$Y = \alpha D + \beta' W + \epsilon$ → the TRUE GDP

$\tilde{Y} = Y - \gamma'_{\dots} W$; $\tilde{D} = D - \gamma_{D\dots} W$

Let's propose: $\tilde{Y} = a\tilde{D} + u$

$\therefore \quad \alpha = \arg\min_{a \in \mathbb{R}} E(\tilde{Y} - a\tilde{D})^2 = E(\tilde{D}^2)^{-1} E(\tilde{D}Y)$

α can be recovered from the population linear regression of $\tilde{Y}$ on $\tilde{D}$.

Note that $a = \alpha$ solves the equation:

$$\boxed{E(\tilde{Y} - a\tilde{D})\tilde{D} = 0}$$

So if we have random sample $(Y_i, X_i)_{i=1}^{n}$, we need to mimic in the Partialling-out procedure in the population.

NEYMAN ORTHOGONALITY

α → Target Parameter

η → nuisance Projection parameters with True Value

$$\eta^0 = (\gamma'_{DW}, \gamma'_{YW})'$$

the learned value of "α" depends on the values of the nuisance parameters. It is useful to consider the dependence of α on the nuisance parameters!

$$\alpha = \alpha(\eta)$$

Main idea of Double lasso:-

learning "α" that is first order insensitive to local perturbations of the nuisance parameters around their true values, η^0:

$$D = \partial_\eta \alpha(\eta^0) = 0$$

"α" is $\boxed{\text{LOCALLY INSENSITIVE}}$ to perturbations of the nuisance parameters $\boxed{\text{AROUND}}$ their true values.

Double lasso exploits the empirical analogue of the following moment condition:

$$M(a, \eta) = E\left[(\tilde{Y}(\eta_1) - a\tilde{D}(\eta_2))\tilde{D}(\eta_2)\right] = 0$$

the true parameter $\boxed{a = \alpha}$ solves this equation when:

$$\eta := (\eta_1', \eta_2') = \eta^0 := (\gamma_{YW}', \gamma_{DW}')'$$

where: $\tilde{Y}(\eta_1) = Y - \eta_1' W \; ; \; \tilde{D}(\eta_2) = D - \eta_2' W$

the true residualized Quantities correspond to $\eta := \eta^0$

$$\tilde{Y} = Y - \gamma_{YW}' W \; ; \; \tilde{D} = D - \gamma_{DW}' W$$

BY THE IMPLICIT FUNCTION theorem!

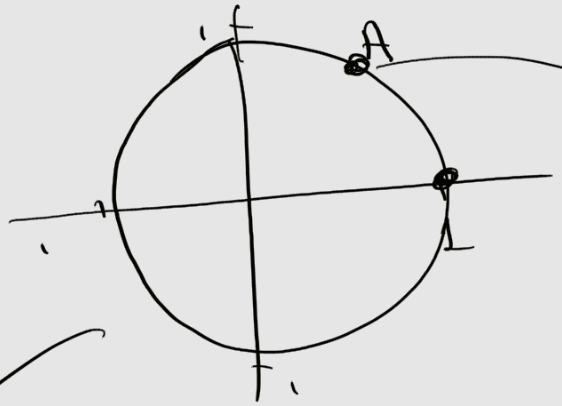
$$D = \partial_a M(\alpha, \eta^0)^{-1} \partial_\eta M(\alpha, \eta^0)$$

and

$$\partial_\eta M(\alpha, \eta^0) \begin{cases} \partial_{\eta_1} M(\alpha, \eta^0) = E(W\tilde{D}) = 0 \\ \partial_{\eta_2} M(\alpha, \eta^0) = -E[W\tilde{Y}] + 2E[\alpha W \tilde{D}] = 0 \end{cases}$$

BUT WHAT IS THE IMPLICIT FUNCTION THEOREM?

If we have $f(x,y) = x^2+y^2$ then $f(x,y) = 1$ → this is the unit circle



around point "A" ↳ y can be expressed as a function of x

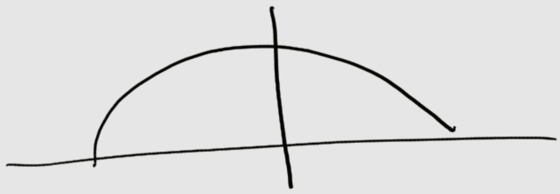
$$g_1(x) = \sqrt{1-x^2}$$

Using this circle there is no way to represent y as a function of x

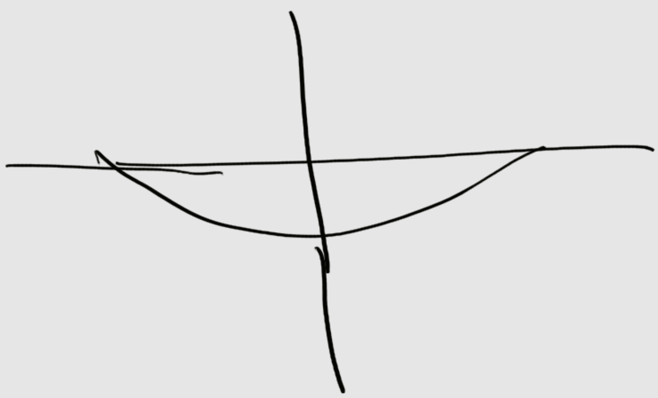
for each choice of $x \in (-1,1)$ → there are two choices of y → $\pm\sqrt{1-x^2}$

$\boxed{\text{BUT!}}$ → In a PART of the circle there a way to represent y as function of x

Example 1 $\quad g_1(x) = \sqrt{1-x^2}$ for $-1 \le x \le 1$



Example 2 $\quad g_2(x) = -\sqrt{1-x^2}$ for $-1 \le x \le 1$



the purpose of IFT is to Tell us evidence of $g_1(x)$ ^ $g_2(x)$ even when we can not write explicit formulas, it garantees that g_1 ^ g_2 are differentiable even when we do not have a formula for $f(x,y)$

Proof for 2D CASE

Suppose $F: \mathbb{R}^2 \to \mathbb{R}$ is a continuously differentiable function defining a curve $F(z) = F(x,y) = 0$
Let (x_0, y_0) be a point in the curve.

Theorem → If $\left.\frac{\partial F}{\partial y}\right|_{(x_0,y_0)} \neq 0$

then for The curve around (x_0, y_0) we can write $y = f(x)$ when f is a real function

We take differ...
$$dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy = 0$$

we can express:
$$\frac{\partial F}{\partial y} dy = -\left(\frac{\partial F}{\partial x} dx\right)$$

$$\boxed{\frac{dy}{dx} = -\left[\frac{\partial F}{\partial y}\right]^{-1}\left(\frac{\partial F}{\partial x}\right)}$$

This function is around (x_0, y_0)

* Now we use the ITF to proof Neyman Orthogonality
now we have new variables, but the process is the same

$$\left.\begin{aligned} F &\xrightarrow{in} M \\ (X,Y) &\xrightarrow{in} (a,\eta) \\ (X_0,Y_0) &\xrightarrow{in} (\alpha,\eta^0) \end{aligned}\right\} M(a,\eta) = E\left[\left(\tilde{Y}(\eta_1) - \tilde{D}(\eta_2)\right)\tilde{D}(\eta_2)\right] = 0$$

$\wedge\ \eta^0 = (\gamma_{yw}, \gamma_{Dw})$

So using $\boxed{\text{ITF}}$

$$\frac{\partial M}{\partial a}(\alpha,\eta^0)\,da + \frac{\partial M}{\partial \eta}(\alpha,\eta^0)\,d\eta = 0$$

$$\therefore \boxed{\frac{da}{\partial \eta} = -\underbrace{\left[\frac{\partial M}{\partial a}(\alpha,\eta_0)\right]^{-1}}_{S_1}\underbrace{\left[\frac{\partial M}{\partial \eta}(\alpha,\eta^0)\right]}_{S_2}}$$

the second component of the equation above is:

$$\left[\frac{\partial M}{\partial \eta}(\alpha,\eta_0)\right] = \frac{\partial M}{\partial \eta_1}(\alpha,\eta_0) + \frac{\partial M}{\partial \eta_2}(\alpha,\eta_0)$$

$$= \underbrace{\frac{\partial M}{\partial \eta_1}(\alpha,\gamma'_{yw},\gamma'_{Dw})}_{S_1} + \underbrace{\frac{\partial M}{\partial \eta_2}(\alpha,\gamma'_{yw},\gamma'_{Dw})}_{S_2}$$

$$S_1 = \left.\frac{\partial M}{\partial \eta_1}\right|_{\alpha,\gamma'_{yw},\gamma'_{Dw}} = \frac{\partial E\left[\left[\tilde{Y}(\eta_1) - \tilde{D}(\eta_2)\right]\tilde{D}(\eta_2)\right]}{\partial \eta_1}$$

$$= \frac{\partial E\left[\tilde{Y}(\eta_1)\tilde{D}(\eta_2)\right]}{\partial \eta_1} \wedge\ \tilde{Y}(\eta_1) = Y - \eta_1 W$$

$$S_1 = E\left[(W)\, \tilde{D}(\eta_2) \right] \Big|_{(\alpha, \gamma_{YW}, \gamma_{DW})}$$

$$= E\left[(W)(D - \eta_2' W) \right] \Big|_{(\alpha, \gamma_{YW}', \gamma_{DW}')}$$

$$= E\left[(W)\,(D - \gamma_{DW}' W) \right]$$

$$= E\left[WD - W \gamma_{DW}' W \right]$$

$$= E\left[WD - W \left[(W'W)^{-1} (W'D) \right]' W \right]$$

$$= E\left[WD - WD \right] = 0$$

SIMILAR WITH S_2.

$$S_2 = \frac{\partial M}{\partial \eta_2} \Big|_{(\alpha, \gamma_{YW}', \gamma_{DW}')} = -E[W \tilde{Y}] + 2E[\alpha W \tilde{D}] = 0$$

$$\therefore \quad \frac{\partial M}{\partial \eta} \Big|_{(\alpha, \eta_0) = (\alpha, \gamma_{YW}', \gamma_{DW}')} = 0$$

then $\boxed{\frac{\partial a}{\partial \eta} = 0 \quad \vee \quad \frac{\partial \alpha}{\partial \eta}(\eta^0) = 0}$

For work group!

Do again all the derivation/proof of the Neyman orthogonality condition! please show why S_2 is $= 0$!, you need to use matrix algebra.