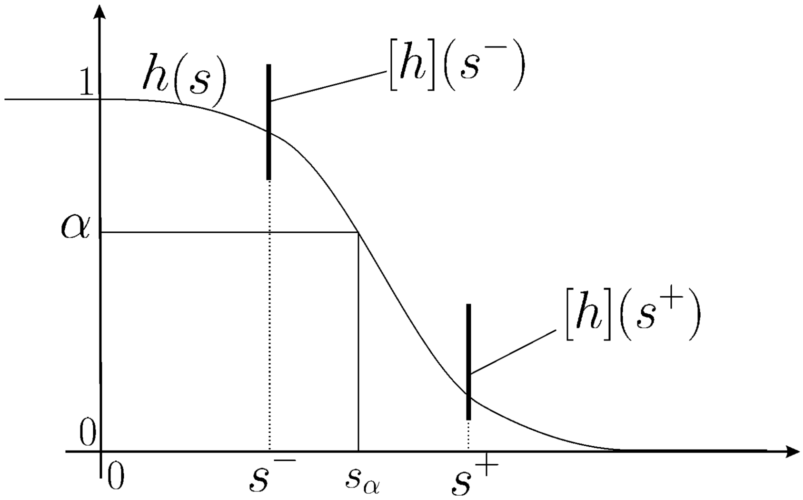
$$\alpha = h(s_\alpha) \triangleq \frac{\int_{f^{-1}([s_\alpha,\infty[)} f(\mathbf{p})d\mathbf{p}}{\int_{\mathbb{R}^n} f(\mathbf{p})d\mathbf{p}}$$

The function $h(s)$ is decreasing. Moreover,

$$h(s) \in [h](s) \triangleq \frac{\int_{[\hat{f}]^{-1}([s,\infty[)}[f](\mathbf{p})d\mathbf{p}}{\int_{\mathbb{R}^n}[\hat{f}](\mathbf{p})d\mathbf{p}}.$$

Thus

$$\begin{array}{lll} (a) & \alpha < lb([h](s^-)) & \Rightarrow s^- < s_\alpha \\ (b) & \alpha > ub([h](s^+)) & \Rightarrow s^+ > s_\alpha \end{array}$$



1. Take a paving $\mathcal{Q}$ of $\mathbb{R}^n$; $s^- := +\infty$; $s^+ := 0$;

2. Compute an interval staircase function $[\hat{f}]$ enclosing f;

3. Decrease s^- until $\alpha < lb([h](s^-))$

4. Increase s^+ until $\alpha > ub([h](s^+))$;

5. $\left[\mathcal{K}_\alpha^-, \mathcal{K}_\alpha^+\right] := ([\hat{f}] - [s^-, s^+])^{-1}([0, \infty[)$.

Theorem : After completion of this algorithm, we have

$$\mathbb{S}_\alpha \in \left[\mathcal{K}_\alpha^-, \mathcal{K}_\alpha^+\right] \text{ and } s_\alpha \in [s^-, s^+].$$

8 Application to Bayesian estimation

Model:

$$y(t) = p_1 \sin(p_2 t) + n(t)$$

where $n(t)$ is a white normal random signal with:

$$\pi_n(n) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{n^2}{2\sigma^2}\right),$$

where the standard deviation is $\sigma = \frac{1}{2}$.

Sampling times and data:

$$\begin{aligned} \mathbf{t} &= (1, 2, 3), \\ \mathbf{y} &= (0.8,\ 1.0,\ 0.2)^{\mathsf{T}}. \end{aligned}$$

Therefore

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}}_{\mathbf{y}} = \underbrace{\begin{pmatrix} p_1 \sin(p_2) \\ p_1 \sin(2p_2) \\ p_1 \sin(3p_2) \end{pmatrix}}_{\phi(\mathbf{p})} + \underbrace{\begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}}_{\mathbf{n}}$$

Since $n(t)$ is white,

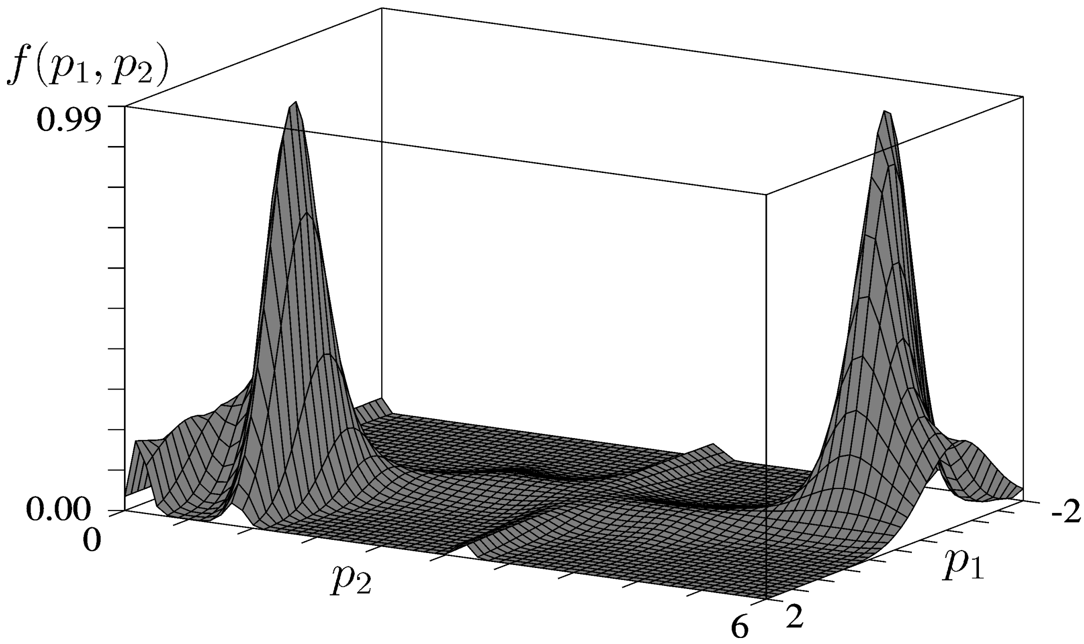
$$\begin{aligned}\pi_n(\mathbf{n}) &= \pi_n(n_1).\pi_n(n_2).\pi_n(n_3) \\ &= \frac{1}{\left(\sqrt{2\pi}\right)^3} \exp(-2n_1^2) \exp(-2n_2^2) \exp(-2n_3^2).\end{aligned}$$

If

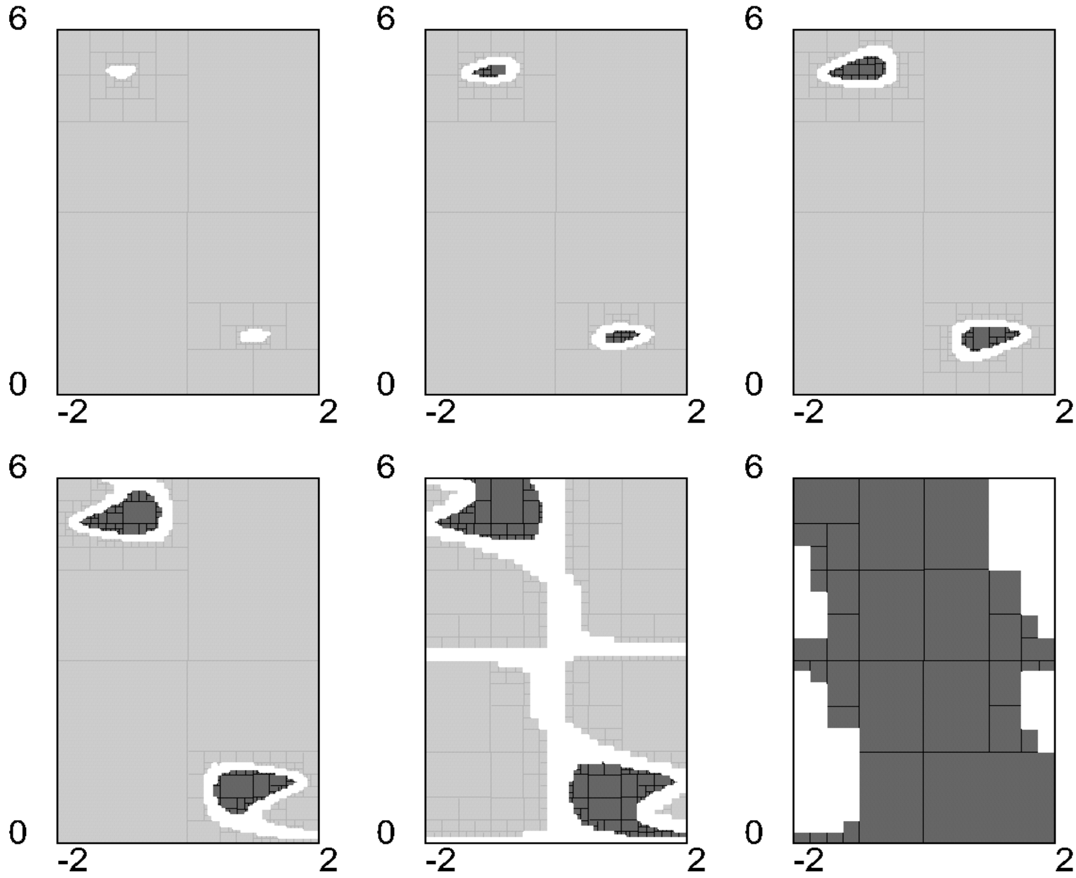
$$\pi_{\text{prior}}(\mathbf{p}) = \frac{\text{door}_{[-2,2]}(p_1).\text{door}_{[0,6]}(p_2)}{24}.$$

The posterior unnormalized pdf for $\mathbf{p}$:

$$f(\mathbf{p}) = \left(\prod_{k=1}^{3} \exp(-2\,(y_k - p_1 \sin(kp_2))^2)\right) . \text{door}_{[-2,2]}(p_1) . \text{door}_{[0,6]}(p_2).$$



$f(p_1, p_2)$
0.99
0.00
0
p_2
6
2
p_1
-2



$\left[\mathcal{K}_\alpha^-, \mathcal{K}_\alpha^+\right]$ for $\alpha \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$;