Robust Weighting and Matching Techniques for Causal Inference in Observational Studies with Continuous Treatment

Universität Stuttgart

Ioan Scheffel

October 22, 2022

Contents

Chapter 1

Chapter One Title

hello $\mathbb{R}$

Chapter 2

Convex Analysis

We begin by defining convex sets

Definition 1. *A subset* $\Omega \subseteq \mathbb{R}^n$ *is called CONVEX if we have* $\lambda x + (1-\lambda)y \in \Omega$ *for all* $x, y \in \Omega$ *and* $\lambda \in (0,1)$.

Clearly, the line segment $[a,b] := \{\lambda a + (1-\lambda)b \mid \lambda \in [0,1]\}$ is contained in Ω for all $a, b \in \Omega$ if and only if Ω is a convex set.

Next we define convex functions.

The concept of convex functions is closely related to convex sets.

The line segment between two points on the graph of a convex function lies on or above and does not intersect the graph.

In other words: The area above the graph of a convex function f is a convex set, i.e. the *epigraph* $\mathrm{epi}(f) := \{(x,\alpha) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq \alpha\}$ is a convex set in $\mathbb{R}^{n+1}$.

Often an equivalent characterisation of convex functions is more useful.

Theorem 1. *The convexity of a function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *on* $\mathbb{R}^n$ *is equivalent to the following statement:*

For all $x, y \in \mathbb{R}^n$ *and* $\lambda \in (0,1)$ *we have*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y). \tag{2.1}$$

Chapter 3

Random Matrix Inequality

Theorem 2. *Let* $(A_k)_{1\leq k\leq n} \subseteq \mathbb{R}^{d_1\times d_2}$ *be a finite sequence of independent, random matrices. Assume that*

$$\mathbb{E}(A_k) = 0 \quad \textit{and} \quad \|A_k\| \leq L \quad \textit{for each} \quad k \in \{1,\dots,n\}. \tag{3.1}$$

Introduce the random matrix

$$S := \sum_{k=1}^{n} A_k. \tag{3.2}$$

Let $v(S)$ *be the matrix variance statistic of the sum:*

$$\begin{aligned} v(S) :&= \max\left\{\left\|\mathbb{E}(SS^T)\right\|, \left\|\mathbb{E}(S^TS)\right\|\right\} && (3.3)\\ &= \max\left\{\left\|\sum_{k=1}^{n}\mathbb{E}(A_kA_k^T)\right\|, \left\|\sum_{k=1}^{n}\mathbb{E}(A_k^TA_k)\right\|\right\}. && (3.4)\end{aligned}$$

Then

$$\mathbb{E}\|S\| \leq \sqrt{2v(S)\log(d_1+d_2)} + \frac{1}{3}L\log(d_1+d_2). \tag{3.5}$$

Furthermore, for all $t \geq 0$,

$$\mathbb{P}(\|S\| \geq t) \geq (d_1+d_2)\exp\left(\frac{-t^2/2}{v(S)+Lt/3}\right). \tag{3.6}$$

Chapter 4

Simple yet useful Calculations

Proposition 1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuous such that a minimum x^* exists and is unique. Then for all $y \in \mathbb{R}^n$ and $C > 0$ it follows*

$$\inf_{\|\Delta\|=C} f(y+\Delta) - f(y) > 0 \qquad \Rightarrow \qquad \|x^* - y\| \leq C. \tag{4.1}$$

Proof. Since $\mathcal{C} := \{\|\Delta\| \leq C\}$ is compact and

$$f(x^*) \leq f(y) < \inf_{\|\Delta\|=C} f(y+\Delta)$$

the continious function $f(y + \,\cdot\,)$ has a minimum in $\mathring{\mathcal{C}} := \{\|\Delta\| < C\}$. Since x^* is the unique minimum of f there exists $\Delta^* \in \mathring{\mathcal{C}}$ such that $x^* - y = \Delta^*$. We conclude that $\|x^* - y\| \leq C$. $\square$

Proposition 2. *Let $f \in C^2(\mathbb{R})$. Then for all $a, x, \Delta \in \mathbb{R}^n$ there exist $\xi_1, \xi_2 \in (0, 1)$ such that it holds*

$$f(a^T(x+\Delta)) - f(a^T x) = f^{'}(a^T x)\,\Delta^T a + \xi_1 f^{''}(a^T(x + \xi_1\xi_2\Delta))\,\Delta^T A\ \Delta, \tag{4.2}$$

where $A := aa^T \in \mathbb{R}^{n\times n}$.

Proof. By the chain rule we have for all $a, x, \Delta \in \mathbb{R}^n$ and $\xi \in (0, 1)$

$$\nabla_x f(a^T(x+\xi\Delta)) = f^{'}(a^T(x+\xi\Delta))\,a. \tag{4.3}$$

By the mean value theorem and (4.3) there exist $\xi_1, \xi_2 \in (0,1)$ such that

$$\begin{aligned} f(a^T(x+\Delta)) - f(a^T x) &= \Delta^T \nabla_x f(a^T(x+\xi_1\Delta)) \\ &= f^{'}(a^T(x+\xi_1\Delta))\, \Delta^T a. \end{aligned} \tag{4.4}$$

and

$$f^{'}(a^T(x+\xi_1\Delta)) - f^{'}(a^T x) = \xi_1 f^{''}(a^T(x+\xi_1\xi_2\Delta))\, a^T\Delta. \tag{4.5}$$

Plugging (4.5) in (4.4) yields (4.2). $\square$