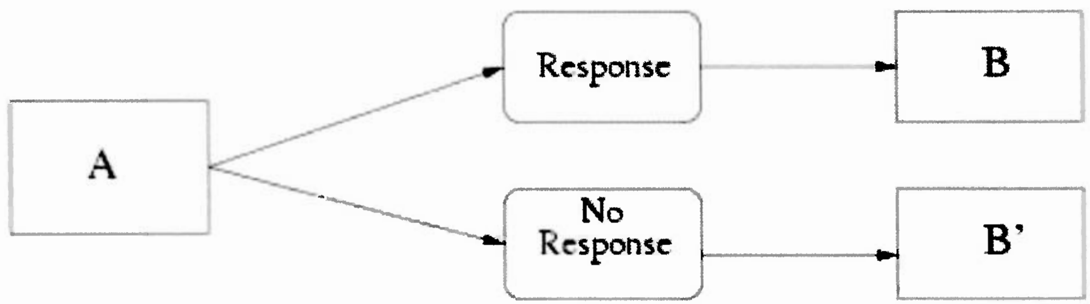
Schematically: The specific regime "Give first-line induction therapy A followed by maintenance B if response else if no response give second-line therapy B'"



Step 1 *Step 2*

(Induction Trt) (Intermediate Outcome) (Maintenance or Second-line Trt)

(M. Davidian slides)

Eight possible regimes:

1. A_1 followed by B_1 if response, else B_1'
2. A_1 followed by B_1 if response, else B_2'
3. A_1 followed by B_2 if response, else B_1'
4. A_1 followed by B_2 if response, else B_2'
5. A_2 followed by B_1 if response, else B_1'
6. A_2 followed by B_2 if response, else B_2'
7. A_2 followed by B_1 if response, else B_1'
8. A_2 followed by B_2 if response, else B_2'

Natural questions:

- What would be the *mean outcome* (e.g., *mean survival time*) if the *population* were to *follow* a particular regime?
- How do these mean outcomes *compare* among the possible regimes?

Background: Murphy '03

Data iid (S_j, A_j, X_j), $j = 0, \dots, J+1$, $Y := S_{J+1}$, $A_j \in \{-1, 1\}$

$$S_0 \to A_0 \to S_1 \to \dots \to A_{J-1} \to S_J \to A_J \to Y$$

potential outcomes of treatment $S(\bar{a}), Y(\bar{a})$ corresponding to treatment regime $\bar{a}$, link through consistency $Y = Y(A)$, $S_j = S_j(A)$

Consider "data-dependent" index into the potential outcomes

$$Y(\bar{d}_J) = Y(\bar{d}_J(\bar{S}(\bar{a}_{J-1}), \bar{a}_{J-1}))\Big|_{\bar{a}_{J-1} = \bar{d}_{J-1}}$$

$$d_j : \underbrace{(\bar{S}_j(\bar{a}_{j-1}), \bar{a}_{j-1})}_{\text{"past"}} \mapsto -1 \text{ or } 1$$

We seek to maximize the expected mean response

$$\mathbb{E}(Y(\bar{d}_J)) = \mathbb{E}(Y(\bar{a}_J)|_{\bar{a}_J = \bar{d}_J}) = \mathbb{E}\left\{ Y(\bar{a}_J)\big|_{a_0 = d_0(S_0), \dots, a_J = d_J(\bar{S}_J(\bar{a}_{J-1}), \bar{a}_{J-1})} \right\}$$

Identification is by an analogue of the "G-Formula"

~~$\mathbb{E}(Y(\bar{d}_J))$~~ under SRA: $\bar{S}_{J+1}(\bar{a}_J) \perp A_j \mid \bar{A}_{j-1}, \bar{S}_j$

$$\mathbb{E}(Y(\bar{d}_J)) = \mathbb{E}\left(\dots \mathbb{E}\left(\underbrace{\mathbb{E}(Y \mid \bar{S}_J, \bar{A}_{J-1}, A_J = d_J)} \mid \bar{S}_{J-1}, \bar{A}_{J-2}, A_{J-1} = d_{J-1} \right) \dots \right)$$

$\bar{A}_{J-1}$

$$= \mathbb{E}(Y(d_J(\bar{S}_J, \bar{A}_{J-1})) \mid \bar{S}_J, \bar{A}_{J-1}, A_J = d_J) \qquad \text{(consistency)}$$

$$= \mathbb{E}(Y(d_J(\bar{S}_J, \bar{A}_{J-1})) \mid \bar{S}_J, \bar{A}_{J-1}) \qquad \text{(SRA)}$$

$$\vdots$$

$$\mathbb{E}(Y(\bar{d}_J))$$

find optimal treatment regime "greedily"

given treatment $\bar{a}_j$. let

$$Q_0(\bar{S}_j, \bar{A}_{j-1}, a_j) := \mathbb{E}(Y \mid \bar{S}_j, \bar{A}_{j-1}, A_j = a_j)$$

$$J_0(\bar{S}_j, \bar{A}_{j-1}) := \sup_{a_j} Q_0(\bar{S}_j, \bar{A}_{j-1}, a_j)$$

$$Q_1(\bar{S}_{j-1}, \bar{A}_{j-2}, a_{j-1}) := \mathbb{E}(Y \mid \bar{S}_{j-1}, \bar{A}_{j-2}, a_{j-1})$$

$$J_1(\bar{S}_{j-1}, \bar{A}_{j-2}) := \sup_{a_{j-1}} Q_1(\bar{S}_{j-1}, \bar{A}_{j-2}, a_{j-1})$$

$$\vdots$$

$$J_K(S_0) := \sup_{a_0} Q_K(S_0, a_0)$$

Theorem (Murphy '03)

$$\sup_{\bar{d}} \mathbb{E}(\mathbb{E}(\ldots \mathbb{E}(\mathbb{E}(Y \mid \bar{S}_j, \bar{A}_{j-1}, A_j = d_j) \mid \bar{S}_{j-1}, \bar{A}_{j-2}, A_{j-1} = d_{j-1}) \ldots)))$$

$$= \mathbb{E}(J_K(S_0))$$

Outcome-Weighted Learning ("OWL", Zhao et al. 2012)

point exposure setting

Data iid (Y, A, X), $A \perp X$, $A \in \{-1, 1\}$, $Y \geq 0$

Problem: find rule $d: \mathcal{X} \to \{-1, 1\}$ to maximize Y in expectation if everyone followed it

First, estimate potential outcome $Y(d(X)) = Y(a)(\omega)\big|_{a = d(X(\omega))}$, only observed when $d(X) = A$

Why not use $\mathbb{P}(Y \mid A = d(X))$? Still possibility of confounding unless $\mathbb{P}(A=1 \mid X) \equiv \frac{1}{2}$)

d observed

	x_1	x_2
$A=1$	$Y=1$	$Y=2$
$A=2$	$Y=1$	$Y=2$

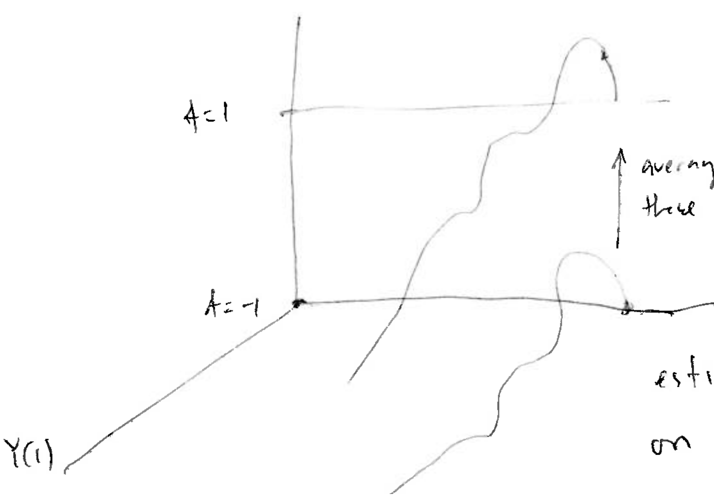
$\mathbb{P}(X = x_1) = \mathbb{P}(X = x_2) = \frac{1}{2}$

$$d(x) = \begin{cases} 1, & x = x_2 \\ -1, & x = x_1 \end{cases}$$

$$\mathbb{E}(Y(d(X))) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 2 = \frac{3}{2}$$

if $\mathbb{P}(A=1) = 1 - \mathbb{P}(A=-1) = \frac{1}{3}$, $\mathbb{P}(Y \mid A = d(X)) = \dfrac{\frac{1}{2}\left(2 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3}\right)}{\frac{1}{2}} = \frac{4}{3}$

usual picture for non-random treatment eg $a=1$



← observed

← not observed

SRA:
$\mathbb{P}(Y=y, A=1, X=x) \propto \mathbb{P}(Y=y, A=0, X=x)$

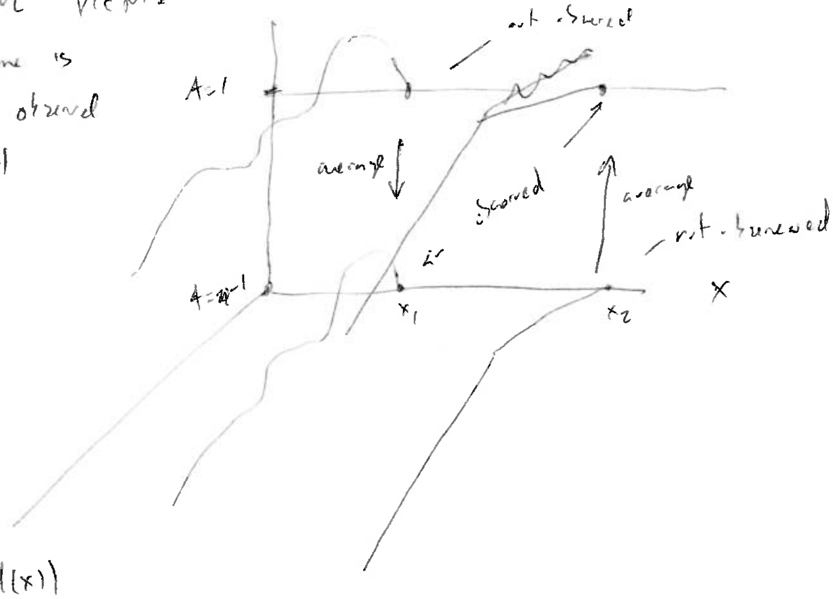
estimate $\mathbb{E}(Y(1))$ with $\tilde{\mathbb{P}}$ concentrated on $\{A=1\}$ using density $\dfrac{\{A=1\}}{\pi(X)}$

(#4)

OWL picture

potential outcome is $Y(d(x))$, only observed when $A=d(x)$

$d(x_1) = -1, \quad d(x_2) = 1$



$$\frac{d\tilde{P}}{dP} = \frac{1\{A=d(x)\}}{f(A)}$$

$f(A) = A\pi + (1-A)/2$

note constant at for $\pi = \frac{1}{2}$

straightforward to extend OWL to MAR → go to code

Given rule $d(x)$ the expected ~~reward~~ outcome if everyone were to follow it is therefore estimatable as

$$\mathbb{E}(Y(d(x))) = \mathbb{E}\left(\frac{Y1\{A=D(x)\}}{\pi A + (1-A)/2}\right) =: V(d)$$

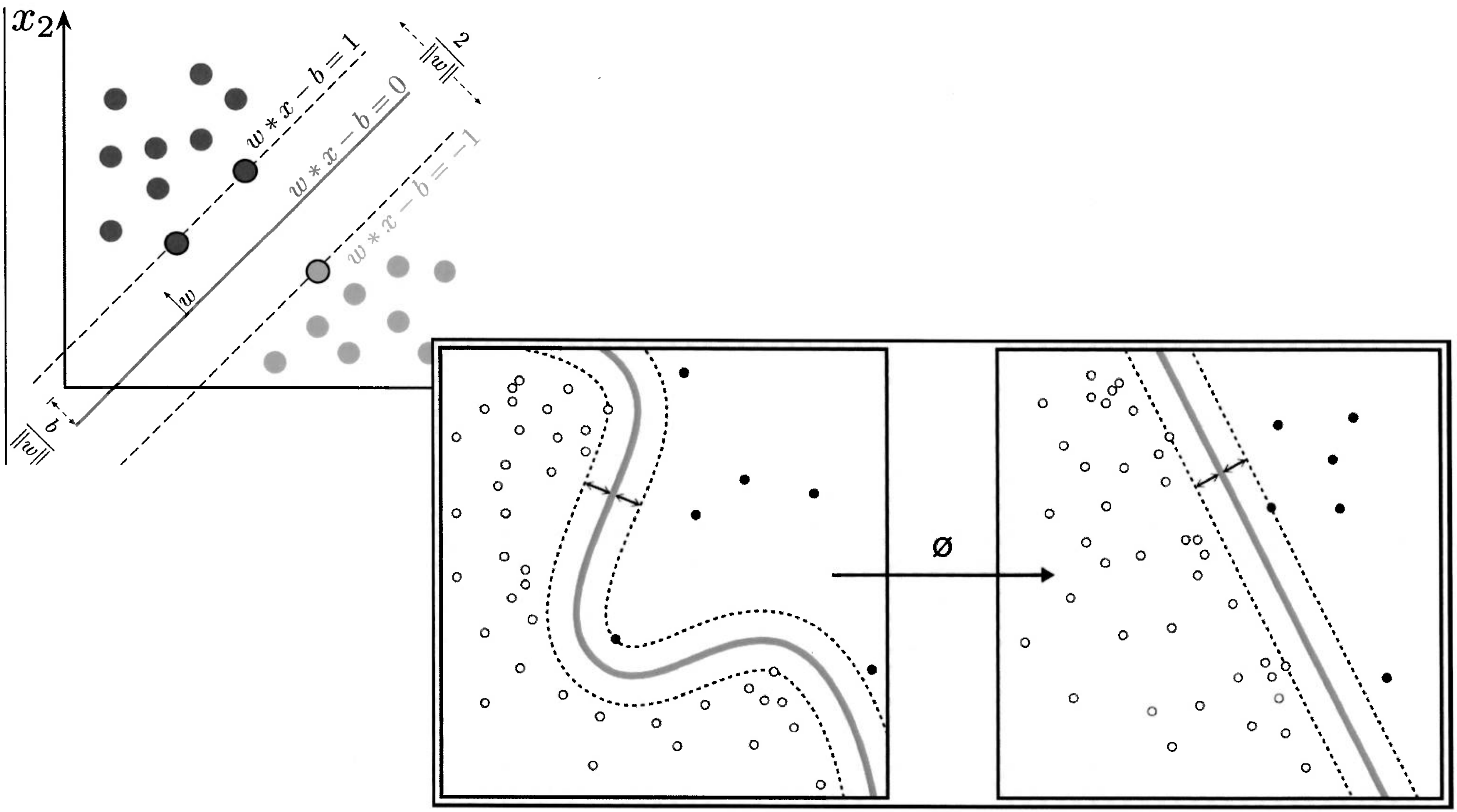
and we seek ~~to~~

$$\underset{d}{\arg\max}\ \mathbb{E}\left(\frac{Y1\{A=d(x)\}}{\pi A + (1-A)/2}\right)$$

$$= \underset{d}{\arg\max}\ \mathbb{E}\left(\frac{Y}{\pi A + (1-A)/2}\right) - \mathbb{E}\left(\frac{Y1\{A\neq d(x)\}}{\pi A + (1-A)/2}\right)$$

$$= \underset{d}{\arg\min}\ \mathbb{E}\left(\frac{Y1\{A\neq d(x)\}}{\pi A + (1-A)/2}\right)$$

interpretation: ~~expected loss~~ objective is expected loss ~~of using d(x)~~ under the classification rule $d(x)$, $\mathbb{E}(A\neq d(x))$, but weighted by $\frac{Y}{f(A)}$, so the penalty is more severe for large outcome Y (or rarer treatment A...)


x_2
$w * x - b = 1$
$w * x - b = 0$
$w * x - b = -1$
$\frac{2}{\|w\|}$
$\frac{b}{\|w\|}$
w
Ø

#5

~~e~~ empirically, we wish to minimize over ~~d~~ classifiers d

$$\sum_j \frac{Y_j \{A_j \neq d(X_j)\}}{\pi(A_j)}$$

more generally, taking $d(x) := \text{sign}(f(x))$, $f : \mathcal{X} \to \mathbb{R}$, minimize over f

$$\sum_j \frac{Y_j \{A_j \neq \cancel{\text{sign}}\,\sigma(f(x))\}}{f(A_j)}$$

"when solving a given problem, try to avoid solving a more general problem as an intermediate step" Zhou et al. 2012 quoting Vapnik (in another context - they make this same generalization)

done

computationally better to use a convex function eg ~~positive part~~ $\phi(\cdot) := (\cdot)^+$ or "hinge loss"

to approximate the indicator $\{A_j \neq \sigma(f(x))\}$:

$$\sum_j \frac{Y_j \, (1 - A_j f(X_j))^+}{f(A)} + \lambda \|f\|^2$$

and add regularizer
λ a tuning parameter governing strength

SVM (Vapnik et al '90s) objective : ~~function~~

$$\arg\min_{f \in \mathcal{F}} \sum_j (1 - A_j f(x_j))^+ + \lambda \|f\|^2$$

OWL differs in adding weights to the loss terms
Theoretical results of Zhou '12 parallel main SVM results

QP solvers take the dual ~~problem~~ optimization problem

$$\arg\max_\alpha \sum \alpha_j - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j A_i A_j (x_i, x_j)$$

#6

Theorems (Zhao '12)

given decision fn f define risk as the expected loss under our objective

$$R(f) := \mathbb{E}\left(\frac{Y}{P(A)} \{A \neq \cancel{\text{sgn}}(f(X))\} \right)$$

$$R(f) = \mathbb{E}\left(\mathbb{E}\left(\frac{Y}{P(A)} \{A \neq \sigma(f(x))\} \,\middle|\, X \right) \right)$$

$$= \mathbb{E}\left(\mathbb{E}(Y \mid X, A=1) \{\sigma(f(x)) \neq 1\} + \mathbb{E}(Y \mid X, A=-1) \{\sigma(f(x)) \neq -1\} \right)$$

$$= \mathbb{E}\left(\mathbb{E}(Y \mid X, A=1) \{\sigma(f(x)) = -1\} + \mathbb{E}(Y \mid X, A=-1) \{\sigma(f(x)) = 1\} \right)$$

f^* s.t. $\sigma(f^*(x)) = \sigma\left(\mathbb{E}(Y \mid X, A=1) - \mathbb{E}(Y \mid X, A=0) \right)$ ("bayes optimal classifier")

<u>Thm</u> The minimizer of the hinge loss risk

$$R_\phi(f) = \mathbb{E}\left(\frac{Y}{P(A)} (1 - A f(x))^+ \right)$$

is the bayes optimal classifier

<u>Thms</u> consistency and rates of convergence (cf Steinwart & Scovel and Tsybakov 2005)

Inference on parameters of f? ⇒ go to code

Theorem 3.3. *Assume that we choose a sequence $\lambda_n > 0$ such that $\lambda_n \to 0$ and $\lambda_n n \to \infty$. Then for all distributions P, we have that in probability,*

$$\lim_{n\to\infty} \left\{ \mathcal{R}_\phi(\hat{f}_n) - \inf_{f \in \bar{\mathcal{H}}_k} \mathcal{R}_\phi(f) \right\} = 0,$$

where $\bar{\mathcal{H}}_k$ denotes the closure of $\mathcal{H}_k$. Thus, if f^* belongs to the closure of $\limsup_{n\to\infty} \mathcal{H}_k$, where $\mathcal{H}_k$ can potentially depend on n, we have $\lim_{n\to\infty} \mathcal{R}_\phi(\hat{f}_n) = \mathcal{R}_\phi^*$ in probability. It then follows that $\lim_{n\to\infty} \mathcal{R}(\hat{f}_n) = \mathcal{R}^*$ in probability.

Theorem 3.4. *Let P be a distribution of (X, A, R) satisfying condition (3.4) with noise exponent $q > 0$. Then for any $\delta > 0, 0 < \nu < 2$, there exists a constant C (depending on ν, δ, d and π) such that for all $\tau \geq 1$ and $\sigma_n = \lambda_n^{-1/(q+1)d}$,*

$$Pr^*(\mathcal{R}(\hat{f}_n) \leq \mathcal{R}^* + \epsilon) \geq 1 - e^{-\tau},$$

where Pr^* denotes the outer probability for possibly nonmeasurable sets, and

$$\epsilon = C\left[\lambda_n^{-\frac{2}{2+\nu}+\frac{(2-\nu)(1+\delta)}{(2+\nu)(1+q)}} n^{-\frac{2}{2+\nu}} + \frac{\tau}{n\lambda_n} + \lambda_n^{\frac{q}{q+1}}\right].$$

In particular, when data are well separated, q can be sufficiently large and we can let (δ, ν) be sufficiently small. Then the convergence rate almost achieves the rate $n^{-1/2}$. However, if the

X^+

$\mathbb{E}(R \mid X=x, A=1) - \mathbb{E}(R \mid X=x, A=0) = 0$

X^-