شبکه‌های کانولوشنی مستقل از مختصات

کانولوشن‌های یکسان نسبت به ایزومتری و پیمانه روی منیفلدهای ریمانی

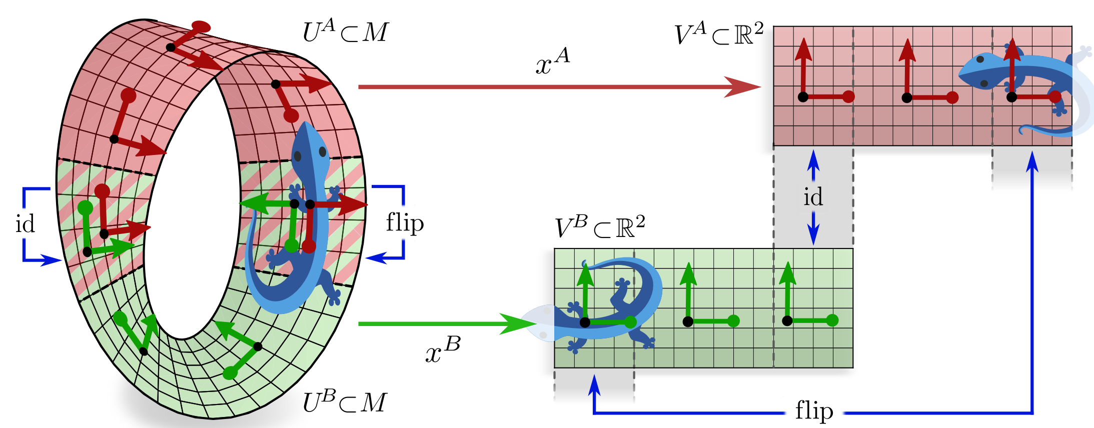
موریس وایلر[1] پاتریک فوره[1] اریک ورلینده[1] مکس ولینگ[1,2]

[1]دانشگاه آمستردام
[2]پژوهشگاه هوش مصنوعی کوالکام

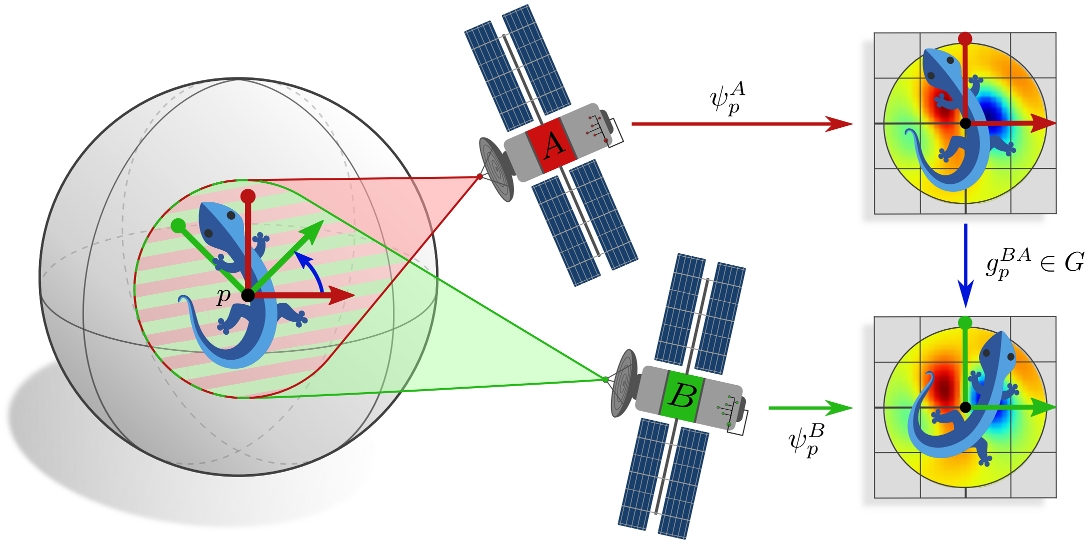
p.d.forre@uva.nl m.weiler.ml@gmail.com
m.welling@uva.nl e.p.verlinde@uva.nl

□□□□□□□□

انگیزه‌ای از موفقیت گسترده شبکه‌های عمیق کانولوشنی، علاقه زیادی برای تعمیم کانولوشن‌ها به منیفلدهای غیراقلیدسی وجود دارد. یک پیچیدگی عمده در مقایسه با فضاهای مسطح این است که مشخص نیست کرنل کانولوشن باید در کدام تراز روی یک منیفلد اعمال شود. دلیل اساسی این ابهام آن است که منیفلدهای عمومی دارای انتخاب متعارف چارچوب‌های مرجع (گیج) نیستند. بنابراین کرنل‌ها و ویژگی‌ها باید نسبت به مختصات دلخواه بیان شوند. ما استدلال می‌کنیم که انتخاب خاص مختصات‌بندی نباید بر استنتاج شبکه تأثیر بگذارد □ آن باید مستقل از مختصات باشد. تقاضای همزمان برای استقلال مختصات و اشتراک وزن منجر به الزامی روی شبکه می‌شود تا تحت تبدیل‌های گیج محلی (تغییرات چارچوب‌های مرجع محلی) تناوب‌پذیر باشد. ابهام چارچوب‌های مرجع بدین‌گونه به G-ساختار منیفلد بستگی دارد، به طوری که سطح لازم تناوب‌پذیری گیج توسط گروه ساختار G متناظر تجویز می‌شود. کانولوشن‌های مستقل از مختصات ثابت می‌شوند که نسبت به آن ایزومتری‌هایی که تقارن‌های G-ساختار هستند تناوب‌پذیر باشند. نظریه حاصل به شکل آزاد از مختصات بر حسب بندل‌های فیبر فرمول‌بندی می‌شود. برای نمونه‌سازی طراحی کانولوشن‌های مستقل از مختصات، ما شبکه کانولوشنی روی نوار موبیوس پیاده‌سازی می‌کنیم. عمومیت فرمول‌بندی هندسه دیفرانسیل ما از شبکه‌های کانولوشنی با بررسی گسترده ادبیات نشان داده می‌شود که تعداد زیادی از شبکه‌های کانولوشنی اقلیدسی، شبکه‌های کانولوشنی کروی و شبکه‌های کانولوشنی روی سطوح عمومی را به عنوان نمونه‌های خاص کانولوشن‌های مستقل از مختصات توضیح می‌دهد.



شکل ۱: نمایش کانولوشن روی نوار موبیوس با گیج‌های مختلف



شکل ۲: مشاهده‌گران مختلف A و B ممکن است یک الگوی از ویژگی‌ها را از «دیدگاه» متفاوتی درک کنند. ماهواره‌ها در کاربرد ما کرنل‌های کانولوشنی هستند که میدان دید محلی خود را در اطراف p به یک بردار ویژگی در p خلاصه می‌کنند. «دیدگاه» آن‌ها انتخاب یک چارچوب مرجع محلی (گیج) در p است که کرنل بر اساس آن تراز می‌شود. از آنجا که مشاهدات هر دو دیدگاه نشان‌دهنده یک الگوی یکسان هستند، پاسخ‌های کرنل باید حاوی اطلاعات معادل باشند، یعنی استنتاج باید مستقل از مختصات باشد. این امر کرنل‌های کانولوشنی را ملزم می‌کند تا تحت تبدیل‌های گیج محلی (یعنی تغییرات چارچوب‌های مرجع) تناوب‌پذیر باشند. سطح تناوب‌پذیری گیج توسط گروه ساختار G تعیین می‌شود، که هم به منیفولد و هم به کاربرد بستگی دارد.

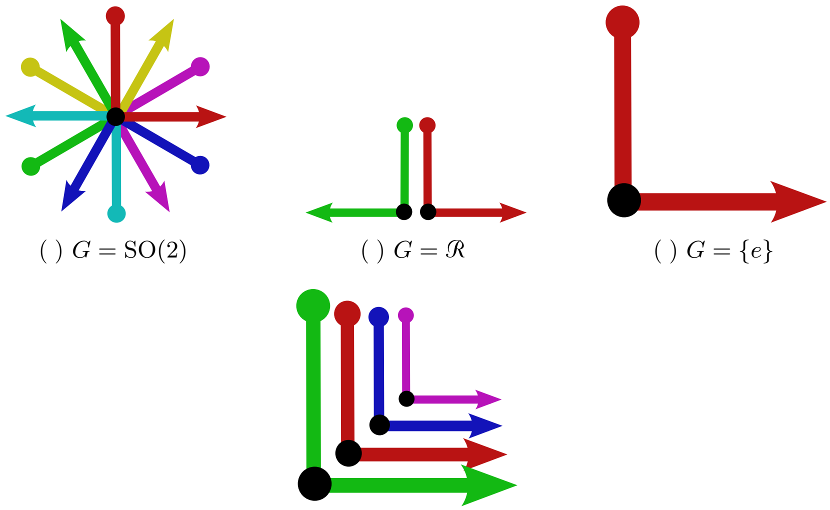
۱ مقدمه

در سال‌های اخیر، شبکه‌های عصبی عمیق به مدل‌های منتخب برای طیف گسترده‌ای از وظایف در یادگیری ماشین تبدیل شده‌اند. موفقیت مدل‌های عمیق اغلب ریشه در طراحی وظیفه–خاص دارد که منعکس‌کننده ساختار ریاضی داده‌های مورد پردازش است. یک مثال برجسته شبکه‌های کانولوشنی (CNN) هستند که از طریق اتصال محلی و اشتراک وزن فضایی، ساختار مکانی داده‌ها را بهره‌برداری می‌کنند. از آنجا که یک کرنل یکسان در هر نقطه از فضا اعمال می‌شود، شبکه‌های کانولوشنی نسبت به انتقال تناوب‌پذیر هستند، به این معنی که الگوهای آموخته‌شده را به طور خودکار بر روی موقعیت‌های فضایی تعمیم می‌دهند. با توجه به موفقیت تجربی قابل توجه شبکه‌های کانولوشنی اقلیدسی، علاقه زیادی به تعمیم مدل‌های کانولوشنی برای پردازش سیگنال‌ها در دامنه‌های عمومی‌تر و تناوب‌پذیر ساختن آن‌ها تحت گروه‌های تقارن بزرگ‌تر وجود دارد.

این کار به بررسی تعمیم شبکه‌های کانولوشنی به منیفلدهای ریمانی می‌پردازد. یک پیچیدگی عمده در تعمیم شبکه‌های کانولوشنی از فضاهای اقلیدسی $\mathbb{R}^d$ به منیفلدهای ریمانی عمومی این است که منیفلدهایی دارای انتخاب مرجع جهت ارجح نیستند که کرنل کانولوشن بتواند بر اساس آن برای اندازه‌گیری ویژگی‌ها تراز شود. از آنجا که هیچ جهت مرجعی ترجیح داده نمی‌شود، کرنل باید به صورت دلخواه بر روی منیفولد تراز شود. موضوع اصلی این کار تنظیم این اختیار با مستقل ساختن استنتاج شبکه‌ها از تراز خاص کرنل‌های کانولوشن است. مشخص می‌شود که این امر مستلزم آن است که کرنل‌ها تناوب‌پذیر گیج باشند، یعنی تحت تبدیل‌های تراز کرنل، تناوب‌پذیر باشند. پاسخ یک کرنل تناوب‌پذیر گیج، هنگام تغییر تراز آن، به طور قابل پیش‌بینی تغییر می‌کند – بنابراین محتوای اطلاعات استخراج شده برای هر انتخاب (دلخواه) تراز، یکسان تضمین می‌شود.

ما تراز یک کرنل در نقطه‌ای p از یک منیفولد M را به عنوان انتخاب یک چارچوب مرجع محلی □ یا گیج □ از فضای مماس مربوطه T_pM رسمی‌سازی می‌کنیم. بنابراین تبدیل‌های گیج، تبدیلاتی بین انتخاب‌های چارچوب‌های مرجع هستند. شکل ۵ مفهوم تراز کردن کرنل‌ها در امتداد چارچوب‌های مرجع را بصری‌سازی می‌کند. تراز کردن کرنل نسبت به میدان چارچوب کانونی (منحصر به فرد ارجح) صفحه اقلیدسی $\mathbb{R}^۲$، همانطور که در بالا نشان داده شده است، منجر به میدان کرنل معمول شبکه‌های کانولوشنی اقلیدسی می‌شود. یک میدان چارچوب متفاوت، همانطور که در پایین نشان داده شده است، یک میدان کرنل و در نتیجه یک شبکه جایگزین را نشان می‌دهد. همانطور که در بالا ذکر شد، در اکثر منیفلدهایی، انتخاب چارچوب‌ها ذاتاً مبهم است به طوری که هیچ تراز کرنل خاصی ترجیح داده نمی‌شود. شکل ۲ این مشکل را برای کره $S^۲$ بصری‌سازی می‌کند، جایی که چارچوب‌ها فقط تا چرخش‌ها منحصر به فرد هستند.

سطح ابهام در انتخاب چارچوب‌های مرجع به ساختار هندسی منیفولد بستگی دارد. چنین ساختاری اغلب امکان رفع ابهام چارچوب‌های مرجع تا تبدیل‌های تقارنی خاص (تبدیل‌های گیج) را فراهم می‌کند؛ به شکل ۳ مراجعه کنید. این گزاره با چند مثال بهتر توضیح داده می‌شود:



( ) $G = \mathrm{SO}(2)$ ( ) $G = \mathscr{R}$ ( ) $G = \{e\}$

( ) $G = \mathscr{S}$

شکل ۳: انتخاب چارچوب‌های مرجع یک فضای مماس T_pM همیشه منحصر به فرد نیست. ساختار هندسی (G-ساختار) یک منیفولد، زیرمجموعه‌ای ترجیحی از چارچوب‌های مرجع را نشان می‌دهد به طوری که تبدیل‌های گیج بین این چارچوب‌ها در گروه ساختار $G \leq \mathrm{GL}(d)$ قرار می‌گیرند. اشکال ??، ??، ?? و ?? چنین زیرمجموعه‌هایی از چارچوب‌ها را به ترتیب برای گروه بدیهی $G = \{e\}$، گروه بازتاب $G = \mathscr{R}$، گروه چرخش $G = \mathrm{SO}(۲)$ و گروه مقیاس $G = \mathscr{S}$ نشان می‌دهند. ویژگی‌ها، اندازه‌گیری‌ها را نسبت به هر یک از چارچوب‌های متمایز کدگذاری می‌کنند. ضرایب عددی آن‌ها نسبت به چارچوب‌های مختلف با عمل یک نمایش گروهی ρ از G مرتبط هستند.

- یک منیفولد هموار هیچ اولویتی در انتخاب چارچوب‌ها ندارد. تبدیل‌های گیج بین چارچوب‌های عمومی، نگاشت‌های خطی معکوس‌پذیر دلخواه هستند، یعنی مقادیری در گروه خطی عمومی $G = \mathrm{GL}(d)$ می‌گیرند.
- یک جهت‌گیری منیفولد امکان تشخیص چارچوب‌های راست‌گرد از چپ‌گرد را می‌دهد. تبدیل‌های گیج بین چارچوب‌های هر دو دست، جهت‌گیری را حفظ می‌کنند، یعنی اعضایی از $G = \mathrm{GL}^+(d)$ هستند (نگاشت‌های خطی معکوس‌پذیر با دترمینان مثبت).
- یک فرم حجم امکان تشخیص چارچوب‌های واحد حجم را می‌دهد. تبدیل‌های گیج سپس حجم را حفظ می‌کنند، یعنی مقادیری در گروه خطی خاص $G = \mathrm{SL}(d)$ می‌گیرند.
- ساختار متریک یک منیفولد ریمانی امکان اندازه‌گیری فواصل و زوایا در فضاهای مماس را فراهم می‌کند و بنابراین امکان تشخیص چارچوب‌های متعامد را می‌دهد. تبدیل‌های گیج بین چارچوب‌های متعامد، چرخش‌ها و بازتاب‌ها در گروه متعامد $G = d$ هستند.
- همچنین، یک جهت‌گیری و متریک به معنای چارچوب‌های متعامد جهت‌دار است. تبدیل‌های گیج سپس فقط چرخش‌ها در گروه متعامد خاص $G = \mathrm{SO}(d)$ هستند.
- یک میدان چارچوب بر روی منیفولد شامل یک چارچوب منحصر به فرد در هر نقطه از منیفولد است. تبدیل‌های گیج در این حالت بدیهی هستند، که توسط گروه بدیهی $G = \{e\}$ توصیف می‌شود.

همه این ساختارهای هندسی یک ویژگی مشترک دارند که یک زیرمجموعه (زیربندل) ترجیحی از چارچوب‌ها را تعریف می‌کنند به طوری که تبدیل‌های گیج مقادیری در یک گروه ساختار $G \leq \mathrm{GL}(d)$ می‌گیرند. برای تأکید بر نقش مرکزی گروه ساختار G، چنین ساختارهایی به عنوان G-ساختارها GM نامیده می‌شوند. مثال‌های بصری از G-ساختارها برای گروه‌های ساختار G و منیفلدهای M مختلف در شکل ?? ارائه شده‌اند.

از آنجا که انتخاب چارچوب‌های مرجع ذاتاً مبهم است، هر کمیت هندسی و عملیات شبکه باید به همان اندازه به خوبی نسبت به چارچوب‌های دلخواه G-ساختار GM قابل نمایش باشند، یعنی باید مستقل از مختصات GM باشند. بنابراین بردارهای ویژگی با یک نمایش گروهی (عمل گروهی خطی) ρ از گروه ساختار G مرتبط هستند که قانون تبدیل آن‌ها را تحت تبدیل‌های گیج (انتقال‌های با مقدار G بین چارچوب‌های مرجع) تعیین می‌کند. انتخاب خاص نمایش گروهی، نوع هندسی میدان بردار ویژگی را تعیین می‌کند. مثال‌های معمول شامل میدان‌های اسکالر، بردار یا تانسور هستند، با این حال، انواع میدان‌های عمومی‌تر نیز در عمل استفاده می‌شوند. شکل ۶ استقلال مختصات کمیت‌های هندسی را در مثال شناخته شده بردارهای مماس بصری‌سازی می‌کند.

هر لایه شبکه باید قوانین تبدیل ویژگی‌ها را رعایت کند، یعنی باید تضمین کند که خروجی‌های آن طبق انتظار تبدیل می‌شوند. به طور خاص برای کانولوشن‌ها، استقلال مختصات GM ایجاب می‌کند که اعمال کرنل مشترک نسبت به چارچوب‌های مختلف G-ساختار در نقطه‌ای $p \in M$ باید

همان پاسخ را تا یک تبدیل گیج ایجاد کند. ما نشان می‌دهیم که این امر نیازمند G–استیریبل بودن (تناوب‌پذیری گیج، معادله (۸۶)) کرنل‌های کانولوشن است. به طور شهودی، می‌توان کرنل‌های G–استیریبل را به عنوان اندازه‌گیری ویژگی‌ها نسبت به چارچوب‌های مرجع در نظر گرفت، که ضروری است زیرا هیچ انتخابی از چارچوب، یعنی تراز کرنل مطلق، نباید ترجیح داده شود.[1] مثال‌هایی از کرنل‌های G–استیریبل برای گروه بازتاب $G = \mathcal{R}$ در شکل ؟؟ نشان داده شده‌اند. شکل ؟؟ اشتراک چنین کرنل‌هایی را نسبت به چارچوب‌های مختلف یک $\mathcal{R}$–ساختار بصری‌سازی می‌کند. قید G–استیریبل بودن، تقارن‌هایی را بر کرنل‌ها تحمیل می‌کند، به طوری که ترازهای مختلف در واقع منجر به پاسخ‌هایی می‌شوند که دقیقاً با تبدیل‌های گیج $\rho(g)$ تفاوت دارند. ما کانولوشن‌های مستقل از مختصات GM را در ادامه به اختصار GM–کانولوشن‌ها می‌نامیم.

علاوه بر اعمال کرنل‌های تناوب‌پذیر گیج، GM–کانولوشن‌ها ممکن است تناوب‌پذیر ایزومتری باشند، به این معنی که آن‌ها با عمل ایزومتری‌ها بر میدان‌های ویژگی جابجا می‌شوند، همانطور که در شکل ؟؟ نشان داده شده است. فرض کنید $\phi \in \mathrm{Isom}(M)$ یک ایزومتری (تقارن) از منیفولد M باشد. یک شبکه عصبی دقیقاً زمانی نسبت به عمل این ایزومتری تناوب‌پذیر است که الگوها در هر نقطه $p \in M$ به همان شیوه الگوها در $\phi(p)$ پردازش شوند. تناوب‌پذیری ایزومتری یک شبکه، بنابراین با ناوردا بودن ایزومتری اتصال عصبی (میدان کرنل) آن مطابقت یک به یک دارد؛ به شکل ؟؟ مراجعه کنید. از آنجا که کانولوشن‌های ما کرنل‌ها را نسبت به چارچوب‌های (دلخواه) G–ساختار GM اعمال می‌کنند، تقارن‌های میدان کرنل با تقارن‌های G–ساختار منطبق هستند. با نشان دادن تقارن‌های یک G–ساختار GM (حفظ‌کننده فاصله) به عنوان $\mathrm{Isom}_{GM} \leq \mathrm{Isom}(M)$، این بدان معنی است که کانولوشن‌های ما دقیقاً Isom_{GM}–تناوب‌پذیر هستند. شکل ؟؟ این واقعیت را بصری‌سازی می‌کند که G–ساختارها و میدان‌های کرنل مربوطه دارای تقارن‌های یکسانی هستند. خواننده تشویق می‌شود که G–ساختارهای شکل ؟؟ را با توجه به تقارن‌هایشان و ویژگی‌های تناوب‌پذیری ضمنی کانولوشن‌های GM مربوطه بررسی کند.

طراحی شبکه‌های کانولوشنی مستقل از مختصات GM در منیفلدهای ریمانی نیازمند انتخاب یک G–ساختار است که به ملاحظات متعددی بستگی دارد. اولاً، انتخاب گروه ساختار G تناوب‌پذیری گیج محلی کانولوشن را تعیین می‌کند: یک کرنل G–استیریبل به طور خودکار الگوهای آموخته‌شده را بر روی تمام حالت‌های الگوهای مرتبط با G تعمیم می‌دهد؛ به شکل ؟؟ مراجعه کنید. ثانیاً، انتخاب خاص G–ساختار تناوب‌پذیری ایزومتری سراسری کانولوشن را تعیین می‌کند. در کاربردهای تصویربرداری پزشکی، الگوها اغلب در چرخش‌ها، بازتاب‌ها و موقعیت‌های دلخواه ظاهر می‌شوند □ بنابراین باید یک d–ساختار ناوردا با $\mathrm{Isom}_{GM} = \mathrm{E}(d)$ بر روی $\mathbb{R}^d$ انتخاب کرد، مشابه $\mathrm{SO}(۲)$–ساختار که در شکل ؟؟ نشان داده شده است. تصاویر مانند عکس‌های پرتره یک محور عمودی متمایز دارند، با این حال، بازتاب‌ها در امتداد این محور، آمار تصویر را ناوردا می‌گذارند □ این امر به یک $\mathcal{R}$–ساختار همانند شکل ؟؟ نیاز دارد. علاوه بر این ملاحظات تقارنی، توجه به این نکته مهم است که هر منیفولد (توپولوژی) برای هر انتخابی از گروه ساختار G، ساختارهای G هموار را نمی‌پذیرد. یک مثال نوار موبیوس است که توپولوژی پیچ‌خورده (غیرجهت‌پذیری) آن مانع از اختصاص هموار جهت‌گیری‌های چارچوب می‌شود. بنابراین یک عملیات کانولوشن مستقل از مختصات هموار بر روی نوار موبیوس لزوماً بر کرنل‌های استیریبل بازتابی تکیه دارد.

این کار شامل یک بررسی گسترده ادبیات در مورد شبکه‌های کانولوشنی است که عمومیت نظریه ما را نشان می‌دهد. این کار شامل انواع مختلف شبکه‌های کانولوشنی در فضاهای اقلیدسی، شبکه‌های کانولوشنی کروی و شبکه‌های کانولوشنی روی سطوح عمومی (مانند شبکه‌های سطحی) است. ما انتخاب‌های خاص G–ساختارها را که به طور ضمنی توسط نویسندگان با تجزیه و تحلیل ویژگی‌های تناوب‌پذیری سراسری و محلی مدل‌هایشان انجام شده است، شناسایی می‌کنیم. جدول ؟؟ خلاصه‌ای از طبقه‌بندی حاصل از شبکه‌های کانولوشنی مستقل از مختصات GM را ارائه می‌دهد.

برای ارائه یک مثال دقیق در مورد چگونگی پیاده‌سازی نظریه ما در عمل، یک پیاده‌سازی از GM–کانولوشن‌ها بر روی نوار موبیوس برای $G = \mathcal{R}$ را مورد بحث قرار می‌دهیم. این شامل استخراج کرنل‌های استیریبل بازتابی برای انواع میدان‌های مختلف (نمایش‌های گروهی) و ارزیابی تجربی تناوب‌پذیری ایزومتری پیش‌بینی شده نظری است. همانطور که انتظار می‌رود، کانولوشن‌های مستقل از مختصات GM عملکرد بهتری نسبت به یک پیاده‌سازی ساده وابسته به مختصات دارند. کد در `https://github.com/mauriceweiler/MobiusCNNs` موجود است.

یک فرمول‌بندی بدون مختصات از نظریه ما در زبان بندل‌های فیبر ابداع شده است. G–ساختارها GM زیربندل‌های اصلی G از بندل چارچوب FM بر روی M هستند. میدان‌های ویژگی برش‌هایی از بندل‌های بردار ویژگی مرتبط با G هستند. گیج‌ها تسهیم‌های بندل محلی هستند، در حالی که تبدیل‌های گیج، نگاشت‌های انتقال بین چنین تسهیم‌هایی هستند. ایزومتری‌هایی که یک کانولوشن GM نسبت به آن‌ها تناوب‌پذیر است، اتومورفیسم‌های بندل اصلی از G–ساختار هستند.

شبکه‌های کانولوشنی مستقل از مختصات ما تعمیم‌هایی از شبکه‌های کانولوشنی استیریبل [؟ ؟ ؟ ؟] از فضاهای اقلیدسی (یا همگن) به منیفلدهای ریمانی هستند. در حالی که شبکه‌های کانولوشنی استیریبل بر تبدیل‌های فعال، سراسری میدان‌های ویژگی تمرکز می‌کنند، شبکه‌های کانولوشنی مستقل از مختصات تبدیل‌های منفعل، محلی بین چارچوب‌های مرجع را در نظر می‌گیرند.[2] ما نسخه‌های اولیه نظریه شبکه‌های کانولوشنی مستقل از مختصات ("شبکه‌های کانولوشنی تناوب‌پذیر گیج") را در کارهای قبلی [؟ ؟] پیشنهاد کردیم. برخلاف این انتشارات، کار حاضر نظریه را با جزئیات بسیار بیشتری توسعه می‌دهد، آن را بر حسب بندل‌های فیبر فرمول‌بندی می‌کند، تناوب‌پذیری تحت عمل ایزومتری‌ها را اثبات می‌کند و یک مرور ادبیات ارائه می‌دهد.

[1] به شباهت با اصل نسبیت خاص انیشتین توجه کنید، که به جای چارچوب‌های G–ساختار، بر برابری چارچوب‌های اینرسی تکیه دارد.
[2] این شبیه تغییر تمرکز از کوواریانسی لورنتس سراسری در نسبیت خاص به کوواریانسی لورنتس محلی در نسبیت عام است.

فهرست مطالب

این کار در قالب یک مقدمه، سه بخش اصلی و یک پیوست سازماندهی شده است.

بخش ۱.۱.۲ تلاش می‌کند شبکه‌های عصبی مستقل از مختصات را با زبانی آسان معرفی کند. میدان‌های ویژگی و لایه‌های شبکه نسبت به مختصات محلی (تسهیم‌های بندل) بیان می‌شوند. استقلال مختصات مورد نیاز، مستلزم آن است که ویژگی‌ها با یک قانون تبدیل خاصی مرتبط باشند. لایه‌های شبکه ملزم به تضمین رفتار تبدیل صحیح ویژگی‌ها هستند.

بخش ؟؟ نظریه شبکه‌های عصبی مستقل از مختصات را بر حسب بندل‌های فیبر رسمی‌سازی می‌کند. این امر امکان فرمول‌بندی سراسری و مستقل از مختصات را فراهم می‌کند، که به ویژه هنگام بررسی تناوب‌پذیری ایزومتری شبکه‌ها مفید است. تعاریف از بخش ۱.۱.۲ با بیان عملیات مستقل از مختصات در تسهیم‌های بندل محلی (مختصات) بازیابی می‌شوند.

بخش ؟؟ نظریه ما را در کارهای مرتبط جای می‌دهد. این بخش بررسی‌های مفصلی از معماری‌های شبکه کانولوشنی بر روی هندسه‌های مختلف ارائه می‌دهد و آن‌ها را به عنوان شبکه‌های کانولوشنی مستقل از مختصات فرمول‌بندی مجدد می‌کند. برای تسهیل توسعه معماری‌های شبکه جدید، ما ویژگی‌های مربوط به هندسه‌های خاص را قبل از بررسی شبکه‌هایی که بر روی آن‌ها عمل می‌کنند، مورد بحث قرار می‌دهیم.

خواننده می‌تواند بخش ؟؟ را در مرحله اول نادیده بگیرد فرمول‌بندی از بخش ۱.۱.۲ برای خواندن مرور ادبیات در بخش ؟؟ کاملاً کافی است.

مروری بر مفاهیم و نتایج اصلی کار ما در بخش ۲ زیر ارائه شده است. این مرور از معادلات پرهیز می‌کند و بر شهود هندسی از طریق بصری‌سازی‌ها تکیه دارد. امیدواریم این بخش به مخاطبان غیرتخصصی کمک کند تا ایده‌ای از محتوای کار ما به دست آورند.

مرور کلی جزئیات

بخش ۱.۱.۲: هدف بخش ۳ ابداع فضاهای ویژگی مستقل از مختصات است. به طور خاص، بخش ۱.۳ گیج‌ها، تبدیل‌های گیج و G–ساختارها را معرفی می‌کند. گیج‌ها روشی رسمی برای بیان بردارهای مماس (مستقل از مختصات) و توابع روی فضاهای مماس نسبت به چارچوب‌های مرجع هستند. تبدیل‌های گیج بین این عبارات مختصاتی در گیج‌های مختلف ترجمه می‌کنند. بخش ۲.۳ میدان‌های بردار ویژگی مستقل از مختصات را معرفی می‌کند. همانند مورد بردارهای مماس، ضرایب عددی بردارهای ویژگی هنگام انتقال بین چارچوب‌های مرجع تغییر می‌کنند. قوانین تبدیل بردارهای ویژگی به ویژه انتقال موازی آن‌ها و پیش‌برداری آن‌ها هنگام عمل توسط ایزومتری‌ها را تعیین می‌کنند، که به ترتیب در بخش‌های ۳.۳ و ۴.۳ توضیح داده شده‌اند.

بخش ۴ شبکه‌های عصبی را که بین میدان‌های ویژگی نگاشت می‌کنند، توسعه می‌دهد. عملیات نقطه‌ای، مانند جمع بایاس، کانولوشن‌های 1×1 و غیرخطی‌ها، در بخش ۱.۴ مورد بحث قرار می‌گیرند. بخش ۲.۴ بر روی کانولوشن‌ها با کرنل‌های گسترده فضایی تمرکز می‌کند. هر یک از این عملیات در ابتدا بدون فرض اشتراک وزن معرفی می‌شوند، یعنی به عنوان مثال، اجازه یک کرنل متفاوت در هر نقطه از منیفولد را می‌دهند. این کرنل‌ها (یا بایاس‌ها یا غیرخطی‌ها) به هیچ وجه محدود نمی‌شوند. با این حال، هنگام نیاز به اشتراک وزن فضایی، آن‌ها مجبور می‌شوند که تناوب‌پذیر گیج باشند زیرا تنها مقادیر تناوب‌پذیر می‌توانند به صورت مستقل از مختصات به اشتراک گذاشته شوند. بخش ۳.۴ یک اثبات مختصر از تناوب‌پذیری ایزومتری کانولوشن‌های GM را بر حسب عبارات مختصاتی محلی ارائه می‌دهد. ایده اصلی در اینجا این است که ایزومتری‌ها را می‌توان به عنوان القاکننده تبدیل‌های گیج (تفسیر منفعل) در نظر گرفت، که با تناوب‌پذیری گیج کرنل‌ها توضیح داده می‌شود.

بخش ؟؟ پیاده‌سازی کانولوشن‌های مستقل از جهت‌گیری بر روی نوار موبیوس را توصیف می‌کند. پس از بررسی هندسه نوار موبیوس در بخش ؟؟، انواع مختلفی از میدان‌های ویژگی در بخش ؟؟ تعریف می‌شوند. بخش ؟؟ بعدی، شبکه‌های کانولوشنی مستقل از جهت‌گیری را به صورت تحلیلی توصیف می‌کند. به طور خاص، ما کرنل‌های کانولوشنی تناوب‌پذیر گیج، بایاس‌ها و غیرخطی‌ها را برای هر یک از انواع میدان استخراج می‌کنیم. بخش ؟؟ با یک پیاده‌سازی و ارزیابی عددی مدل‌های مربوطه به پایان می‌رسد.

بخش ؟؟: بخش ؟؟ محتوای بخش ۳ را بازتاب می‌دهد، با این حال، به صورت سراسری و بر حسب بندل‌های فیبر. مقدمه‌ای کلی بر بندل‌های فیبر در بخش ؟؟ ارائه شده است. بخش‌های ؟؟ و ؟؟ بندل مماس TM، بندل چارچوب FM، G–ساختارها GM و بندل‌های بردار ویژگی مرتبط با G ($\mathcal{A}$) را معرفی می‌کنند. میدان‌های ویژگی به صورت سراسری به عنوان برش‌هایی از بندل‌های بردار ویژگی تعریف می‌شوند. تسهیم‌های بندل محلی (گیج‌ها)، که در بخش ؟؟ مورد بحث قرار گرفته‌اند، این بندل‌ها را در مختصات بیان می‌کنند و بدین ترتیب تعاریف ما از بخش ۳ را بازیابی می‌کنند. ما به طور خاص نشان می‌دهیم که چگونه تسهیم‌های محلی بندل‌های مختلف یکدیگر را القا می‌کنند، به طوری که تبدیل‌های گیج (نگاشت‌های انتقال) آن‌ها همگام‌سازی می‌شوند. بخش ؟؟ انتقال‌دهنده‌های موازی در G–بندل‌ها را مورد بحث قرار می‌دهد.

بخش ؟؟ شبکه‌های مستقل از مختصات از بخش ۴ را بر حسب بندل‌های فیبر بازنویسی می‌کند. کانولوشن‌های 1×1 در بخش ؟؟ به عنوان مورفیسم‌های مشخص بندل برداری M توصیف می‌شوند. به طور جایگزین، آن‌ها را می‌توان به عنوان برش‌هایی از یک بندل همومورفیسم مشاهده کرد. بخش ؟؟ میدان‌های کرنل مستقل از مختصات و تبدیل‌های میدان کرنل را معرفی می‌کند. این عملیات مشابه کانولوشن‌های GM هستند اما نیازی به اشتراک وزن ندارند، یعنی ممکن است در هر مکان فضایی یک کرنل متفاوت اعمال کنند. یک میدان کرنل کانولوشنی GM با اشتراک یک کرنل G–استیریبل

(تناوب‌پذیر گیج) منفرد در سراسر منیفولد ساخته می‌شود. سپس کانولوشن‌های مستقل از مختصات GM به عنوان تبدیل‌های میدان کرنل با میدان‌های کرنل GM–کانولوشنی تعریف می‌شوند. هنگام بیان فرمول‌بندی مستقل از مختصات کانولوشن‌های GM نسبت به تسهیم‌های محلی (گیج‌ها)، عبارات مختصاتی کانولوشن‌های GM را از بخش ۲.۴ بازیابی می‌کنیم.

تناوب‌پذیری ایزومتری کانولوشن‌های GM در بخش ؟؟ بررسی می‌شود. پس از معرفی ایزومتری‌ها، بخش ؟؟ عمل پیش‌برنده آن‌ها را بر روی بندل‌های فیبر مورد بحث قرار می‌دهد. این عمل نیز می‌تواند در تسهیم‌های محلی بیان شود، که منجر به فرمول‌بندی از بخش ۴.۳ می‌شود. بخش ؟؟ عمل ایزومتری‌ها بر روی میدان‌های کرنل را تعریف می‌کند و ثابت می‌کند که تناوب‌پذیری ایزومتری یک تبدیل میدان کرنل، ناوردایی ایزومتری میدان کرنل آن را نتیجه می‌دهد و بالعکس. کانولوشن‌های GM ثابت شده‌اند که تحت عمل آن ایزومتری‌هایی که اتومورفیسم‌های بندلی (تقارن‌ها) از G–ساختار GM هستند، تناوب‌پذیر می‌باشند. بخش ؟؟ میدان‌های کرنل ناوردا با ایزومتری را با جزئیات بیشتری بررسی می‌کند و ثابت می‌کند که آن‌ها معادل میدان‌های کرنل در فضاهای خارج‌قسمتی از عمل ایزومتری هستند □ به طور شهودی، میدان‌های کرنل ناوردا با ایزومتری ملزم به اشتراک کرنل‌ها بر روی مدارات ایزومتری هستند. این نتیجه به ویژه دلالت دارد که تبدیل‌های میدان کرنل تناوب‌پذیر ایزومتری در فضاهای همگن لزوماً کانولوشن‌های GM هستند.

بخش ؟؟: بخش سوم این کار نشان می‌دهد که تعداد زیادی از شبکه‌های کانولوشنی از ادبیات را می‌توان به عنوان اعمال کانولوشن‌های GM برای انتخاب خاصی از G–ساختار و انواع میدان تفسیر کرد. این بخش با بحثی کلی در مورد انتخاب‌های طراحی شبکه‌های کانولوشنی مستقل از مختصات آغاز می‌شود. جدول ؟؟ مروری و طبقه‌بندی از مدل‌های مورد بررسی را ارائه می‌دهد. خواننده دعوت می‌شود که به G–ساختارهای بصری‌سازی شده در بخش ؟؟ نگاهی بیندازد زیرا اینها ایده شهودی در مورد ویژگی‌های کانولوشن‌های GM مربوطه ارائه می‌دهند.

شبکه‌های کانولوشنی اقلیدسی که نه تنها تناوب‌پذیر ایزومتری هستند بلکه به طور کلی‌تر تحت عمل گروه‌های آفین تناوب‌پذیر هستند، در بخش ؟؟ بررسی می‌شوند. این مدل‌ها اساساً معادل شبکه‌های کانولوشنی استیریبل در فضاهای برداری اقلیدسی $\mathbb{R}^d$[؟ ؟ ؟] هستند. بخش ؟؟ شبکه‌های کانولوشنی استیریبل را بررسی می‌کند و ارتباط آن‌ها با کانولوشن‌های GM را مورد بحث قرار می‌دهد. این رویکرد تا حدی نامطلوب است زیرا $\mathbb{R}^d$ با یک میدان چارچوب کانونی (ساختار $\{e\}$) همراه است، که به طور ضمنی توسط مدل‌های تناوب‌پذیر نادیده گرفته می‌شود. بخش ؟؟ رویکردی اصولی‌تر را در پیش می‌گیرد و فضاهای آفین اقلیدسی $\mathbb{E}_d$ را تعریف می‌کند که دقیقاً با G–ساختارهایی مجهز شده‌اند که منجر به کانولوشن‌های GM تناوب‌پذیر $\operatorname{Aff}(G)$ می‌شوند. کانولوشن‌های واقعی GM در بخش ؟؟ تعریف می‌شوند. بخش ؟؟ شبکه‌های کانولوشنی اقلیدسی تناوب‌پذیر آفین را که در ادبیات یافت می‌شوند، بررسی می‌کند. آن‌ها عمدتاً در انتخاب‌های مفروض گروه‌های ساختار و نمایش‌های گروهی متفاوت هستند.

بخش ؟؟ شبکه‌های کانولوشنی را در فضاهای اقلیدسی سوراخ‌دار $\mathbb{E}_d \backslash \{۰\}$، که مبدأ $\{۰\}$ آن‌ها حذف شده است، پوشش می‌دهد. این مدل‌ها نسبت به چرخش حول مبدأ تناوب‌پذیر هستند، با این حال، نسبت به انتقال تناوب‌پذیر نیستند. آن‌ها بر اساس G–ساختارهایی هستند که با مختصات قطبی، مختصات لگاریتمی–قطبی یا مختصات کروی مطابقت دارند.

شبکه‌های کانولوشنی کروی در بخش ؟؟ پوشش داده می‌شوند. بخش ؟؟ هندسه ۲–کره (جاسازی شده) $S^۲$ را مورد بحث قرار می‌دهد. با تفسیر فضاهای مماس به عنوان زیرفضاهای دو بعدی یک فضای جاسازی $\mathbb{R}^۳$، عبارات فرم بسته نگاشت‌های نمایی و لگاریتمی، چارچوب‌ها، گیج‌ها، انتقال‌دهنده‌ها و عمل ایزومتری را استخراج می‌کنیم. بخش ؟؟ شبکه‌های کانولوشنی کروی تناوب‌پذیر $\mathrm{SO}(۳)$ و ۳ را بررسی می‌کند. ما به طور خاص ثابت می‌کنیم که نظریه ما فرمول‌بندی عمومی کانولوشن‌های کروی توسط ؟] را به عنوان یک حالت خاص شامل می‌شود. شبکه‌های کانولوشنی کروی که فقط نسبت به چرخش $\mathrm{SO}(۲)$ حول یک محور ثابت تناوب‌پذیر هستند در بخش ؟؟ توصیف می‌شوند. بخش ؟؟ شبکه‌های کانولوشنی بیست‌وجهی را بررسی می‌کند. بیست‌وجهی کره را تقریب می‌زند اما از وجوه محلی مسطح تشکیل شده است که امکان پیاده‌سازی کارآمد عملیات کانولوشن را فراهم می‌کند.

یک بررسی از شبکه‌های کانولوشنی بر روی سطوح دو بعدی عمومی در بخش ؟؟ یافت می‌شود. بخش ؟؟ مقدمه‌ای کوتاه بر هندسه دیفرانسیل کلاسیک سطوح جاسازی شده و گسسته‌سازی آن‌ها بر حسب شبکه‌های مثلثی ارائه می‌دهد. کانولوشن‌های سطحی در ادبیات به دو دسته طبقه‌بندی می‌شوند: دسته اول، که در بخش ؟؟ پوشش داده شده است، بر اساس کرنل‌های $G = \mathrm{SO}(۲)$–استیریبل است. این مدل‌ها مستقل از انتخاب خاص چارچوب متعامد راست‌گرد هستند. بخش ؟؟ دسته دوم مدل‌ها را بررسی می‌کند که بر اساس کرنل‌های $\{e\}$–استیریبل، یعنی غیرتناوب‌پذیر هستند. این مدل‌ها به صراحت بر انتخاب یک میدان چارچوب تکیه دارند. بنابراین آن‌ها عمدتاً در روش‌های اکتشافی که برای تعیین چارچوب‌های مرجع استفاده می‌شوند، تفاوت دارند. توجه داشته باشید که چنین مدل‌هایی لزوماً در منیفلدهای غیرقابل موازی‌سازی مانند کره‌های توپولوژیکی ناپیوسته هستند.

پیوست: پیوست شامل اطلاعات اضافی و اثبات‌های طولانی است.

گیج‌ها به صورت رسمی تخصیص فوری چارچوب‌های مرجع به فضاهای مماس را انجام می‌دهند اما به نقاط روی منیفولد به صورت مستقل از مختصات اشاره دارند. یک جایگزین محبوب انتخاب چارت‌های مختصاتی است که به اصطلاح پایه‌های مختصاتی (پایه‌های هولونومیک) فضاهای مماس را القا می‌کنند. پیوست ؟؟ مقدمه‌ای بر فرمالیسم چارت‌ها ارائه می‌دهد و آن را با فرمالیسم گیج کلی‌تر مرتبط می‌کند.

پیوست ؟؟ درباره استقلال مختصات کرنل‌ها و اشتراک وزن در امتداد چارچوب‌های مرجع توضیح می‌دهد. اشتراک وزن مستقل از مختصات GM فقط برای کرنل‌های G–استیریبل امکان‌پذیر است.

کانولوشن‌های GM با بیان میدان‌های ویژگی در مختصات نرمال ژئودزیک محاسبه می‌شوند، جایی که آن‌ها با کرنل‌های کانولوشنی G–استیریبل مطابقت داده می‌شوند. این فرآیند شامل یک انتگرال‌گیری بر روی فضاهای مماس است که در پیوست ؟؟ توصیف شده است.

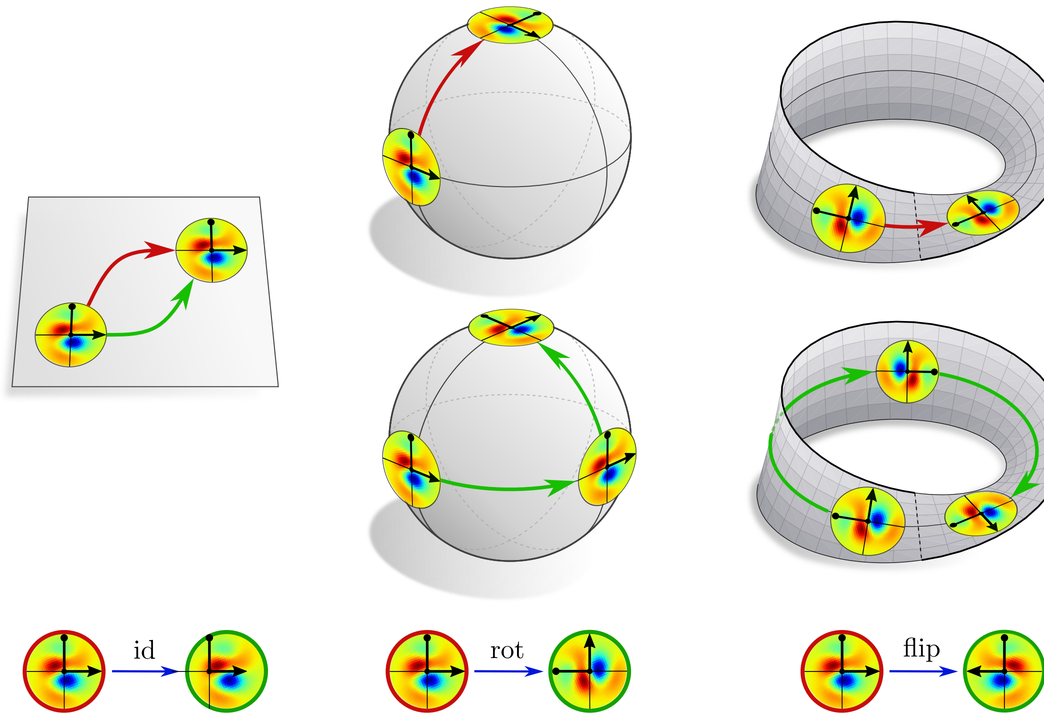
؟]، ؟]و ؟] نظریه‌های نسبتاً عمومی از کانولوشن‌ها بر روی فضاهای همگن را پیشنهاد کردند. از آنجا که این مدل‌ها وزن‌ها را از طریق عمل یک گروه تقارنی به اشتراک می‌گذارند، آن‌ها بسیار شبیه به تبدیل‌های میدان کرنل تناوب‌پذیر ایزومتری ما از بخش‌های ؟؟ و ؟؟ هستند. پیوست ؟؟ این مدل‌ها را بررسی می‌کند و توضیح می‌دهد که چگونه آن‌ها با کانولوشن‌های GM ما مرتبط هستند.

پیوست ؟؟ ثابت می‌کند که میدان‌های کرنل ناوردا با ایزومتری در منیفولد معادل میدان‌های کرنل در فضاهای خارج‌قسمتی از عمل ایزومتری هستند. حالت خاص فضاهای همگن، که در آن‌ها تبدیل‌های میدان کرنل تناوب‌پذیر ایزومتری معادل کانولوشن‌های GM هستند، در پیوست ؟؟ پوشش داده شده است.

کانولوشن‌های کروی ؟] در پیوست ؟؟ ثابت شده‌اند که یک حالت خاص از کانولوشن‌های GM کروی ما هستند بنابراین هر شبکه کانولوشنی کروی که توسط نظریه آن‌ها پوشش داده می‌شود، توسط نظریه ما نیز توضیح داده می‌شود.

پیوست ؟؟ تأکید می‌کند که تبدیل‌های میدان کرنل و کانولوشن‌های GM ما به خوبی تعریف شده‌اند اگر میدان کرنل هموار باشد و شامل کرنل‌های با پشتیبانی فشرده باشد. "به خوبی تعریف شده" در اینجا به این معنی است که انتگرال‌ها وجود دارند و میدان‌های ویژگی حاصل هموار هستند.

در نهایت، پیوست ؟؟ استدلال می‌کند که میدان‌های ویژگی که مطابق با نمایش منظم گروه ساختار G تبدیل می‌شوند، معادل میدان‌های اسکالر بر روی G–ساختار هستند. این موضوع از آن جهت اهمیت دارد که برخی مدل‌ها، به ویژه کانولوشن‌های گروهی، این دیدگاه را اتخاذ می‌کنند.



شکل ۴: شهودی در مورد ابهام ذاتی اشتراک وزن در منیفلدهایی. چپ: یک تفسیر رایج از اشتراک وزن در صفحه، جابجایی یک کرنل در کل فضا است. از آنجا که انتقال موازی در فضاهای مسطح مستقل از مسیر است، این امر بدون ابهام است. وسط: در فضاهای خمیده، مانند کره، انتقال موازی وابسته به مسیر است. مسیرهای مختلف منجر به کرنل‌هایی می‌شوند که نسبت به یکدیگر چرخانده شده‌اند. راست: نوار موبیوس یک منیفولد غیرقابل جهت‌گیری است. بنابراین، مسیرهای مختلف می‌توانند منجر به کرنل‌هایی شوند که نسبت به یکدیگر بازتاب یافته‌اند. پایین: ما ترازهای مختلف کرنل را با انتخاب‌های مختلف چارچوب‌های مرجع محلی از فضاهای مماس مربوطه رسمی‌سازی می‌کنیم. به خوبی شناخته شده است که هیچ انتخابی از چارچوب‌های مرجع (گیج) در منیفلدهای عمومی ارجح نیست. مختصات‌بندی‌های مختلف با تبدیل‌های گیج مرتبط هستند، که مقادیری در گروه ساختار G منیفولد (گروه بدیهی $G = \{e\}$ برای صفحه، گروه چرخش $G = \mathrm{SO}(۲)$ برای کره و گروه بازتاب $G = \mathcal{R}$ برای نوار موبیوس) می‌گیرند. شبکه‌های کانولوشنی مستقل از مختصات، ابهام چارچوب‌های مرجع را با اعمال کرنل‌های کانولوشنی G-استیریبل (تناوب‌پذیر گیج) برطرف می‌کنند.

۲ مرور کلی و شهود بصری

فرمول‌بندی جبری شبکه‌های کانولوشنی مستقل از مختصات نیازمند آشنایی با نظریه گروه‌ها، نظریه نمایش‌ها و هندسه دیفرانسیل است که ممکن است برای مخاطبان غیرتخصصی مانعی ایجاد کند. با این حال، بیشتر ساختارها و نتایج ما از نظر هندسی بسیار شهودی هستند و می‌توان آن‌ها را با چند مثال بصری توضیح داد. این بخش تلاش می‌کند تا مروری کلی و شهود بصری در مورد شبکه‌های کانولوشنی مستقل از مختصات ارائه دهد.

بخش ۱.۲ زیر، کانولوشن‌های مستقل از مختصات GM را در منیفلدهای ریمانی معرفی می‌کند. تناوب‌پذیری آن‌ها تحت عمل ایزومتری‌ها در بخش ؟؟ مورد بحث قرار می‌گیرد. بخش ؟؟ در مورد عواملی که بر انتخاب G-ساختار در طراحی شبکه‌های کانولوشنی مستقل از مختصات تأثیر می‌گذارند، توضیح می‌دهد.

۱.۲ ￼-ساختارها و کانولوشن‌های مستقل از مختصات ￼￼

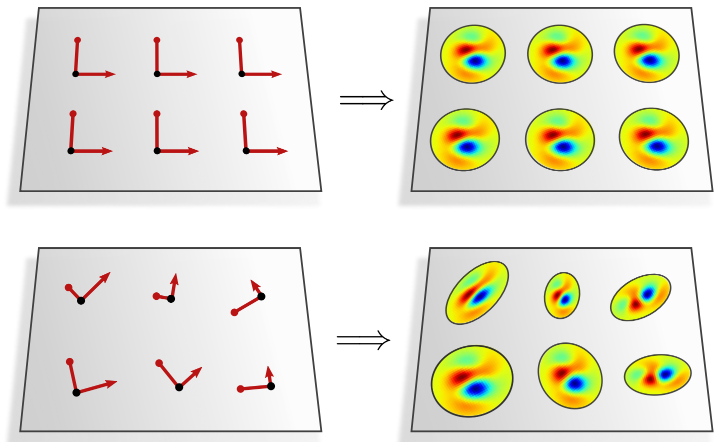
از آنجا که کانولوشن‌ها اساساً با خاصیت اشتراک وزن خود مشخص می‌شوند، یک سؤال اصلی در این کار این است:

چگونه باید کرنل‌های کانولوشن بر روی منیفلدهای ریمانی به اشتراک گذاشته شوند؟ [۳]

اشتراک وزن از طریق تقارن‌ها

یک رویکرد رایج به اشتراک گذاشتن وزن‌ها از طریق عمل یک گروه تقارنی از فضای زیرین است [؟ ؟]. به عنوان مثال، شبکه‌های کانولوشنی سنتی وزن‌ها را با جابجایی کرنل‌ها روی صفحه به اشتراک می‌گذارند، در حالی که شبکه‌های کانولوشنی کروی وزن‌ها را با چرخش کرنل‌ها روی کره

[۳] این سؤال به طور کلی‌تر برای هر تابع قالب محلی، به عنوان مثال بایاس‌ها یا غیرخطی‌های نقطه‌ای نیز کاربرد دارد.



شکل ۵: یک ویژگی کلیدی کانولوشن‌ها این است که آن‌ها وزن‌ها را در سراسر منیفولد به اشتراک می‌گذارند. ما تراز یک کرنل در $p \in M$ را با انتخاب یک چارچوب مرجع □ یا گیج □ از فضای مماس مربوطه T_pM شناسایی می‌کنیم. بنابراین میدان‌های چارچوب مختلف دلالت بر میدان‌های کرنل (کانولوشنی) متفاوتی دارند.

انتخاب چارچوب‌ها اغلب منحصر به فرد نیست. ابهام در این انتخاب با G-ساختارها رسمی‌سازی می‌شود؛ به شکل ؟؟ مراجعه کنید. برای در نظر گرفتن اختیاری بودن چارچوب‌ها، کرنل‌ها ملزم می‌شوند که G-استیریبل (تناوب‌پذیر) باشند، همانطور که در اشکال ؟؟ و ؟؟ بصری‌سازی شده است.

به اشتراک می‌گذارند. برای به اشتراک گذاشتن یک کرنل در کل فضا، عمل گروه تقارنی باید متعدی باشد. از آنجا که این امر به طور کلی برای ایزومتری‌های منیفلدهای ریمانی صادق نیست، این استراتژی برای هدف ما رد می‌شود.

اشتراک وزن از طریق انتقال

اشتراک وزن در فضاهای اقلیدسی اغلب به عنوان "جابجایی" یک کرنل در فضا در نظر گرفته می‌شود. از آنجا که در فضاهای مسطح انتقال موازی مستقل از مسیر انتخاب شده است، این امر منجر به تراز بدون ابهام کرنل‌ها می‌شود؛ به شکل ۴ (چپ) مراجعه کنید. با این حال، در فضاهای خمیده یا غیرقابل جهت‌گیری، انتقال موازی وابسته به مسیر می‌شود و بنابراین برای اشتراک وزن نامناسب است. شکل ۴ (وسط و راست) این مشکل را برای کره و نوار موبیوس مثال می‌زند، جایی که مسیرهای مختلف منجر به تراز متفاوتی از کرنل می‌شوند.

اشتراک وزن در امتداد چارچوب‌ها

از آنجا که مفهوم "ترازهای کرنل" تا حدی مبهم است ابتدا باید آن را از نظر ریاضی دقیق کنیم:

ما انتخاب تراز کرنل در نقطه‌ای $p \in M$
را به عنوان انتخاب یک چارچوب مرجع محلی (گیج) از فضای مماس مربوطه T_pM رسمی‌سازی می‌کنیم.

یک چارچوب مرجع در $p \in M$ یک تاپل مرتب $[e_1, \ldots, e_d]$ از $d := \dim(M)$ بردار مماس خطی مستقل $e_i \in T_pM$ است که به عنوان محورهای چارچوب نامیده می‌شوند. از آنجا که چارچوب‌های مختلف در p با تبدیل‌های خطی مرتبط هستند، انتخاب‌های مختلف چارچوب‌ها با تغییر شکل‌های کرنل خطی مطابقت دارند. شکل ۵ دو انتخاب مختلف از میدان‌های چارچوب را در $M = \mathbb{R}^2$ نشان می‌دهد. به اشتراک گذاشتن یک کرنل کانولوشنی در امتداد این میدان‌های چارچوب منجر به میدان‌های کرنل (کانولوشنی) مربوطه می‌شود.

شناسایی ترازهای کرنل با چارچوب‌های مرجع این سؤال را مطرح می‌کند:

انتخاب چارچوب‌های مرجع محلی بر روی یک منیفولد (ریمانی) تا چه حد مبهم است؟

همانطور که در ادامه توضیح داده می‌شود، ابهام چارچوب‌های مرجع توسط یک G-ساختار که منیفولد به آن مجهز است، تعیین می‌شود.

۱.۱.۲ □-ساختارها

بندل چارچوب FM

فضای همه چارچوب‌های ممکن T_pM با عنوان F_pM نشان داده می‌شود. با هم، چارچوب‌های همه فضاهای مماس، بندل چارچوب FM را تشکیل می‌دهند؛ به شکل ؟؟ مراجعه کنید. هیچ انتخاب خاصی از چارچوب‌ها در FM بر روی یک منیفولد هموار "عریان" (بدون ساختار هندسی اضافی) ترجیح داده نمی‌شود، که ابهام تراز کرنل‌ها را حداکثری می‌کند. برای رفع ابهام چارچوب‌ها و ترازهای کرنل، منیفولد باید با ساختار هندسی اضافی مجهز شود.

ساختار متریک OM

یک منیفولد ریمانی با ساختار متریک مجهز شده است. با ارائه یک حاصل‌ضرب داخلی (متریک ریمانی) بر روی فضاهای مماس، این ساختار امکان شناسایی چارچوب‌های خاصی را می‌دهد که محورهای آن‌ها متعامد با یکدیگر هستند. تبدیل‌های گیج، یعنی تبدیل‌ها بین انتخاب‌های چارچوب‌های مرجع (به اشکال ؟؟ (چپ) و ۶ مراجعه کنید)، سپس مقادیری در گروه متعامد d می‌گیرند. در منیفلدهای ریمانی، تراز کرنل‌ها همیشه تا چرخش‌ها و بازتاب‌ها بدون ابهام است.

G-ساختارها GM

شبکه‌های کانولوشنی اقلیدسی کرنل‌های کانولوشن (و در نتیجه چارچوب‌ها) را تراز می‌کنند موازی با یکدیگر همانطور که در شکل ۵ (بالا) بصری‌سازی شده است. شبکه‌های کانولوشنی کروی معمولاً ترازهای کرنل را تا چرخش‌ها بدون ابهام می‌کنند، یعنی آن‌ها یک دست‌پری ترجیحی از چارچوب‌های مرجع را فرض می‌کنند. ساختار متریک به تنهایی برای توصیف این تنظیمات کافی نیست، که نشان می‌دهد این منیفلدهایی با ساختار هندسی اضافی علاوه بر ساختار متریک مجهز شده‌اند. ما پیشنهاد می‌کنیم که چارچوب ریاضی مناسب، G-ساختارها است و فرض می‌کنیم:

ابهام در انتخاب چارچوب‌های مرجع (و در نتیجه ترازهای کرنل)
بر روی یک منیفولد M با G-ساختار GM آن رسمی‌سازی می‌شود.

G-ساختارها GM بندل‌هایی از چارچوب‌های مرجع متمایز بر روی M هستند به طوری که تبدیل‌های گیج بین چارچوب‌های یک فضای مماس واحد، مقادیری در گروه ساختار $G \leq \mathrm{GL}(d)$ می‌گیرند. به طور شهودی، می‌توان مجموعه G_pM چارچوب‌های T_pM را "شبیه" G دانست، با این حال، بدون یک مبدأ متمایز.[۴]

بندل چارچوب FM خود یک G-ساختار با $G = \mathrm{GL}(d)$ است، در حالی که بندل چارچوب‌های متعامد $\mathrm{O}M$ یک G-ساختار (ساختار متریک) با $G = d$ است. شبکه‌های کانولوشنی اقلیدسی سنتی بر روی میدان چارچوب کانونی در $\mathbb{R}^d$ که در شکل ؟؟ نشان داده شده است، تکیه دارند که یک G-ساختار برای گروه بدیهی $G = \{e\}$ است. شکل ؟؟ G-ساختارها را برای منیفلدهای بیشتر و گروه‌های ساختار دیگر بصری‌سازی می‌کند. مروری بر گروه‌های ساختار رایج در جدول ۱ در بخش ۳.۱.۳ یافت می‌شود.

ما در ادامه همیشه منیفلدهای ریمانی را مجهز به یک G-ساختار اضافی در کنار ساختار متریک خود فرض خواهیم کرد.[۵] انتخاب خاص G-ساختار خواص شبکه عصبی را تعیین می‌کند؛ ما در مورد این انتخاب در بخش ؟؟ زیر توضیح خواهیم داد.

---

[۴] G_pM یک فضای همگن اصلی از G (یا G(□□□□□□- است.

[۵] G-ساختار باید با ساختار متریک سازگار باشد به این صورت که چارچوب‌های متمایز در GM زیرمجموعه‌ای از چارچوب‌های متعامد در $\mathrm{O}M$ باشند زمانی که $G < d$. به طور خاص برای $G = d$، G-ساختار GM با ساختار متریک $\mathrm{O}M$ منطبق است و اطلاعات هندسی اضافی را اضافه نمی‌کند.

شبکه‌های کانولوشنی سلسله‌مراتبی از میدان‌های ویژگی را از یک سیگنال ورودی بر روی یک منیفولد استخراج می‌کنند. ویژگی‌ها از طریق کرنل‌ها محاسبه می‌شوند، که برای شناسایی الگوهای فضایی مشخص در ویژگی‌های سطح پایین‌تر بهینه شده‌اند. ما تقاضا می‌کنیم که این فرآیند استنتاج صرفاً بر اساس آرایش نسبی ویژگی‌ها باشد و مستقل از انتخاب خاص مختصات‌بندی باشد. بنابراین ویژگی‌ها باید کمیت‌های هندسی مستقل از مختصات باشند، مشابه اسکالرها، بردارها یا تانسورها. در حالی که چنین کمیت‌های هندسی مستقل از مختصات وجود دارند، یک پیاده‌سازی کامپیوتری (غیرنمادین) نیازمند آن است که آن‌ها بر حسب ضرایب عددی در برخی گیج، یعنی نسبت به برخی انتخاب از چارچوب مرجع، بیان شوند. انتخاب خاص مختصات بی‌اهمیت است □ این تنها یکی از چندین توصیف معادل است. چارچوب ریاضی مناسب برای تنظیم چنین درجات آزادی اضافی، نظریه‌های گیج هستند. یک نظریه گیج، برابری گیج‌های مختلف را با مرتبط کردن سازگار آن‌ها با یکدیگر از طریق تبدیل‌های گیج در نظر می‌گیرد. میدان‌های ویژگی مستقل از مختصات بنابراین با یک قانون تبدیل خاص، یعنی یک عمل گروهی از گروه ساختار که توصیف می‌کند ویژگی‌ها چگونه تحت تبدیل‌های گیج تبدیل می‌شوند، مرتبط هستند. هر لایه شبکه عصبی که چنین میدان‌های ویژگی را پردازش می‌کند، ملزم به رعایت قوانین تبدیل آن‌ها برای حفظ استقلال از مختصاتشان است.

هدف این بخش اول از کار ما، معرفی شبکه‌های کانولوشنی مستقل از مختصات با زبانی آسان است. بنابراین، شهود هندسی و بصری‌سازی‌ها بر فرمالیسم ریاضی ترجیح داده می‌شوند. یک تبیین رسمی‌تر از تعاریف و نتایج ارائه شده در بخش ؟؟ آورده شده است.

بخش ۳ گیج‌ها و تبدیل‌های گیج را معرفی می‌کند که بر اساس آن‌ها میدان‌های بردارهای ویژگی مستقل از مختصات تعریف می‌شوند. شبکه‌های عصبی که بین چنین میدان‌های ویژگی نگاشت می‌کنند، در بخش ۴ توسعه یافته‌اند. بخش ؟؟ یک نمونه پیاده‌سازی از میدان‌های ویژگی و لایه‌های شبکه را در نوار موبیوس ارائه می‌دهد.

۳ مقدمه‌ای بر شبکه‌های کانولوشنی مستقل از مختصات

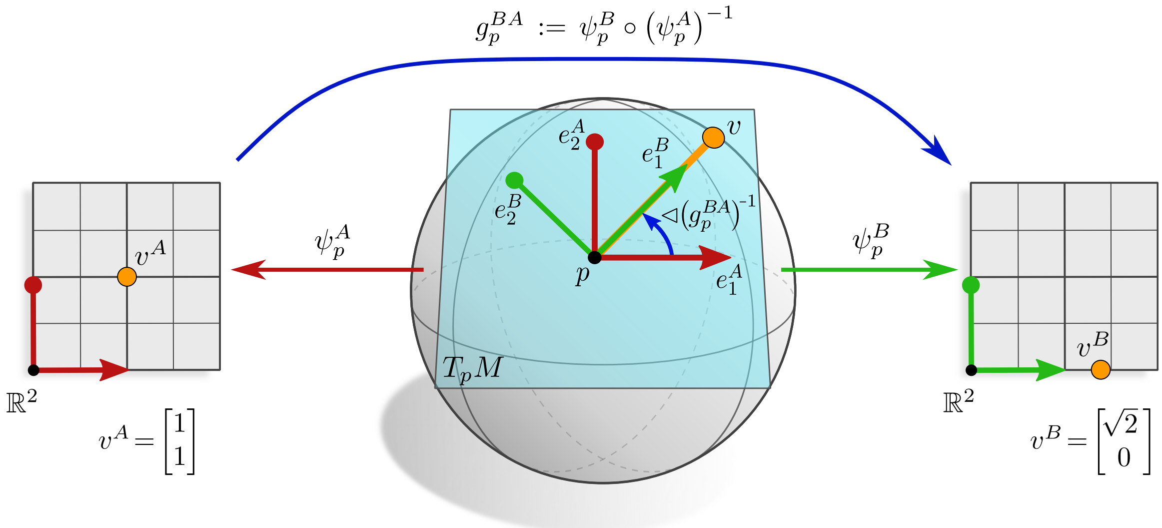
شبکه‌های کانولوشنی سلسله‌مراتبی از میدان‌های ویژگی را از یک سیگنال ورودی بر روی یک منیفولد استخراج می‌کنند. ویژگی‌ها از طریق کرنل‌ها محاسبه می‌شوند، که برای شناسایی الگوهای فضایی مشخص در ویژگی‌های سطح پایین‌تر بهینه شده‌اند. ما تقاضا می‌کنیم که این فرآیند استنتاج صرفاً بر اساس آرایش نسبی ویژگی‌ها باشد و مستقل از انتخاب خاص مختصات‌بندی باشد. بنابراین ویژگی‌ها باید کمیت‌های هندسی مستقل از مختصات باشند، مشابه اسکالرها، بردارها یا تانسورها. در حالی که چنین کمیت‌های هندسی مستقل از مختصات وجود دارند، یک پیاده‌سازی کامپیوتری (غیرنمادین) نیازمند آن است که آن‌ها بر حسب ضرایب عددی در برخی گیج، یعنی نسبت به برخی انتخاب از چارچوب مرجع، بیان شوند. انتخاب خاص مختصات بی‌اهمیت است □ این تنها یکی از چندین توصیف معادل است. چارچوب ریاضی مناسب برای تنظیم چنین درجات آزادی اضافی، نظریه‌های گیج هستند. یک نظریه گیج، برابری گیج‌های مختلف را با مرتبط کردن سازگار آن‌ها با یکدیگر از طریق تبدیل‌های گیج در نظر می‌گیرد. بنابراین میدان‌های ویژگی مستقل از مختصات با یک قانون تبدیل خاص، یعنی یک عمل گروهی از گروه ساختار که توصیف می‌کند ویژگی‌ها چگونه تحت تبدیل‌های گیج تبدیل می‌شوند، مرتبط هستند. هر لایه شبکه عصبی که چنین میدان‌های ویژگی را پردازش می‌کند، ملزم به رعایت قوانین تبدیل آن‌ها برای حفظ استقلال از مختصاتشان است. هدف این بخش اول از کار ما، معرفی شبکه‌های کانولوشنی مستقل از مختصات با زبانی آسان است. بنابراین، شهود هندسی و بصری‌سازی‌ها بر فرمالیسم ریاضی ترجیح داده می‌شوند. یک تبیین رسمی‌تر از تعاریف و نتایج ارائه شده در بخش ؟؟ آورده شده است.

بخش ۳ گیج‌ها و تبدیل‌های گیج را معرفی می‌کند که بر اساس آن‌ها میدان‌های بردارهای ویژگی مستقل از مختصات تعریف می‌شوند. شبکه‌های عصبی که بین چنین میدان‌های ویژگی نگاشت می‌کنند، در بخش ۴ توسعه یافته‌اند. بخش ؟؟ یک نمونه پیاده‌سازی از میدان‌های ویژگی و لایه‌های شبکه را در نوار موبیوس ارائه می‌دهد.

۱.۳ گیج‌ها، تبدیل‌های گیج و G-ساختارها

۱.۱.۳ فضاهای مماس و چارچوب‌های مرجع

یک منیفولد هموار d-بعدی M دارای یک فضای مماس $T_pM \cong \mathbb{R}^d$ متصل به هر نقطه $p \in M$ است. فضاهای مماس، فضاهای برداری d-بعدی هستند، با این حال، برخلاف $\mathbb{R}^d$، آن‌ها به طور کلی با هیچ انتخاب ارجح چارچوب مرجع همراه نیستند. یک بردار مماس $v \in T_pM$ یک شی مستقل از مختصات است و بنابراین بلافاصله به صورت عددی با یک تاپل مختصاتی $(v_1, \ldots, v_d) \in \mathbb{R}^d$ نمایش داده نمی‌شود. به طور انتزاعی‌تر، هر



شکل ۶: شناسایی $T_pM \cong \mathbb{R}^۲$ با $\mathbb{R}^۲$ از طریق گیج‌های مختلف. یک بردار مماس (مستقل از مختصات) $v \in T_pM$ (نارنجی) را می‌توان به صورت عددی با یک تاپل مختصاتی $v^A = \psi_p^A(v) = (۱, ۱)^\top$ نسبت به گیج ψ_p^A (قرمز) یا، به طور معادل، با $v^B = \psi_p^B(v) = (\sqrt{۲}, ۰)^\top$ نسبت به گیج ψ_p^B (سبز) نمایش داد. انتخاب یک گیج مربوط به انتخاب $[e_۱^A, e_۲^A]$ یا $[e_۱^B, e_۲^B]$ از چارچوب مرجع است. در یک منیفولد عمومی هیچ انتخابی از گیج یا مختصات‌بندی پیش‌فرض ارجح نیست. گیج‌های مختلف، و در نتیجه چارچوب‌های مرجع، با تبدیل‌های گیج $g_p^{BA} := \psi_p^B \circ (\psi_p^A)^{-۱}$ (آبی) مرتبط هستند که مقادیری در گروه ساختار G تعریف شده می‌گیرند. این شکل تفسیری گرافیکی از نمودارهای جابجایی در معادله (۸) و شکل ؟؟ است. توجه داشته باشید که گیج‌ها بلافاصله به فضاهای مماس مختصات می‌دهند. شکل ؟؟ در بخش ؟؟ یک نمودار مشابه برای چارت‌های (آفین) را نشان می‌دهد که به منیفولد مختصات می‌دهند و بدین ترتیب گیج‌ها ("پایه‌های مختصات") را القا می‌کنند.

فضای مماس T_pM با $\mathbb{R}^d$ هم‌ریخت است اما به طور کلی هیچ هم‌ریختی کانونی بین آن‌ها وجود ندارد. بنابراین هر دو فضا از نظر ساختاری معادل هستند اما به هیچ روش ارجح به یکدیگر شناسایی نمی‌شوند.

یک گیج (تسهیم محلی بندل مماس) بر روی $U^A \subseteq M$ به عنوان مجموعه‌ای از نگاشت‌های خطی معکوس‌پذیر که به صورت هموار وابسته به موقعیت هستند، تعریف می‌شود:

$$\psi_p^A : T_pM \to \mathbb{R}^d\,, \;\; p \in U^A\,, \tag{۱}$$

که هم‌ریختی‌های فضای برداری گمشده بین T_pM و $\mathbb{R}^d$ را مشخص می‌کند. همانطور که در شکل ۶ بصری‌سازی شده است، آن‌ها با تخصیص یک بردار ضریب به فضاهای مماس مختصات می‌دهند:

$$v^A \;:=\; \psi_p^A(v) \;\in \mathbb{R}^d \tag{۲}$$

به هر بردار مماس مستقل از مختصات $v \in T_pM$. معکوس این رابطه می‌دهد:

$$v \;=\; \left(\psi_p^A\right)^{-1}\!\left(v^A\right) \;=\; \left(\psi_p^A\right)^{-1}\!\Big(\sum\nolimits_i v_i^A \epsilon_i\Big) \;=\; \sum\nolimits_i v_i^A \,\left(\psi_p^A\right)^{-1}\!\left(\epsilon_i\right) \;=:\; \sum\nolimits_i v_i^A\, e_i^A\,, \tag{۳}$$

که در آن $\{\epsilon_۱, \ldots, \epsilon_d\}$ را پایانه استاندارد $\mathbb{R}^d$ نامگذاری کرده‌ایم و از خطی بودن گیج برای بیرون کشیدن جمع استفاده کرده‌ایم. این نشان می‌دهد که گیج را می‌توان به عنوان مجهز کردن هر فضای مماس T_pM با یک چارچوب مرجع در نظر گرفت:

$$\left[e_۱^A,\,\ldots,\,e_d^A\right] \;:=\; \left[(\psi_p^A)^{-۱}(\epsilon_۱),\;\ldots\;,\;(\psi_p^A)^{-۱}(\epsilon_d)\right], \tag{۴}$$

که به عنوان یک d-تاپل از بردارهای مماس خطی مستقل تعریف می‌شود که با نگاشت پایه استاندارد $\mathbb{R}^d$ به عقب از طریق نگاشت معکوس گیج به دست می‌آید. برای اختصار، ما در ادامه از نماد کوتاه‌شده $\left[e_i^A\right]_{i=۱}^d$ برای چارچوب‌ها $\left[e_۱^A, \ldots, e_d^A\right]$ استفاده خواهیم کرد. ضرایب v^A مختصات v نسبت به این چارچوب هستند. مجموعه‌ای از چارچوب‌های القا شده توسط ψ_p^A بر روی U^A (هموار) میدان چارچوب نامیده می‌شود؛ برای بصری‌سازی به شکل ۷ مراجعه کنید.

گیج‌های ψ^X فضاهای مماس را تنها در همسایگی‌های محلی $U^X \subseteq M$ مختصات می‌دهند، و به دلیل موانع توپولوژیکی به طور کلی نمی‌توانند به کل منیفولد بدون نقض فرض همواری گسترش یابند. بنابراین یک اطلس در نظر گرفته می‌شود:

$$\mathscr{A} \;=\; \left\{ \left(U^X, \psi^X \right) \right\}_{X \in \mathfrak{X}} \,, \tag{۵}$$

متشکل از گیج‌های هموار بر روی مجموعه‌ای از همسایگی‌های U^X که منیفولد را پوشش می‌دهند، یعنی شرط $\bigcup_{X \in \mathfrak{X}} U^X \;=\; M$ را برآورده می‌کنند، که در آن $\mathfrak{X}$ یک مجموعه اندیس است.[۶] در نواحی همپوشانی $U^A \cap U^B \neq \varnothing$ از همسایگی‌ها، گیج‌های مختلف ψ_p^A و ψ_p^B توسط توابع گذار هموار به هم متصل می‌شوند:

$$g^{BA}:\; U^A \cap U^B \to \mathrm{GL}(d), \quad p \mapsto g_p^{BA} := \psi_p^B \circ \left(\psi_p^A \right)^{-1} . \tag{۶}$$

در اینجا ما دامنه مشترک (فعلاً) را با گروه خطی عمومی $\mathrm{GL}(d)$ در نظر می‌گیریم، متشکل از همه ماتریس‌های معکوس‌پذیر در $\mathbb{R}^{d\times d}$، که رابطه بین هر جفت از هم‌ریختی‌های فضای برداری (گیج‌ها) یا چارچوب‌های مرجع را توضیح می‌دهند. عمل چنین تابع گذاری بر روی یک گیج داده شده یک تبدیل گیج را تعریف می‌کند:

$$\psi_p^B \;=\; g_p^{BA} \cdot \psi_p^A . \tag{۷}$$

از نظر یک نمودار جابجایی، رابطه بین گیج‌های مختلف به صورت زیر بصری‌سازی می‌شود:[۷]

$$\begin{tikzcd}
\mathbb{R}^d \arrow[rr, bend left, "g_p^{BA}\cdot"] & T_pM \arrow[l, "\psi_p^A"'] \arrow[r, "\psi_p^B"] & \mathbb{R}^d \arrow[ll, bend left, "g_p^{AB}\cdot \;=\; \left(g_p^{BA}\right)^{-1}\cdot"]
\end{tikzcd} \tag{۸}$$

این نمودار را با تفسیر گرافیکی آن در شکل ۶ مقایسه کنید.

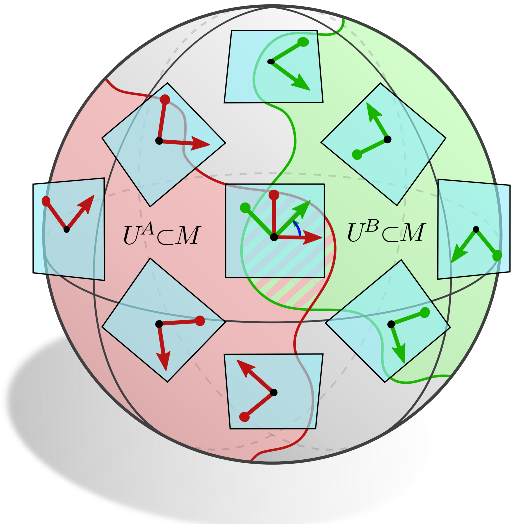
یک تبدیل گیج، مختصات‌بندی فضاهای مماس را تغییر می‌دهد به طوری که همان بردار مماس مستقل از مختصات v با یک بردار مولفه‌ای متفاوت نمایش داده می‌شود:

$$v^B \;=\; g_p^{BA} v^A \,. \tag{۹}$$

---

[۶] اطلس گیج‌ها بسیار شبیه به اطلس‌های معمول چارت‌های یک منیفولد است (پیوست ؟؟). تفاوت این است که اطلس‌های مورد نظر در اینجا مستقیماً به بندل مماس TM مختصات می‌دهند به جای منیفولد M.

[۷] نمودارها یک نمای کلی بصری از توابع و فضاهایی را که بین آن‌ها نگاشت می‌کنند، ارائه می‌دهند. به عنوان مثال، نمودار
$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "h"'] & Y \arrow[d, "g"] \\ & Z \end{tikzcd}$
به این معنی است که توابع $f:\, X \to Y$، $g:\, Y \to Z$ و $h:\, X \to Z$ وجود دارند. اگر ترکیبات توابع در امتداد همه مسیرهایی با شروع و پایان یکسان مطابقت داشته باشند، نمودار جابجایی نامیده می‌شود. نمودار مثال ما اگر (و تنها اگر) $h = g \circ f$ صادق باشد، جابجایی است.



شکل ۷: هر نقطه p از یک منیفولد ریمانی M دارای یک فضای مماس T_pM متصل است. یک گیج هموار ψ^A بر روی یک زیرمجموعه مناسب انتخاب شده $U^A \subseteq M$ (قرمز) تمام فضاهای مماس T_pM را برای p در U^A مختصات می‌دهد همانطور که در شکل ۶ نشان داده شده است. این معادل انتخاب یک میدان چارچوب هموار بر روی U^A است. از آنجا که به طور کلی امکان گسترش یک گیج به صورت سراسری بر روی کل منیفولد وجود ندارد، لازم است که یک G-اطلس در نظر گرفته شود، متشکل از گیج‌هایی که M را پوشش می‌دهند. مختصات‌بندی‌های مختلف ψ^A بر روی U^A (قرمز) و ψ^B بر روی U^B (سبز) از طریق تبدیل‌های گیج (یا نگاشت‌های گذار) $g^{BA} : U^A \cap U^B \to G$ به هم متصل می‌شوند که بر روی همپوشانی $U^A \cap U^B$ (خطدار) تعریف شده‌اند و مقادیری در گروه ساختار $G \leq \mathrm{GL}(d)$ می‌گیرند.

از آنجا که یک گیج مربوط به انتخاب یک میدان چارچوب است، یک تبدیل گیج مربوط به تبدیل بین میدان‌های چارچوب است. به طور خاص، یک چارچوب $\left[e_i^A\right]_{i=1}^d = \left[e_1^A, \dots, e_d^A\right]$ در $p \in M$ به یک چارچوب دیگر تبدیل می‌شود:

$$
\begin{aligned}
\left[e_i^B\right]_{i=1}^d &:= \left[\left(\psi_p^B\right)^{-1}(\epsilon_i)\right]_{i=1}^d && \text{(چارچوب القا شده توسط گیج، □□ (۴))}\\
&= \left[\left(g_p^{BA}\cdot\psi_p^A\right)^{-1}(\epsilon_i)\right]_{i=1}^d && \text{(تبدیل گیج، □□ (۷))}\\
&= \left[\left(\psi_p^A\right)^{-1}\left(\left(g_p^{BA}\right)^{-1}\epsilon_i\right)\right]_{i=1}^d && \text{(معکوس گسترش یافته)}\\
&= \left[\left(\psi_p^A\right)^{-1}\left(\sum\nolimits_j \epsilon_j\epsilon_j^\top \left(g_p^{BA}\right)^{-1}\epsilon_i\right)\right]_{i=1}^d && \text{(هویت وارد شده } \mathbb{1} = \textstyle\sum_j \epsilon_j\epsilon_j^\top\text{)}\\
&= \left[\left(\psi_p^A\right)^{-1}\left(\sum\nolimits_j \epsilon_j \left(\left(g_p^{BA}\right)^{-1}\right)_{ji}\right)\right]_{i=1}^d && \text{(عناصر ماتریس } \left(g_p^{BA}\right)^{-1} \text{ شناسایی شده)}\\
&= \left[\sum\nolimits_j \left(\psi_p^A\right)^{-1}(\epsilon_j)\left(\left(g_p^{BA}\right)^{-1}\right)_{ji}\right]_{i=1}^d && \text{(خطی بودن } \psi_p^A\text{)}\\
&= \left[\sum\nolimits_j e_j^A\left(\left(g_p^{BA}\right)^{-1}\right)_{ji}\right]_{i=1}^d && \text{(چارچوب القا شده توسط گیج، □□ (۴))}\\
&=: \left[e_i^A\right]_{i=1}^d \blacktriangleright \left(g_p^{BA}\right)^{-1}
\end{aligned}
\tag{۱۰}
$$

از طریق عمل راست تعریف شده:

$$
\blacktriangleright : \left(\left[e_i\right]_{i=1}^d,\, g\right) \mapsto \left[e_i\right]_{i=1}^d \blacktriangleright g := \left[\sum\nolimits_j e_j\, g_{ji}\right]_{i=1}^d
\tag{۱۱}
$$

از عناصر گروه بر روی چارچوب‌ها. توجه داشته باشید که معکوس در این عمل در معادله (۱۰) به دلیل تعریف معادله (۷) بدون معکوس است.[۸] معمولاً به رفتار تبدیل چارچوب‌های مرجع، تبدیل کوواریانت گفته می‌شود در حالی که تبدیل گیج‌ها و ضرایب برداری به عنوان تبدیل کنتروواریانت نامیده می‌شود؛ به پیوست ؟؟ مراجعه کنید.

[۸] قراردادهای دیگر ممکن است انتخاب معکوس‌ها را در $\psi^B = g^{BA}\psi^A$ و $[e_i^B]_{i=1}^d = [e_i^A]_{i=1}^d \blacktriangleright \left(g^{BA}\right)^{-1}$ تغییر دهند. یک معکوس در هر یک از دو معادله برای سازگاری عمل چپ $\cdot$ بر روی گیج‌ها و عمل راست $\blacktriangleright$ بر روی چارچوب‌ها ضروری است.

از آنجا که رفتار تبدیل ضرایب در معادله (۹) و پایه در معادله (۱۰) معکوس یکدیگر هستند، آن‌ها یکدیگر را خنثی می‌کنند، یعنی بردار مماس $v = \sum_i v_i^A e_i^A = \sum_i v_i^B e_i^B$ را ناوردا می‌گذارند:

$$\begin{aligned} v \;=\; \sum\nolimits_i v_i^B e_i^B \;&=\; \sum\nolimits_i v_i^B \sum\nolimits_j e_j^A \left(\left(g_p^{BA}\right)^{-1} \right)_{ji} \\ &=\; \sum\nolimits_j \left(\sum\nolimits_i \left(\left(g_p^{BA}\right)^{-1} \right)_{ji} v_i^B \right) e_j^A \\ &=\; \sum\nolimits_j v_j^A e_j^A \,. \end{aligned} \tag{۱۲}$$

این ساختار تضمین می‌کند که هر محاسبه‌ای در نهایت مستقل از گیج انتخاب شده است، که معمولاً به آن استقلال از مختصات گفته می‌شود. به طور کلی، هر نمایش مختصاتی از یک شی یا تابع مستقل از مختصات برای دلایل سازگاری باید مستقل از مختصات باشد.

برای کامل بودن می‌خواهیم اشاره کنیم که فرمالیسم ارائه شده در اینجا پایه‌های عمومی فضاهای مماس را تعریف می‌کند، که گاهی اوقات به عنوان پایه‌های غیرمختصاتی (پایه‌های غیرهولونومیک) از نظر گیج‌های محلی نامیده می‌شوند. یک جایگزین بسیار محبوب اما کمتر عمومی پایه‌های مختصاتی (پایه‌های هولونومیک) هستند:

$$\left[\frac{\partial}{\partial x_1^A}\bigg|_p , \, \dots \, , \, \frac{\partial}{\partial x_d^A}\bigg|_p \right] , \tag{۱۳}$$

که توسط چارت‌های مختصاتی منیفولد القا می‌شوند [؟]:

$$x^A : U^A \to V^A \subseteq \mathbb{R}^d \tag{۱۴}$$

گیج‌های مربوطه توسط دیفرانسیل‌های چارت داده می‌شوند، یعنی:

$$\psi_p^A \;=\; \hat{d}x_p^A \;=\; \left(\hat{d}x_{p,1}^A , \dots , \, \hat{d}x_{p,d}^A \right)^\top \;:\; T_pM \to \mathbb{R}^d \,. \tag{۱۵}$$

تبدیل‌های گیج در این تنظیم با ژاکوبیان‌ها مطابقت دارند:

$$g_p^{BA} \;=\; \frac{\partial x^B}{\partial x^A}\bigg|_{x^A(p)} \;\in\; \mathrm{GL}(d) \tag{۱۶}$$

از نگاشت‌های گذار چارت. یک چارت نمونه و پایه‌های مختصاتی القا شده آن در شکل ۸ بصری‌سازی شده‌اند. پیوست ؟؟ رابطه بین هر دو فرمالیسم را با جزئیات مورد بحث قرار می‌دهد؛ یک مرور کلی در جدول ؟؟ ارائه شده است.

در ادامه این مقاله ما عمدتاً در فرمالیسم گیج کار خواهیم کرد، که چارچوب‌های مرجع را مستقیماً به فضاهای مماس تخصیص می‌دهد به جای اینکه آن‌ها را از چارت‌ها القا کند. استثناها عبارتند از کانولوشن‌های موبیوس در بخش ؟؟، CNNهای اقلیدسی در بخش ؟؟، مختصات لگاریتمی–قطبی در بخش ؟؟ و CNNهای بیست‌وجهی در بخش ؟؟. در همه این موارد، منیفلدهایی به صورت محلی مسطح هستند و چارت‌ها ایزومتریک هستند، به طوری که چارچوب‌های متعامد را القا می‌کنند. کانولوشن‌های GM در U^A سپس می‌توانند به روشی کارآمد با اجرای کانولوشن‌های اقلیدسی با کرنل‌های G–استیریبل در دامنه‌های مشترک چارت‌ها V^A محاسبه شوند.

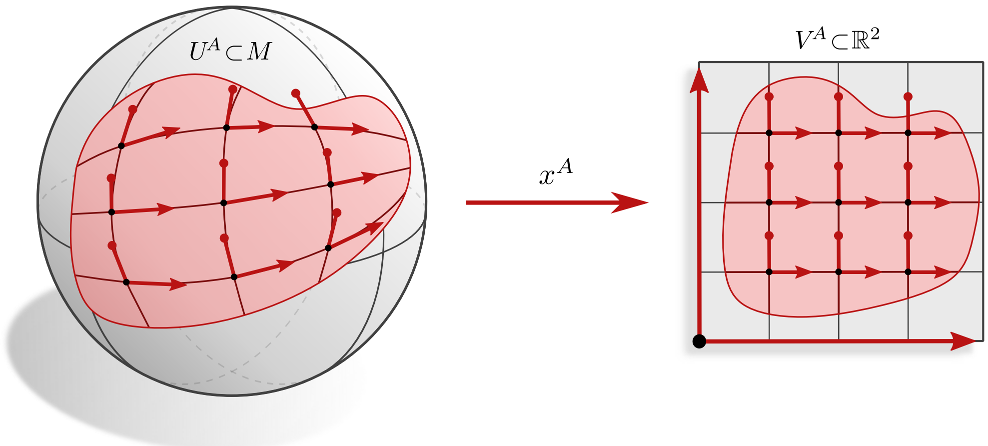
۲.۱.۳ توابع مستقل از مختصات در فضاهای مماس

همانطور که بردارهای $v \in T_pM$، توابع بر روی فضاهای مماس مستقل از مختصات هستند، یعنی بدون اشاره به هیچ چارچوب مرجع تعریف می‌شوند. یک گیج انتخاب شده امکان نمایش چنین نگاشت‌های مستقل از مختصات را توسط توابعی فراهم می‌کند که بر روی بردارهای ضریب در $\mathbb{R}^d$ عمل می‌کنند. مشابه بردارهای ضریب، مختصات‌بندی توابع باید به روشی خاص تحت تبدیل‌های گیج تبدیل شوند تا به طور سازگار تعریف شوند، یعنی برای رعایت استقلال از مختصات. ما بعداً مفهوم ارائه شده در اینجا را از بیان نگاشت‌های مستقل از مختصات بر حسب مختصات محلی برای تعریف کانولوشن‌های مستقل از مختصات GM به کار خواهیم برد.

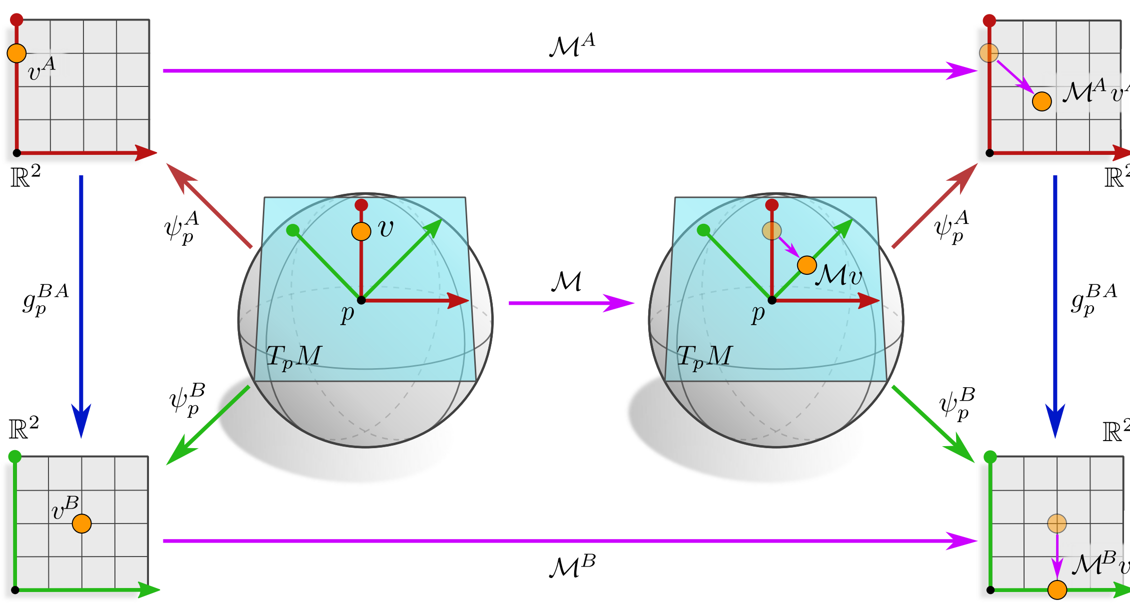
به عنوان یک مثال ساده برای یک عملیات مستقل از مختصات، بیایید حالت یک نگاشت خطی را در نظر بگیریم:

$$\mathcal{M} : T_pM \to T_pM \,. \tag{۱۷}$$

فرض کنید $v_{\square\square} \in T_pM$ یک بردار مماس باشد که توسط $\mathcal{M}$ به $v_{\square\square\square} = \mathcal{M} v_{\square\square} \in T_pM$ نگاشت می‌شود. نگاشت‌های خطی در پیاده‌سازی‌های عددی معمولاً توسط ماتریس‌های ضریب مدل می‌شوند که بین بردارهای ضریب نسبت به یک انتخاب از چارچوب مرجع نگاشت می‌کنند. برای دقیق



شکل ۸: یک چارت $x^A : U^A \to V^A$ مختصات $V^A \subseteq \mathbb{R}^d$ را به نواحی $U^A \subseteq M$ از منیفولد تخصیص می‌دهد. این چارت پایه‌های مختصاتی $\left[\frac{\partial}{\partial x_1^A}\big|_p, \ldots, \frac{\partial}{\partial x_d^A}\big|_p\right]$ و گیج‌های مربوطه $\psi_p^A = \hat{d}x_p^A$ را برای فضاهای مماس T_pM در U^A القا می‌کند. ما عمدتاً با چارت‌ها کار نخواهیم کرد بلکه به نقاط $p \in M$ به صورت مستقل از مختصات ارجاع خواهیم داد. گیج‌ها (چارچوب‌ها) سپس مستقیماً به فضاهای مماس تخصیص داده می‌شوند به جای اینکه از چارت‌ها القا شوند.



شکل ۹: تفسیر گرافیکی نمودار جابجایی در معادله (۲۳). یک نگاشت مستقل از مختصات $\mathcal{M} : T_pM \to T_pM$ را می‌توان به طور معادل با توابع $\mathcal{M}^A : \mathbb{R}^d \to \mathbb{R}^d$ یا $\mathcal{M}^B : \mathbb{R}^d \to \mathbb{R}^d$ نسبت به گیج‌های مختلف ψ_p^A یا ψ_p^B نمایش داد. این مختصات‌بندی‌های $\mathcal{M}$ با پیش و پس ترکیب با گیج‌ها در دامنه و دامنه مشترک تعریف می‌شوند، به عنوان مثال، با دنبال کردن فلش‌ها، $\mathcal{M}^A := \psi_p^A \circ \mathcal{M} \circ \left(\psi_p^A\right)^{-1}$. در نتیجه، تبدیل‌های گیج $\mathcal{M}^B = g_p^{BA} \mathcal{M}^A \left(g_p^{BA}\right)^{-1}$ بین مختصات‌بندی‌ها توسط یک پیش و پس ترکیب با نگاشت‌های گذار g_p^{BA} در دامنه و دامنه مشترک داده می‌شوند. همه کمیت‌ها و نگاشت‌ها در این کار یا مستقل از مختصات خواهند بود (مانند $\mathcal{M}$) یا به روشی مستقل از مختصات در گیج‌های مختلف بیان خواهند شد (مانند $\mathcal{M}^A$ و $\mathcal{M}^B$). ما بنابراین باید قوانین تبدیل را برای هر کمیت و تابع تعریف (یا استخراج) کنیم.

کردن این موضوع، فرض کنید گیج ψ_p^A داده شده باشد به طوری که بردارهای مستقل از مختصات $v_{\square\square}$ و $v_{\square\square\square}$ در T_pM با بردارهای ضریب $v_{\square\square}^A = \psi_p^A(v_{\square\square})$ و $v_{\square\square\square}^A = \psi_p^A(v_{\square\square\square})$ در $\mathbb{R}^d$ نمایش داده شوند. نگاشت خطی $\mathcal{M}$ در این گیج توسط ماتریس نمایش داده می‌شود:

$$\mathcal{M}^A \;:=\; \psi_p^A \circ \mathcal{M} \circ \left(\psi_p^A\right)^{-۱} \;\in\; \mathbb{R}^{d\times d} \tag{۱۸}$$

که تعریف آن توسط نمودار جابجایی زیر بصری‌سازی شده است:

$$\begin{tikzcd} \mathbb{R}^d \arrow[rrr, bend right, "\mathcal{M}^A"'] & T_pM \arrow[l, "\psi_p^A"'] \arrow[r, "\mathcal{M}"] & T_pM \arrow[r, "\psi_p^A"] & \mathbb{R}^d \end{tikzcd} \tag{۱۹}$$

ماتریس با نگاشت مستقل از مختصات سازگار است زیرا هر دو یکدیگر را نتیجه می‌دهند:

$$\begin{aligned} \mathcal{M}^A v_{\square\square}^A &= \left[\psi_p^A \circ \mathcal{M} \circ \left(\psi_p^A\right)^{-۱}\right] \circ \left[\psi_p^A(v_{\square\square})\right] \\ &= \psi_p^A\big(\mathcal{M}\, v_{\square\square}\big) \\ &= \psi_p^A(v_{\square\square\square}) \\ &= v_{\square\square\square}^A \end{aligned} \tag{۲۰}$$

البته می‌توان $\mathcal{M}$ را نسبت به هر انتخاب دیگری از گیج ψ_p^B نیز نمایش داد. ما از معادله (۹) می‌دانیم که بردارهای ضریب در گیج‌های مختلف با $v^B = g_p^{BA} v^A$ مرتبط هستند. به طور مشابه، $\mathcal{M}^B$ با $\mathcal{M}^A$ توسط تبدیل گیج مرتبط است:

$$\begin{aligned} \mathcal{M}^B &= \psi_p^B \circ \mathcal{M} \circ \left(\psi_p^B\right)^{-۱} \\ &= \psi_p^B \circ \left(\psi_p^A\right)^{-۱} \circ \mathcal{M}^A \circ \psi_p^A \circ \left(\psi_p^B\right)^{-۱} \\ &= g_p^{BA} \mathcal{M}^A \left(g_p^{BA}\right)^{-۱}, \end{aligned} \tag{۲۱}$$

که در اینجا هم بر دامنه و هم بر دامنه مشترک عمل می‌کند.[۹] این قانون تبدیل دوباره سازگار است زیرا تبدیل‌های متقابل یکدیگر را خنثی می‌کنند:

$$\begin{aligned} \mathcal{M}^B v_{\square\square}^B &= \left[g_p^{BA} \mathcal{M}^A \left(g_p^{BA}\right)^{-۱}\right] \left[g_p^{BA} v_{\square\square}^A\right] \\ &= g_p^{BA} \mathcal{M}^A v_{\square\square}^A \\ &= g_p^{BA} v_{\square\square\square}^A \\ &= v_{\square\square\square}^B \end{aligned} \tag{۲۲}$$

تبدیل‌های گیج استخراج شده بنابراین تأیید می‌کنند که تمام محاسبات مختصات‌بندی شده در نهایت مستقل از مختصات هستند. روابط بین نگاشت مستقل از مختصات و مختصات‌بندی‌های آن توسط نمودار جابجایی زیر خلاصه می‌شود:

$$\begin{tikzcd} \mathbb{R}^d \arrow[rrr, "\mathcal{M}^A"] \arrow[ddd, "g_p^{BA}\cdot"'] & & & \mathbb{R}^d \arrow[ddd, "g_p^{BA}\cdot"] \\ & & & \\ & T_pM \arrow[uul, "\psi_p^A"] \arrow[ddl, "\psi_p^B"'] \arrow[r, "\mathcal{M}"] & T_pM \arrow[uur, "\psi_p^A"'] \arrow[ddr, "\psi_p^B"] & \\ & & & \\ \mathbb{R}^d \arrow[rrr, "\mathcal{M}^B"'] & & & \mathbb{R}^d \end{tikzcd} \tag{۲۳}$$

که در شکل ۹ به صورت گرافیکی تفسیر شده است.

در عمل نمی‌توان نگاشت خطی مستقل از مختصات $\mathcal{M}$ را به صورت عددی بدون اشاره به یک انتخاب مختصات‌بندی پیاده‌سازی کرد. با این حال، وجود آن تنها در صورتی (و تنها در صورتی) دلالت دارد که مختصات‌بندی‌های آن همانطور که در معادله (۲۱) مشخص شده است به یکدیگر مربوط باشند، که تضمین می‌کند رفتار تبدیل صحیح ضرایب بردار ورودی و خروجی در معادله (۹) حفظ می‌شود.

[۹] تبدیل ضرایب ماتریس از طریق ضرب چپ و راست با g^{BA} و $\left(g^{BA}\right)^{-۱}$ به ترتیب، نگاشت خطی را به عنوان یک تانسور از نوع $(۱,۱)$ شناسایی می‌کند.

۳.۱.۳ گروه‌های ساختار، G-ساختارها و G-اطلس‌ها

ما بعداً از شبکه‌های عصبی می‌خواهیم که به روشی مستقل از مختصات عمل کنند، یعنی ما تقاضا می‌کنیم که استنتاج آن‌ها مستقل از انتخاب‌های دلخواه چارچوب‌های مرجع باشد. این سؤال را مطرح می‌کند که انتخاب چارچوب‌های مرجع در یک منیفولد تا چه حد دلخواه است. در بخش‌های ۱.۱.۳ و ۲.۱.۳ قبلی ما هر انتخاب ممکن از گیج یا چارچوب مرجع را مجاز می‌دانستیم، که بنابراین توسط تبدیل‌های گیج با مقدار کلی $\mathrm{GL}(d)$ مرتبط بودند. با این حال، در بسیاری از کاربردها، منیفولد دارای ساختار اضافی است که امکان تشخیص یک زیرمجموعه ارجح از چارچوب‌های مرجع یا گیج‌ها را می‌دهد که توابع گذار آن‌ها مقادیری در یک گروه ساختار کاهش یافته $G \leq \mathrm{GL}(d)$ می‌گیرند. چنین ساختارهای هندسی □ یا بلکه خود زیرمجموعه‌های چارچوب‌های مرجع ترجیحی، که اطلاعات معادل را کدگذاری می‌کنند □ به عنوان G-ساختارها نامیده می‌شوند.

G-ساختارها با در نظر گرفتن چند مثال خاص بهتر درک می‌شوند. لیست زیر چنین مثال‌هایی را ارائه می‌دهد که بر اساس گروه ساختار $G \leq \mathrm{GL}(d)$ طبقه‌بندی شده‌اند:

- d: ساختار متریک یک منیفولد ریمانی را در نظر بگیرید، که امکان اندازه‌گیری فواصل و زوایا را می‌دهد، و بنابراین تشخیص چارچوب‌های متعامد را، یعنی آن چارچوب‌هایی که شرط $\eta(e_i, e_j) = \delta_{ij}$ را برای هر $i, j = ۱, \dots, d$ برآورده می‌کنند. به طور متناظر، یک متریک ریمانی امکان صحبت در مورد گیج‌های ایزومتریک ψ_p^A را می‌دهد، که متریک $\mathbb{R}^d$ را با متریک T_pM شناسایی می‌کنند، یعنی شرط $\eta(v, w) = \langle \psi_p^A(v),\, \psi_p^A(w) \rangle_{\mathbb{R}^d}$ را برای هر $v, w \in T_pM$ برآورده می‌کنند. از آنجا که چارچوب‌های متعامد و گیج‌های ایزومتریک تا چرخش‌ها و بازتاب‌ها تعریف می‌شوند، هر تبدیل گیجی بین آن‌ها مقادیری در گروه متعامد d خواهد گرفت، که آن زیرگروه از $\mathrm{GL}(d)$ است که زوایا و فواصل را حفظ می‌کند.
- $\mathrm{GL}^+(d)$: به طور مشابه، یک جهت‌گیری منیفولد چارچوب‌های راست‌گرد را از چپ‌گرد و گیج‌های حفظ کننده جهت‌گیری را از گیج‌های معکوس کننده جهت‌گیری متمایز می‌کند. تبدیل‌های گیج بین چارچوب‌های یک دست‌پری داده شده مقادیری در $\mathrm{GL}^+(d)$ می‌گیرند، یعنی آن زیرگروه از $\mathrm{GL}(d)$ که جهت‌گیری‌ها را حفظ می‌کند.
- $\mathrm{SO}(d)$: با هم، یک متریک و جهت‌گیری داده شده، چارچوب‌های متعامد با جهت‌گیری خاصی را مشخص می‌کنند. تبدیل‌های گیج بین چنین چارچوب‌هایی تضمین شده‌اند که در زیرگروه $\mathrm{SO}(d)$ از $\mathrm{GL}(d)$ قرار گیرند.
- $\{e\}$: یک میدان چارچوب هموار سراسری یک $\{e\}$-ساختار را در M تعریف می‌کند. در این حالت تنها یک چارچوب متمایز در هر موقعیت وجود دارد، به طوری که تبدیل‌های گیج در گروه بدیهی $\{e\} \leq \mathrm{GL}(d)$ قرار می‌گیرند.
- $\mathrm{GL}(d)$: اگر هیچ ساختار اضافی تحمیل نشود، هر چارچوب مرجع فضاهای مماس به همان اندازه معتبر است. تبدیل‌های گیج در این حالت نگاشت‌های خطی معکوس‌پذیر عمومی در $\mathrm{GL}(d)$ هستند و G-ساختار مربوطه فقط بندل چارچوب FM است.

موضوع مشترک در این مثال‌های انگیزشی این است که همه آن‌ها با موارد زیر تعریف می‌شوند:

۱. یک زیرمجموعه (که به صورت فضایی هموار تغییر می‌کند) از چارچوب‌های مرجع متمایز،
۲. یک زیرمجموعه مربوطه از گیج‌های ترجیحی و
۳. یک زیرگروه $G \leq \mathrm{GL}(d)$ از تبدیل‌های گیج که مفهوم متمایز چارچوب‌ها و گیج‌ها را حفظ می‌کنند.

چنین زیرمجموعه‌هایی از چارچوب‌های مرجع که به صورت هموار تغییر می‌کنند به عنوان G-ساختارها GM بر روی M نامیده می‌شوند و گروه G به عنوان گروه ساختار (کاهش یافته) نامیده می‌شود □ برای تعریفی دقیق‌تر به بخش ؟؟ مراجعه کنید.[۱۰] فرآیند مشخص کردن یک G-ساختار به عنوان کاهش گروه ساختار از $\mathrm{GL}(d)$ به G شناخته می‌شود. یک اطلس $\mathscr{A}^G = \left\{\left(U^X, \psi^X\right)\right\}_{X \in \mathfrak{X}}$ به عنوان G-اطلس نامیده می‌شود اگر تمام توابع گذار آن:

$$g^{BA}:\, U^A \cap U^B \to G, \quad p \mapsto g_p^{BA} := \psi_p^B \circ \left(\psi_p^A\right)^{-۱} \tag{۲۴}$$

در یک گروه ساختار کاهش یافته $G \leq \mathrm{GL}(d)$ قرار گیرند (مقایسه کنید با معادله (۶)). رابطه بین چارچوب‌های مرجع و گیج‌ها در معادله (۴) دلالت دارد بر اینکه هر G-اطلس یک G-ساختار مربوطه را کدگذاری می‌کند.

انتخاب‌های متعددی از G-ساختارها ممکن است برای یک گروه ساختار G داده شده وجود داشته باشند. برای ارتباط با مثال‌های بالا: متریک‌های ریمانی مختلف، زیرمجموعه‌های متفاوتی از چارچوب‌های مرجع را به عنوان متعامد مشخص می‌کنند، یعنی آن‌ها با ساختارهای d مختلف $\mathrm{O}M$ مطابقت دارند. بنابراین انتخاب یک متریک معادل انتخاب یک ساختار d است. به طور مشابه، انتخاب‌های مختلف جهت‌گیری یک منیفولد قابل جهت‌گیری، مجموعه متفاوتی از چارچوب‌ها را به عنوان راست‌گرد مشخص می‌کنند. بنابراین دو انتخاب ممکن جهت‌گیری با دو انتخاب ممکن از ساختارهای $\mathrm{GL}^+(d)$ (GL^+M) مطابقت دارند. ساختارهای $\mathrm{SO}(d)$ ($\mathrm{SO}M$) ممکن است در انتخاب جهت‌گیری و متریک یک هر دو متفاوت باشند. مثال دیگر ساختار $\{e\}$ $\{e\}M$ است. آن‌ها اجازه تبدیل‌های گیج (غیربدیهی) را نمی‌دهند و بنابراین با انتخاب میدان‌های چارچوب هموار سراسری در M مطابقت دارند. جدول ۱ مثال‌های بیشتری از گروه‌های ساختار G و G-ساختارهای مربوطه را ارائه می‌دهد.

[۱۰] به طور رسمی، GM به عنوان یک زیربندل اصلی G از بندل چارچوب FM تعریف می‌شود، که یک بندل اصلی $\mathrm{GL}(d)$ است.

structure group $G \leq \mathrm{GL}(d)$	G-structure GM	equivalent structure on M
$\mathrm{GL}^+(d)$	positively oriented frames	orientation of M
$\mathrm{SL}(d)$	unit volume frames	volume form
$\mathrm{CO}(d)$	conformal frames	—
$\mathrm{Sp}(d)$	symplectic frames	—
$\mathrm{O}(d)$	orthonormal frames	Riemannian metric
$\mathrm{O}(d-n,\, n)$	pseudo-orthonormal frames	pseudo-Riemannian metric
$\mathrm{SO}(d)$	positively oriented orthonormal frames	Riemannian metric + orientation
$\{e\}$	parallelization (global frame field)	—

جدول ۱: مثال‌هایی از G-ساختارها GM بر روی M و گروه‌های ساختار کاهش‌یافته مربوطه $G \leq \mathrm{GL}(d)$. یک G-ساختار به عنوان یک زیرمجموعه هموار متغیر از چارچوب‌های مرجع (یک زیربندل اصلی G از بندل چارچوب FM) تعریف می‌شود، که در آن چارچوب‌های هر فضای مماس به صورت متقابل توسط تبدیل‌های گیج با مقدار G مرتبط هستند. در حالی که این تعریف نسبتاً انتزاعی است، امکان مشاهده بسیاری از ساختارهای هندسی در M را به روشی یکپارچه فراهم می‌کند. به عنوان مثال، یک متریک ریمانی در M امکان تشخیص چارچوب‌های متعامد را می‌دهد. برعکس، مشخص کردن متعامد بودن به طور منحصر به فرد یک متریک را نتیجه می‌دهد. بنابراین یک متریک ریمانی و یک ساختار متعامد با یکدیگر معادل هستند. به طور مشابه، یک تناظر یک به یک بین فرم‌های حجم و چارچوب‌های واحد حجم وجود دارد. توجه داشته باشید که انتخاب یک گروه ساختار G به طور منحصر به فرد یک G-ساختار را مشخص نمی‌کند. به عنوان مثال، متریک‌های ریمانی مختلف می‌توانند به عنوان ساختار d، فرم‌های حجم مختلف به عنوان ساختار $\mathrm{SL}(d)$ یا میدان‌های چارچوب سراسری مختلف به عنوان ساختار $\{e\}$ انتخاب شوند. CNNهای مستقل از مختصات برای رعایت یک G-ساختار داده شده طراحی شده‌اند □ اینکه کدام ساختار خاص است به وظیفه یادگیری بستگی دارد.

کاهش گروه ساختار به G، یعنی وجود یک G-ساختار، ممکن است توسط توپولوژی منیفولد مسدود شود. این دلالت دارد بر اینکه یک "گروه ساختار کاهش‌ناپذیر" وجود دارد که فراتر از آن ابهام چارچوب‌های مرجع را نمی‌توان بدون نقض فرض همواری (یا حتی پیوستگی) G-ساختار برطرف کرد. به عنوان مثال، نوار موبیوس در شکل ؟؟ غیرقابل جهت‌گیری است. این بدان معنی است که یک تعریف هموار و سراسری سازگار از دست‌پری چارچوب و در نتیجه ساختار $\{e\}$ (میدان چارچوب هموار سراسری) را نمی‌پذیرد. همانطور که در شکل ؟؟ بصری‌سازی شده است، یک G-اطلس از گیج‌ها که نوار موبیوس را پوشش می‌دهد، به طور ناگزیر نیازمند یک بازتاب در یکی از نگاشت‌های گذار خواهد بود، که به معنای یک گروه ساختار کاهش‌ناپذیر $G = \mathcal{R}$ است. CNNهای مستقل از مختصات بر روی نوار موبیوس بنابراین حداقل باید تناوب‌پذیر بازتابی باشند. به طور مشابه، گروه ساختار کره را نمی‌توان فراتر از $G = \mathrm{SO}(۲)$ کاهش داد. CNNهای کروی هموار بنابراین لزوماً بر پایه کرنل‌های محلی تناوب‌پذیر چرخش هستند.

توجه داشته باشید که هر منیفولد (دیفرانسیل‌پذیر) با برخی G-ساختار همراه است. به عنوان مثال، یک منیفولد دیفرانسیل‌پذیر خام دارای یک ساختار $\mathrm{GL}(d)$ (حاوی هر چارچوب ممکن) است، یک منیفولد ریمانی یک ساختار d دارد و $\mathbb{R}^d$ به طور کانونی با یک ساختار $\{e\}$ مجهز شده است که در شکل ؟؟ بصری‌سازی شده است. ما بنابراین بدون از دست دادن کلیت، اصطلاح "استقلال از مختصات" را به استقلال از مختصات GM تصحیح خواهیم کرد، یعنی استقلال نسبت به انتخاب چارچوب‌های مرجع در G-ساختار داده شده در M. در طول این کار ما فرض خواهیم کرد که گیج‌ها بخشی از یک G-اطلس هستند:

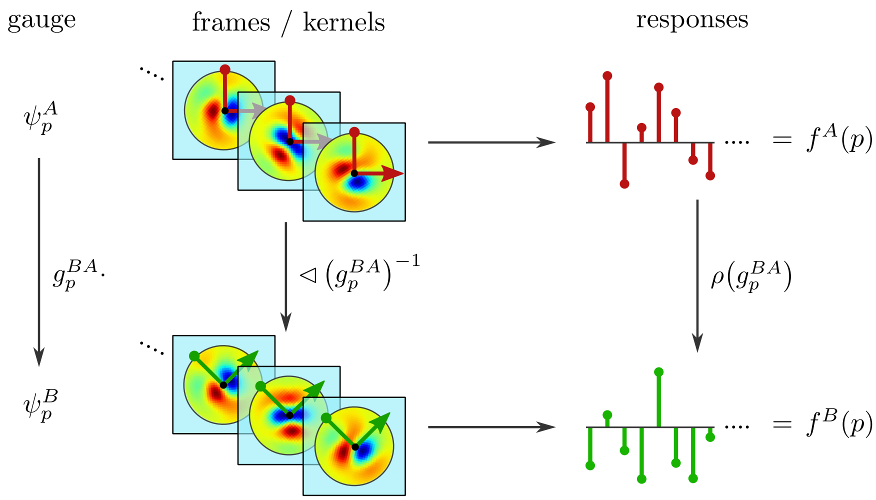
$$\mathcal{A}^G \,=\, \big\{(U^X, \psi^X)\big\}_{X\in\mathfrak{X}} \quad \text{□□□□ □□□□} \quad g_p^{BA} \in G \quad \text{□□□ □□□} \quad \psi_p^A, \psi_p^B \in \mathcal{A}^G,\ p \in U^A \cap U^B\,, \qquad (۲۵)$$

مطابق با G-ساختار داده شده. هر کمیت یا تابع می‌تواند نسبت به هر گیجی از این اطلس بیان شود[۱۱]، و مختصات‌بندی‌ها در گیج‌های مختلف به طور منحصر به فرد توسط یک تبدیل گیج با مقدار G مرتبط هستند. با تضمین استقلال از مختصات همه ساختارها، آن‌ها همیشه با همتایان مستقل از مختصات خود مطابقت خواهند داشت، که ما نظریه سراسری را بر حسب آن‌ها در بخش‌های ؟؟، ؟؟ و ؟؟ فرمول‌بندی خواهیم کرد.

۲.۳ میدان‌های بردار ویژگی مستقل از مختصات

فضاهای ویژگی شبکه‌های کانولوشنی مستقل از مختصات، فضاهای میدان‌های بردار ویژگی هستند. مشابه مورد ضرایب بردارهای مماس، ضرایب عددی بردارهای ویژگی ملزم به تبدیل سازگار تحت تبدیل‌های گیج هستند. قانون تبدیل خاص (نمایش گروهی) یک میدان ویژگی در اینجا نوع میدان آن را مشخص می‌کند □ مثال‌های معمول شامل میدان‌های اسکالر، میدان‌های بردار مماس، میدان‌های تانسور عمومی، میدان‌های ویژگی منظم یا میدان‌های □□□□□ هستند. بخش ۱.۲.۳ چنین میدان‌های ویژگی و قوانین تبدیل آن‌ها را معرفی می‌کند. در بخش ۲.۲.۳، ما به طور مختصر فضاهای ویژگی مستقل از مختصات را تعریف می‌کنیم. مشابه تعریف فضاهای ویژگی شبکه‌های کانولوشنی معمولی به عنوان انباشتی از چندین نقشه ویژگی، فضاهای ویژگی شبکه‌های کانولوشنی مستقل از مختصات شامل چندین میدان ویژگی مستقل هستند.

[۱۱] این یک گزاره غیربدیهی است زیرا هر کمیتی را نمی‌توان نسبت به چارچوب‌های مرجع دلخواه مرتبط با $\mathrm{GL}(d)$ بیان کرد. به عنوان مثال، میدان‌های ویژگی که در بخش ۲.۳ معرفی شده‌اند، تنها تبدیل‌های گیج با مقدار G را می‌پذیرند و بنابراین تنها نسبت به چارچوب‌های ترجیحی در GM تعریف می‌شوند. به عنوان یک مثال شهودی، CNNهای سنتی (غیرتناوب‌پذیر) در $\mathbb{R}^d$ را در نظر بگیرید که نسبت به ساختار $\{e\}$ کانونی $\mathbb{R}^d$ استخراج می‌شوند و اطلاعاتی در مورد پاسخ‌های کرنل نسبت به سایر چارچوب‌های مرجع را حمل نمی‌کنند.



شکل ۱۰: پاسخ‌های عددی $f^A(p) \in \mathbb{R}^c$ و $f^B(p) \in \mathbb{R}^c$ کرنل‌هایی که بر اساس چارچوب‌های مختلف جهت‌گیری شده‌اند، به طور کلی منطبق نیستند. به منظور نمایش ضرایب عددی یکسان بردار ویژگی مستقل از مختصات نسبت به گیج انتخاب شده، آن‌ها ملزم به ارتباط توسط تبدیل‌های گیج $\rho(g_p^{BA})$ هستند اگر گیج‌ها توسط g_p^{BA} مرتبط باشند. همانطور که در بخش ۴ استخراج شده است، این الزام یک قید تناوب‌پذیری گیج را بر کرنل‌های کانولوشن تحمیل می‌کند.

۱.۲.۳ میدان‌های بردار ویژگی منفرد

میدان‌های ویژگی کانولوشنی یک بردار ویژگی، که اطلاعات استنباط شده از یک همسایگی محلی از سیگنال ورودی را کدگذاری می‌کند، به هر نقطه از منیفولد تخصیص می‌دهند. انباشت فضایی اطلاعات توسط یک کرنل کانولوشنی انجام می‌شود که میدان‌های ویژگی را در محیط خود نسبت به چارچوب مرجع محلی خود اندازه‌گیری می‌کند. بنابراین ما گیج ψ^A را فرض می‌کنیم که ترازهای کرنل را در یک همسایگی U^A مشخص می‌کند. نسبت به این گیج، کرنل یک میدان محلی هموار از پاسخ‌ها (مشاهدات) تولید خواهد کرد:

$$f^A : U^A \to \mathbb{R}^c\,, \tag{۲۶}$$

که توسط یک بردار ویژگی عددی c-بعدی $f^A(p)$ در هر موقعیت $p \in U^A$ داده می‌شود. فرض کنید یک میدان پاسخ دوم $f^B : U^B \to \mathbb{R}^c$ داده شده باشد که نسبت به گیج ψ^B در U^B استنباط شده است. از آنجا که پاسخ یک کرنل به طور کلی به تراز آن بستگی دارد، انتظار می‌رود که f^A و f^B در همپوشانی $U^A \cap U^B$ مطابقت نداشته باشند. بدون محدودیت‌های بیشتر، پاسخ‌های یک کرنل کانولوشن به طور دلخواه وابسته به گیج خواهند بود.

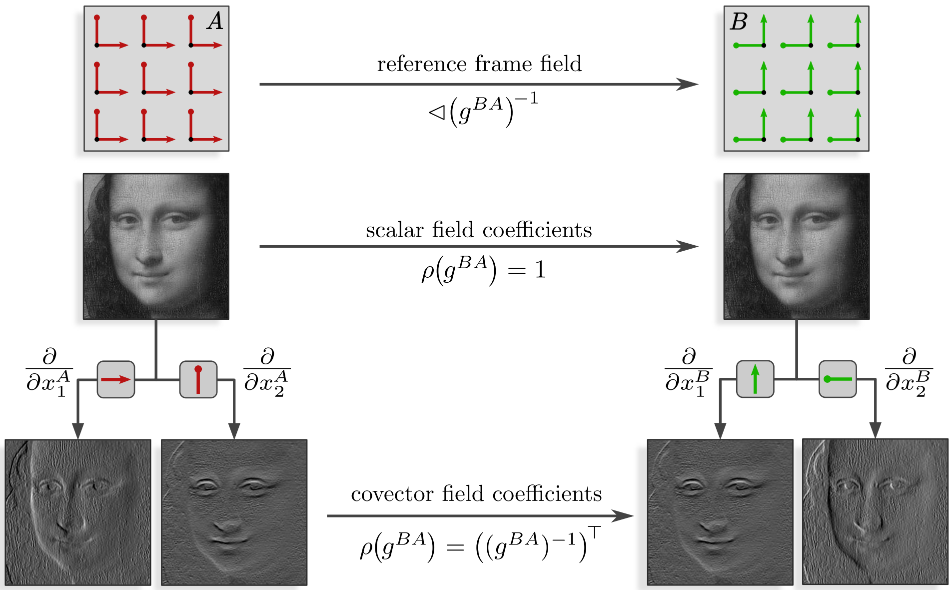
اصل کوواریانس، که توسط آلبرت انیشتین پیشنهاد شد [؟ ؟]، بیان می‌کند که:

> «قوانین جهانشمول طبیعت باید توسط معادلاتی بیان شوند که برای همه سیستم‌های مختصات صادق باشند، یعنی نسبت به هر جایگزینی کوواریانت باشند.»

ما معتقدیم که اصل مشابهی باید در یادگیری عمیق هندسی نیز صادق باشد، یعنی استنتاج باید مستقل از هر دلخواهی در انتخاب چارچوب‌های مرجع باشد. با توجه به اینکه این دلخواهی در مختصات‌بندی‌ها دقیقاً توسط G-ساختار داده شده GM پوشش داده می‌شود، این امر به ویژه ایجاب می‌کند که ویژگی‌ها باید اشیاء هندسی مستقل از مختصات GM باشند.[۱۲] بنابراین ما کرنل‌های کانولوشن را طوری طراحی می‌کنیم که پاسخ‌های آن‌ها f^A و f^B میدان‌هایی از ضرایب بردار ویژگی را کدگذاری کنند که یک میدان بردار ویژگی آزاد از مختصات f را به صورت محلی در گیج‌های مختلف نمایش می‌دهند. مجموعه‌ای از چنین میدان‌های ضریب عددی f^X، که نسبت به G-اطلس گیج‌های ψ^X در همسایگی‌های U^X که M را پوشش می‌دهند بیان شده‌اند، معادل میدان ویژگی سراسری و آزاد از مختصات f در M است.

برای اینکه این میدان ویژگی آزاد از مختصات به خوبی تعریف شود، یعنی مستقل از مختصات GM باشد، میدان‌های ضریب محلی (یا پاسخ‌های کرنل) ملزم به دوخته شدن سازگار از طریق نگاشت‌های گذار با مقدار G هستند. بنابراین آن‌ها باید به روشی اصولی تحت تبدیل‌های گیج تبدیل شوند. از آنجا که ما با فضاهای بردار ویژگی سر و کار داریم، این تبدیل‌ها معمولاً خطی در نظر گرفته می‌شوند، یعنی آن‌ها توسط نمایش‌های گروهی خطی مدل

[۱۲] در این نکته ما از کوواریانس عمومی انیشتین منحرف می‌شویم، که همیشه تبدیل‌های گیج با مقدار $\mathrm{GL}(d)$ را در نظر می‌گیرد (متناظر با کوواریانس دیفئومورفیسم). تنظیم او در فرمول‌بندی ما برای $G = \mathrm{GL}(d)$ شامل شده است، با این حال، ما گروه ساختار فرضی را انعطاف‌پذیر نگه می‌داریم زیرا اکثر کاربردها گروه ساختار کاهش یافته را فرض خواهند کرد.



شکل ۱۱: مثال‌هایی از میدان‌های ضریب ویژگی در $M = \mathbb{R}^2$ از پردازش کلاسیک تصویر. بالا: برای سادگی ما یک میدان چارچوب «موازی» فرض می‌کنیم و همان تبدیل گیج، چرخش به میزان $\pi/2$، را در هر نقطه $p \in M$ در نظر می‌گیریم. وسط: مقادیر شدت یک تصویر خاکستری مستقل از انتخاب چارچوب‌های مرجع هستند. بنابراین آن‌ها توسط میدان‌های اسکالر مدل می‌شوند که با نمایش بدیهی $\rho(g) = 1\ \forall g \in G$ مشخص می‌شوند. پایین: دو کانال ضریب یک تصویر گرادیان از یک تصویر اسکالر با گرفتن مشتق‌ها در امتداد محورهای چارچوب محاسبه می‌شوند □ بنابراین آن‌ها وابسته به گیج هستند. تصاویر گرادیان نسبت به گیج‌های مختلف توسط نمایش گروهی $\rho(g) = (g^{-1})^\top$ مرتبط هستند و بنابراین به عنوان میدان‌های کوبردار (میدان‌های تانسور از نوع (۰, ۱) یا ۱-فرم‌ها) شناسایی می‌شوند. برای چرخش نمایش داده شده به میزان $\pi/2$ این منجر به کانال اول جدید $(\partial/\partial x_1^B)$ معادل کانال دوم قدیمی $(\partial/\partial x_2^A)$ و کانال دوم جدید $(\partial/\partial x_2^B)$ معادل منفی کانال اول قدیمی $(\partial/\partial x_1^A)$ می‌شود. نسبت به چارچوب‌های مرجع مربوطه، هر دو میدان ضریب یکسان میدان گرادیان (آزاد از مختصات) را کدگذاری می‌کنند. بنابراین توصیف به طور خودکار مستقل از مختصات است.

می‌شوند:

$$\rho : G \to \mathrm{GL}(c) \tag{۲۷}$$

از گروه ساختار $G \leq \mathrm{GL}(c)$، که بر $\mathbb{R}^c$ عمل می‌کند و شرط $\rho(gh) = \rho(g)\rho(h)\ \forall\, g,h \in G$ را برآورده می‌کند.[۱۳] مشابه تبدیل ضرایب بردار مماس در معادله (۹)، ضرایب بردار ویژگی سپس تعریف می‌شوند که تحت یک تبدیل گیج با مقدار G مانند $g_p^{BA} = \psi_p^B \circ \big(\psi_p^A\big)^{-1}$ تبدیل شوند:

$$f^B(p) \ :=\ \rho\big(g_p^{BA}\big) f^A(p)\,, \tag{۲۸}$$

که در آن $p \in U^A \cap U^B$؛ برای بصری‌سازی به شکل ۱۰ مراجعه کنید. با ساخته شدن برای تبدیل همگام، فضاهای چارچوب‌های مرجع، ضرایب بردار مماس و ضرایب بردار ویژگی گفته می‌شود که با یکدیگر G-مرتبط هستند. توجه داشته باشید که ساخت از طریق یک G-نمایش ρ به طور کلی تبدیل‌های گیج با مقدار $\mathrm{GL}(d)$ را توصیف نمی‌کند، یعنی ویژگی‌های کاملاً مستقل از مختصات. بنابراین بردارهای ویژگی استخراج شده تنها بیان به خوبی تعریف شده‌ای نسبت به چارچوب‌ها در G-ساختار در نظر گرفته شده GM خواهند داشت، که توسط اصطلاح «استقلال از مختصات GM» پوشش داده می‌شود.

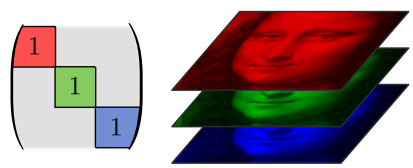
انتخاب‌های مختلف نمایش‌ها ρ_i انواع مختلفی از میدان‌های ویژگی را تولید می‌کنند همانطور که در شکل ۱۱ نمونه‌سازی شده است. به عنوان مثال، نمایش بدیهی، $\rho(g) = 1\ \forall g \in G$، رفتار تبدیل میدان‌های اسکالر $s^A(p) \ \mapsto\ s^B(p) \ :=\ 1 \cdot s^A(p)$ را توصیف می‌کند، که ضرایب عددی آن‌ها تحت تبدیل‌های گیج ناوردا هستند. مثال‌هایی از میدان‌های اسکالر شامل تصاویر خاکستری، میدان‌های دما، میدان‌های فشار یا توزیع‌های احتمال در M هستند. ضرایب میدان‌های بردار مماس مانند $v^A(p) \ \mapsto\ v^B(p) \ :=\ g_p^{BA} v^A(p)$ تبدیل می‌شوند و بنابراین با نمایش گروهی $\rho(g) = g$ مطابقت دارند. مثال‌هایی برای میدان‌های بردار شامل جریان نوری یا میدان‌های سرعت باد هستند. میدان‌های تانسور عمومی‌تر از نوع (r,s) توسط نمایش‌های حاصل‌ضرب تانسوری $\rho(g) \ =\ \otimes^s \left(g^{-1}\right)^\top \ \otimes^r g$ توصیف می‌شوند. آن‌ها به عنوان مثال تصاویر تانسور انتشار، تانسورهای میدان

[۱۳] این شرط تضمین می‌کند که نمایش‌ها همومورفیسم‌های گروهی هستند، یعنی نگاشت‌هایی که ساختار گروهی G را رعایت می‌کنند. بنابراین عمل‌های گروه ساختار بر فضای مماس و فضاهای ضریب بردار ویژگی سازگار هستند.

الکترومغناطیسی یا تانسورهای تنش را مدل می‌کنند. انتخاب رایج برای گروه‌های ساختار گسسته نمایش‌های منظم هستند که مجموعه محدود عملیات گروهی را توسط ماتریس‌های جایگشت تحقق می‌بخشند. نمایش‌های منظم به عنوان تقارن‌های دقیق شبکه‌های کریستالی، شبکه‌های اسپین یا شبکه‌های پیکسل ظاهر می‌شوند [? ? ? ?]. علاوه بر این، آن‌ها معمولاً به عنوان تقریب گسسته گروه‌های ساختار پیوسته استفاده می‌شوند، به عنوان مثال گروه‌های چرخشی $C_N \leq \mathrm{SO}(۲)$ برای تقریب چرخش‌های پیوسته [? ? ? ?]. آن‌ها از اهمیت عملی زیادی برخوردارند زیرا تبدیل ویژگی‌های شبکه‌های کانولوشنی گروهی را توصیف می‌کنند [?]. میدان‌های ویژگی که تحت نمایش‌های کاهش‌ناپذیر (□□□□□) تبدیل می‌شوند در [? ? ? ? ? ?] بررسی شدند.[۱۴] مروری دقیق‌تر و معیاری گسترده از انواع میدان یا نمایش‌های مختلف در یادگیری عمیق در [?] ارائه شد.

برای کامل بودن می‌خواهیم اشاره کنیم که میدان‌های بردار ویژگی آزاد از مختصات به طور رسمی به عنوان برش‌های هموار $f \in \Gamma(\mathcal{A})$ از یک بندل بردار ویژگی $\mathcal{A} \xrightarrow{\pi_\mathcal{A}} M$ تعریف می‌شوند که با G-ساختار GM مرتبط است و فضاهای ضریب بردار ویژگی $\mathbb{R}^c$ را به عنوان فیبرهای معمول دارد. بردارهای ضریب $f^A(p)$ و $f^B(p)$ در $\mathbb{R}^c$ تسهیمات محلی از یک بردار ویژگی آزاد از مختصات $f(p) \in \mathcal{A}_p \cong \mathbb{R}^c$ هستند، و مشابه ضرایب $v^A = \psi_p^A(v)$ و $v^B = \psi_p^B(v)$ از یک بردار مماس $v \in T_pM$ تعریف می‌شوند. توجه داشته باشید که، در حالی که هم‌ریخت هستند، فضاهای ویژگی $\mathcal{A}_p \cong \mathcal{A}_q$ در نقاط مختلف $p \neq q$ از M از یکدیگر متمایز هستند، به طوری که عناصر آن‌ها نمی‌توانند با هم جمع شوند. انتقال‌دهنده‌های موازی، که در بخش‌های ۳.۳ و ؟؟ مورد بحث قرار گرفته‌اند، هم‌ریختی‌هایی بین فضاهای بردار ویژگی مختلف فراهم می‌کنند، که امکان جمع ویژگی‌ها را (پس از انتقال آن‌ها به همان فضای برداری) فراهم می‌کند. از آنجا که این تعاریف کاملاً فنی هستند، ما جزئیات آن‌ها را فعلاً رد می‌کنیم و خواننده علاقه‌مند را به بخش ؟؟ ارجاع می‌دهیم.

۲.۲.۳ میدان‌های ویژگی انباشته و فضاهای ویژگی مستقل از مختصات



شکل ۱۲: سه کانال رنگ یک تصویر □□□ به عنوان میدان‌های اسکالر شناسایی می‌شوند، بنابراین فضای ویژگی کامل طبق $\rho(g) = (۱) \oplus (۱) \oplus (۱)$ تبدیل می‌شود.

فضاهای ویژگی شبکه‌های کانولوشنی معمولی شامل چندین نقشه ویژگی هستند. به طور مشابه، ما فضاهای ویژگی شبکه‌های کانولوشنی مستقل از مختصات را تعریف می‌کنیم که شامل چندین میدان ویژگی f_i از انواع بالقوه متفاوت ρ_i و ابعاد c_i باشند. بنابراین یک میدان کامل از فعالسازی‌های یک فضای ویژگی از یک شبکه کانولوشنی مستقل از مختصات به عنوان مجموع مستقیم تعریف می‌شود[۱۵]

$$f = \bigoplus_i f_i \tag{۲۹}$$

از میدان‌های منفرد. هر نقشه ویژگی از یک شبکه کانولوشنی معمولی موقعیت یک ویژگی خاص را کدگذاری می‌کند و به طور مستقل تبدیل می‌شود هنگامی که ورودی شبکه جابجا می‌شود. میدان‌های ویژگی منفرد f_i شبکه‌های کانولوشنی مستقل از مختصات GM هم موقعیت و هم G-پوز یک ویژگی را کدگذاری می‌کنند. در مقابل نقشه‌های ویژگی معمولی، میدان‌های ضریب آن‌ها، به عنوان مثال f_i^A، علاوه بر این تضمین شده‌اند که به طور مستقل از یکدیگر تحت تبدیل‌های گیج همانطور که توسط نوع آن‌ها $\rho_i : G \to \mathrm{GL}(c_i)$ مشخص شده است، تبدیل شوند. یک نمایش عددی محلی $f^A = \bigoplus_i f_i^A$ از میدان ویژگی کامل در معادله (۲۹) بنابراین طبق مجموع مستقیم نمایش‌های منفرد تبدیل می‌شود، یعنی:

$$\rho = \bigoplus_i \rho_i\,. \tag{۳۰}$$

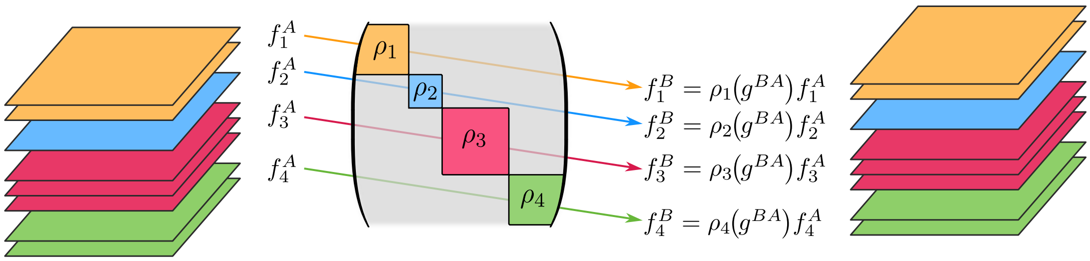
تبدیل مستقل میدان‌های منفرد تحت ρ، که در شکل ۱۳ بصری‌سازی شده است، از ساخت روشن است:

$$\rho(g) f^A \;=\; \big(\bigoplus_i \rho_i(g)\big)\big(\bigoplus_i f_i^A\big) \;=\; \bigoplus_i \big(\rho_i(g) f_i^A\big) \tag{۳۱}$$

به عنوان یک مثال عملی از فضای ویژگی مستقل از مختصات متشکل از چندین میدان، تصویر □□□ را همانطور که در شکل ۱۲ نشان داده شده است، در نظر بگیرید. مانند تصویر خاکستری در شکل ۱۱، کانال‌های رنگ منفرد مقادیر شدت را کدگذاری می‌کنند که تحت تبدیل‌های گیج ناوردا هستند. بنابراین تصویر □□□ کامل باید با سه میدان اسکالر شناسایی شود که هر کدام به طور مستقل تحت نمایش بدیهی «تبدیل» می‌شوند. همه میدان‌های ویژگی منفرد نیازی نیست که از همان نوع ρ_i باشند. به عنوان مثال، در یک کاربرد پیش‌بینی آب و هوا، سیگنال ورودی ممکن است شامل میدان‌های اسکالر کدگذاری کننده ویژگی‌هایی مانند دما یا فشار و میدان‌های برداری مانند سرعت‌های باد باشد. توصیف به عنوان میدان‌های ρ_i از انواع مربوطه، پردازش هندسی صحیح چنین داده‌هایی را تضمین می‌کند. در حالی که انواع میدان ρ_i ورودی و خروجی یک شبکه معمولاً توسط وظیفه یادگیری داده می‌شوند، انواع میدان استفاده شده در لایه‌های مخفی توسط کاربر به عنوان یک فراپارامتر مشابه انتخاب کانال‌ها برای یک شبکه کانولوشنی معمولی انتخاب می‌شوند.

[۱۴] توسط قضیه □□□□–□□□□□، هر نمایش یکانی از یک گروه فشرده می‌تواند از طریق تغییر پایه به مجموع مستقیم □□□□□ تجزیه شود. این دلالت دارد بر اینکه هر عملیات شبکه عصبی خطی بین نمایش‌های عمومی می‌تواند (پس از تغییر پایه) بر حسب عملیات بین □□□□□ درک شود [? ?]. در مقابل، انتخاب خاص نمایش، یعنی تغییر پایه نسبت به □□□□□ موجود در آن، برای هر لایه شبکه غیرخطی اهمیت دارد.

[۱۵] مجموع مستقیم $\oplus$ بردارهای $f_i(p)$ را می‌توان به عنوان «انباشت» آن‌ها در یک بردار پیوسته در نظر گرفت. به طور سازگار با این، مجموع مستقیم نمایش‌های ρ_i را می‌توان به عنوان ساخت یک ماتریس قطری بلوکی حاوی ρ_i به عنوان بلوک‌ها در نظر گرفت؛ به اشکال ۱۲ و ۱۳ مراجعه کنید.



شکل ۱۳: یک فضای ویژگی کامل شامل چندین میدان ویژگی منفرد f_i از انواع بالقوه متفاوت ρ_i و ابعاد c_i است. از طریق گیج ψ^A، به صورت محلی توسط میدان‌های ضریب $f_i^A : U^A \to \mathbb{R}^{c_i}$ نمایش داده می‌شود. میدان‌های ضریب در گیج دیگر ψ_B از طریق تبدیل گیج $f_i^B = \rho_i\big(g^{BA}\big) f_i^A$ مرتبط هستند. ضرایب هر میدان منفرد به طور مستقل تبدیل می‌شوند، بنابراین نمایش مدل‌سازی کل فضای ویژگی توسط مجموع مستقیم داده می‌شود، در اینجا $\bigoplus_i \rho_i = \rho_۱ \oplus \rho_۲ \oplus \rho_۳ \oplus \rho_۴$.

۳.۳ انتقال موازی بردارهای ویژگی

کرنل‌های شبکه‌های کانولوشنی ویژگی‌ها را از تمام نقاط q در یک همسایگی اطراف هر نقطه p از منیفولد جمع‌آوری می‌کنند. از آنجا که ویژگی‌ها در نقاط مختلف در فضاهای بردار ویژگی متفاوتی قرار دارند و نسبت به گیج‌های مختلف بیان می‌شوند، باید قبل از پردازش بیشتر، در امتداد یک مسیر γ از q به p به صورت موازی منتقل شوند. ما ابتدا انتقال بردارهای مماس را مورد بحث قرار می‌دهیم، که توسط یک نگاشت انتقال موازی رسمی‌سازی می‌شود:

$$\mathcal{P}_\gamma : T_qM \to T_pM\,. \tag{۳۲}$$

این انتقال‌دهنده اغلب از اتصال کانونی لوی-چویتا منیفولد محاسبه می‌شود، با این حال، ممکن است در برخی کاربردها مربوط به یک اتصال جایگزین (سازگار با G) باشد، همانطور که در ادامه و در مرور ادبیات ما در بخش ?? بیشتر مورد بحث قرار می‌گیرد. انتقال‌دهنده بردارهای ویژگی (مرتبط با G) از انتقال‌دهنده بردارهای مماس تبعیت می‌کند اگر انتقال سازگار با G باشد.

۱.۳.۳ انتقال‌دهنده‌های بردار مماس

از نظر آموزشی معقول است که ابتدا با حالت خاص انتقال‌دهنده‌های لوی-چویتا در فضاهای اقلیدسی که در شکل ۱۴آ به تصویر کشیده شده‌اند، شروع کنیم، قبل از پرداختن به انتقال‌دهنده‌ها و منیفلدهای عمومی‌تر. در این حالت انتقال موازی مستقل از مسیر انتخاب شده γ است و بردار منتقل شده را به معنای معمول در فضاهای اقلیدسی موازی نگه می‌دارد. توجه داشته باشید که انتقال‌دهنده $\mathcal{P}_\gamma$ بین فضاهای مماس T_qM و T_pM نگاشت می‌کند و بنابراین مستقل از مختصات است. با این حال، می‌توان آن را نسبت به مختصات بیان کرد، سپس به جای بردارهای مماس بر روی بردارهای ضریب عددی عمل می‌کند. شهودی در شکل ۱۴آ آورده شده است، جایی که چارچوب‌ها در q و p موازی نیستند[16] به طوری که ضرایب $(۱,۱)^\top$ در q و $(\sqrt{۲},۰)^\top$ در p متفاوت هستند، حتی اگر بردارهای مماس (مستقل از مختصات) مربوطه با یکدیگر موازی باشند. برای دقیق‌تر کردن این موضوع، گیج‌های $\psi_q^{\widetilde{A}}$ و ψ_p^A را در همسایگی‌های $U^{\widetilde{A}}$ از q (قرمز) و U^A از p (سبز) در نظر بگیرید. فرض کنید یک بردار $v = \big(\psi_q^{\widetilde{A}}\big)^{-۱}(v^{\widetilde{A}}) \in T_qM$ با ضرایب $v^{\widetilde{A}} \in \mathbb{R}^d$ داده شده باشد. ضرایب بردار منتقل شده $\mathcal{P}_\gamma v$ در p سپس توسط: $\psi_p^A \circ \mathcal{P}_\gamma(v) \;=\; \psi_p^A \circ \mathcal{P}_\gamma \circ \big(\psi_q^{\widetilde{A}}\big)^{-۱}(v^{\widetilde{A}})$ داده می‌شود. نتیجه می‌شود که بیان مختصاتی یک انتقال‌دهنده نسبت به گیج‌های $\widetilde{A}$ و A به صورت زیر است:[17]

$$g_\gamma^{A\widetilde{A}} \;:=\; \psi_p^A \circ \mathcal{P}_\gamma \circ \big(\psi_q^{\widetilde{A}}\big)^{-۱} \in\; \mathrm{GL}(d) \tag{۳۳}$$

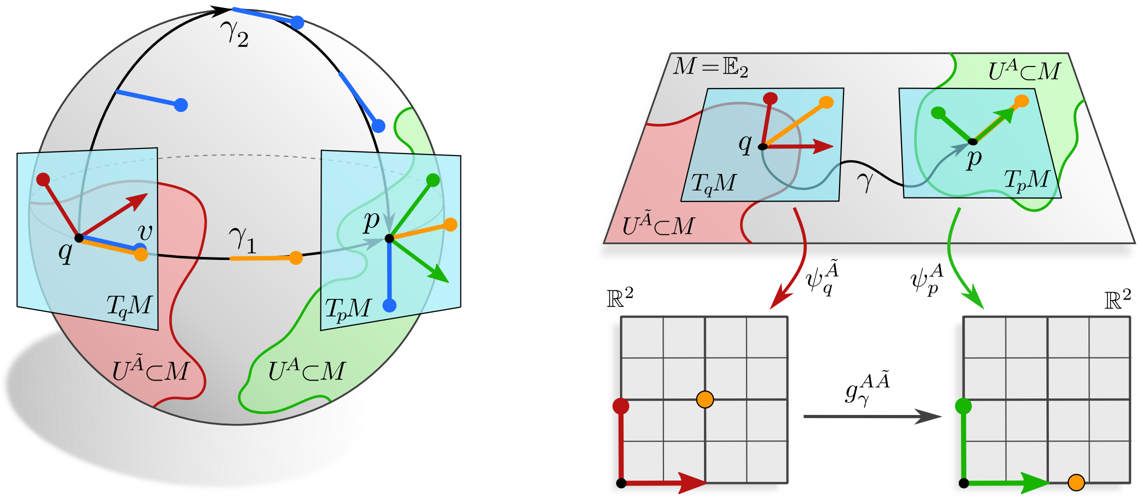
عنصر گروهی $g_\gamma^{A\widetilde{A}}$ انتخاب‌های غیرموازی چارچوب‌های مرجع در q و p را در نظر می‌گیرد. در $\mathbb{R}^d$، معمولاً فرض می‌شود که تمام چارچوب‌ها موازی هستند به طوری که تمام مختصات‌بندی‌های انتقال‌دهنده‌های لوی-چویتا بدیهی می‌شوند.[18]

---

[16] برخلاف منیفلدهای عمومی، $\mathbb{R}^d$ با مفهوم کانونی موازی بودن چارچوب‌های مرجع همراه است.

[17] $g_\gamma^{A\widetilde{A}}$ مقادیری در $\mathrm{GL}(d)$ می‌گیرد اگر ما اتصال‌های دلخواه (با مقدار $\mathfrak{gl}(d)$) و گروه‌های ساختار عمومی $G \leq \mathrm{GL}(d)$ را فرض کنیم. برای اتصال لوی-چویتا با مقدار $\mathfrak{so}(d)$ و چارچوب‌های متعامد، یعنی $G = d$، ما $g_\gamma^{A\widetilde{A}} \in d$ را داریم.

[18] سنتی در $\mathbb{R}^d$ به طور ضمنی این فرض را از چارچوب‌های موازی (شکل ??) و انتقال‌دهنده‌های بدیهی می‌کنند.



(آ) انتقال موازی و مختصات‌بندی آن در یک فضای مسطح.

(ب) انتقال موازی در ۲-کره $S^۲$.

شکل ۱۴: انتقال موازی بردارهای مماس $v \in T_qM$ در q به $\mathcal{P}_\gamma v \in T_pM$ در p. شکل ۱۴آ حالت خاص انتقال‌دهنده‌های لوی-چویتا را در فضاهای اقلیدسی مسطح $M =$ بصری‌سازی می‌کند. مستقل از مسیر انتخاب شده γ انتقال لوی-چویتا بردار (نارنجی) را به معنای معمول در فضاهای اقلیدسی موازی نگه می‌دارد. گیج‌های $\psi_q^{\widetilde{A}}$ (قرمز) و ψ_p^A (سبز) امکان بیان انتقال‌دهنده مستقل از مختصات را توسط یک عنصر گروهی $g_\gamma^{A\widetilde{A}} = \psi_p^A \circ \mathcal{P}_\gamma \circ \left(\psi_q^{\widetilde{A}}\right)^{-۱} \in \mathrm{GL}(d)$ فراهم می‌کنند که تغییر ضرایب بردار را در نظر می‌گیرد اگر چارچوب هدف با چارچوب منبع منتقل شده مطابقت نداشته باشد. شکل ۱۴ب انتقال لوی-چویتا را در ۲-کره $S^۲$ نشان می‌دهد، مقایسه کنید با معادله (؟؟). انتقال‌دهنده‌های $\mathcal{P}_{\gamma_۱}$ و $\mathcal{P}_{\gamma_۲}$ در امتداد مسیرهای مختلف $\gamma_۱$ و $\gamma_۲$ به طور کلی با یکدیگر اختلاف دارند. همانند فضاهای مسطح، انتقال‌دهنده‌های مستقل از مختصات را می‌توان توسط عناصر گروهی که بر روی ضرایب نسبت به چارچوب‌های مختصاتی در q و p عمل می‌کنند، بیان کرد.

از آنجا که انتقال‌دهنده در معادله (۳۳) وابسته به مختصات است، ما علاقه‌مند به تبدیل‌های گیج آن هستیم. گیج‌های $\psi_q^{\widetilde{B}}$ و ψ_p^B را دو گیج جایگزین در همسایگی‌های q و p در نظر بگیرید. از نمودار جابجایی:

$$\begin{tikzcd}
\mathbb{R}^d \arrow[rrrr, "g_\gamma^{A\widetilde{A}}\cdot"] \arrow[dd, "g_q^{\widetilde{B}\widetilde{A}}\cdot"'] & & & & \mathbb{R}^d \arrow[dd, "g_p^{BA}\cdot"] \\
& T_qM \arrow[ul, "\psi_q^{\widetilde{A}}"'] \arrow[dl, "\psi_q^{\widetilde{B}}"] \arrow[rr, "\mathcal{P}_\gamma"] & & T_pM \arrow[ur, "\psi_p^A"] \arrow[dr, "\psi_p^B"'] & \\
\mathbb{R}^d \arrow[rrrr, "g_\gamma^{B\widetilde{B}}\cdot"'] & & & & \mathbb{R}^d
\end{tikzcd} \qquad (۳۴)$$

می‌توان خواند که انتقال‌دهنده‌ها در گیج‌های مختلف توسط:

$$g_\gamma^{B\widetilde{B}} \;=\; g_p^{BA}\, g_\gamma^{A\widetilde{A}} \left(g_q^{\widetilde{B}\widetilde{A}}\right)^{-۱} \qquad (۳۵)$$

مرتبط هستند. توجه داشته باشید که شباهت این قانون تبدیل و نمودار جابجایی به آنچه در معادلات (۲۱) و (۲۳) آمده است. تفاوت بین هر دو این است که انتقال‌دهنده دارای دامنه T_qM و دامنه مشترک T_pM متفاوتی است که توسط گیج‌های مختلف و مستقل از یکدیگر تسهیم شده‌اند و بنابراین به طور مستقل تبدیل می‌شوند.

به طور کلی، انتقال موازی بردارهای مماس توسط یک انتخاب اتصال تعیین می‌شود، به عنوان مثال (اما نه لزوماً) توسط اتصال کانونی لوی-چویتا یک منیفولد ریمانی. یک اتصال را می‌توان به عنوان مجموعه‌ای از انتقال‌دهنده‌های بی‌نهایت کوچک بین فضاهای مماس مجاور در نظر گرفت، به طوری که انتقال‌دهنده کامل $\mathcal{P}_\gamma$ با انتگرال‌گیری اتصال در امتداد مسیر γ به دست می‌آید. انتقال‌دهنده‌ها در امتداد مسیرهای مختلف $\gamma_۱$ و $\gamma_۲$ از q به p نیازی به مطابقت ندارند، که در شکل ۱۴ب با انتقال‌دهنده‌های لوی-چویتا در ۲-کره $S^۲$ مثال زده شده است، مقایسه کنید با معادله (؟؟). همانند فضاهای مسطح، انتقال‌دهنده‌های مستقل از مختصات را می‌توان توسط معادله (۳۳) نسبت به گیج‌ها بیان کرد. تبدیل‌های گیج چنین انتقال‌دهنده‌های مختصات‌بندی شده دوباره توسط معادله (۳۵) داده می‌شوند. انتقال‌دهنده‌ها در یک منیفولد داده شده می‌توانند در اصل به صورت تحلیلی از اتصال محاسبه شوند [؟ ؟] و گاهی اوقات می‌توانند به صورت فرم بسته بیان شوند، به عنوان مثال برای کره $S^۲$، معادله (؟؟). چندین الگوریتم عددی برای محاسبه انتقال‌دهنده‌های

موازی در مش‌ها وجود دارد؛ به بخش ؟؟ مراجعه کنید. ما به جزئیات بیشتر در مورد نحوه محاسبه انتقال‌دهنده‌های بردار مماس $\mathcal{P}_\gamma$ نخواهیم پرداخت بلکه به سادگی آن‌ها را داده شده فرض می‌کنیم.

۲.۳.۳ انتقال‌دهنده‌های بردار ویژگی

معادله (۲۸) قانون تبدیل ضرایب بردار ویژگی را توسط نوع میدان آن‌ها (ρ) تعریف می‌کند. انتقال‌دهنده موازی آن‌ها، که نسبت به گیج‌های $\psi_q^{\widetilde{A}}$ و ψ_p^{A} بیان می‌شود، به طور مشابه با پیچیدن انتقال‌دهنده ضریب بردار مماس در این نمایش میدان داده می‌شود، یعنی توسط:

$$\rho\big(g_\gamma^{A\widetilde{A}}\big)\,. \tag{۳۶}$$

توجه داشته باشید که □ از آنجا که $\rho\ :\ G\ \to\ \mathrm{GL}(c)$ یک نمایش G است □ این ساختار تنها زمانی به خوبی تعریف می‌شود که تمام انتقال‌دهنده‌ها $g_\gamma^{A\widetilde{A}}$ (برای مسیرهای دلخواه γ و چارچوب‌های A، $\widetilde{A}$) واقعاً مقادیری در گروه ساختار انتخاب شده G بگیرند. اینکه آیا این حالت صادق است یا خیر هم به انتخاب خاص G-ساختار (یا G-اطلس) و هم به انتقال‌دهنده‌ها (یا اتصال) مورد نظر بستگی دارد □ آن‌ها باید سازگار باشند [؟].

تمام شبکه‌های کانولوشنی (بنابراین انتقال می‌دهند) بردارهای ویژگی را به نحوی جمع‌آوری می‌کنند، و بنابراین برخی انتخاب از اتصال و G-ساختار را فرض می‌کنند. اگر G-ساختار انتخاب شده با اتصال لوی-چویتا ناسازگار باشد، این به معنای آن است که این مدل‌ها □ اغلب به طور ضمنی □ یک اتصال جایگزین و سازگار با G را برای جمع‌آوری ویژگی‌ها فرض می‌کنند. خواننده فعلاً نباید نگران انتخاب‌های خاص اتصال‌ها باشد، که در مرور ادبیات ما در بخش ؟؟ روشن‌تر خواهد شد. در ادامه این بخش، ما بیشتر در مورد سازگاری G اتصال‌ها و G-ساختارها توضیح خواهیم داد. با فرض اینکه انتقال‌دهنده‌های ویژگی در ادامه همیشه به خوبی تعریف خواهند شد، این بخش را می‌توان با خیال راحت در اولین مطالعه نادیده گرفت.

بحث دقیق‌تر و مستقل از مختصات انتقال‌دهنده‌ها در بندل‌های بردار ویژگی مرتبط را می‌توان در بخش ؟؟ یافت.

۳.۳.۳ سازگاری اتصال‌ها و □-ساختارها

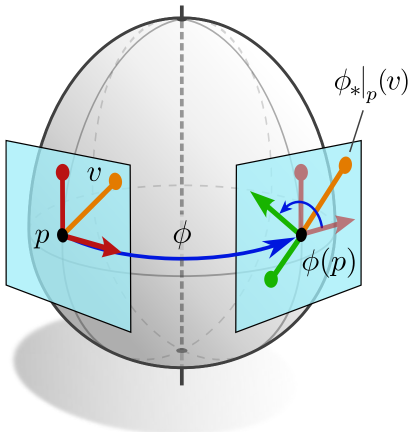
یک اتصال با یک G-ساختار GM سازگار با G نامیده می‌شود اگر عبارات مختصاتی $g_\gamma^{A\widetilde{A}}$ انتقال‌دهنده‌های $\mathcal{P}_\gamma$ آن نسبت به هر چارچوب A، $\widetilde{A}$ از GM مقادیری در گروه ساختار G بگیرند [؟].[19] یک اتصال سازگار با G منجر به انتقال‌دهنده‌های بردارهای ویژگی مرتبط با G می‌شود.

برای روشن کردن این شرط سازگاری تا حدی انتزاعی، چند مثال خاص را مورد بحث قرار می‌دهیم. یک مثال ساده، اتصال لوی-چویتا در $\mathbb{R}^۲$ است، شکل ۱۴آ. دو $\{e\}$-ساختار در $\mathbb{R}^۲$ که در اشکال ؟؟ و ؟؟ نشان داده شده‌اند را در نظر بگیرید. در اینجا $G=\{e\}$ است، به این معنی که نوع میدان $\rho:\{e\}\to\mathrm{GL}(c)$ یک نمایش $\{e\}$ است، به طوری که انتقال موازی بردارهای ویژگی تنها در صورتی می‌تواند تعریف شود که عبارات مختصاتی $g_\gamma^{A\widetilde{A}}$ مقادیری در $\{e\}$ بگیرند، یعنی بدیهی باشند. از آنجا که $\{e\}$-ساختار در شکل ؟؟ از چارچوب‌های "موازی" تشکیل شده است، این امر واقعاً همینطور است □ بنابراین اتصال لوی-چویتا با این $\{e\}$-ساختار سازگار است. در مقابل، چارچوب‌های $\{e\}$-ساختار در شکل ؟؟ نسبت به یکدیگر "چرخانده شده‌اند"، که منجر به عبارات مختصاتی غیربدیهی $g_\gamma^{A\widetilde{A}}$ می‌شود که مقادیری در $\mathrm{SO}(۲)$ می‌گیرند (در شکل ۱۴آ بصری‌سازی شده است). از آنجا که نوع میدان $\rho:\{e\}\to\mathrm{GL}(c)$ چرخش‌ها را مدیریت نمی‌کند، امکان تعریف انتقال لوی-چویتا ویژگی‌های مرتبط با این $\{e\}$-ساختار وجود ندارد □ آن‌ها ناسازگار هستند. به عنوان مثال دوم، اتصال لوی-چویتا در $S^۲$ را در نظر بگیرید، که در شکل ۱۴ب نشان داده شده است. انتقال در این حالت همیشه وابسته به مسیر خواهد بود و منجر به بردارهای چرخانده شده متفاوت خواهد شد، که دلالت دارد بر اینکه $g_\gamma^{A\widetilde{A}}$ مقادیری در $\mathrm{SO}(۲)$ خواهد گرفت. بردارهای ویژگی که باید مطابق با اتصال لوی-چویتا منتقل شوند، بنابراین باید از نوع $\rho:\mathrm{SO}(۲)\to\mathrm{GL}(c)$ باشند که یک نمایش $\mathrm{SO}(۲)$ است. این امر حداقل نیازمند $\mathrm{SO}(۲)$-ساختار در $S^۲$ است که در شکل ؟؟ نشان داده شده است. $\{e\}$-ساختار در $S^۲$ از شکل ؟؟ با اتصال لوی-چویتا ناسازگار است.

از آنجا که اتصال لوی-چویتا یک اتصال متریک است، طول و زاویه بین بردارهای مماس را حفظ می‌کند، و بنابراین چارچوب‌های متعامد را به چارچوب‌های متعامد منتقل می‌کند. نتیجه می‌شود که اتصال لوی-چویتا همیشه با d-ساختار چارچوب‌های متعامد سازگار است، که نسبت به آن $g_\gamma^{A\widetilde{A}}$ مقادیری در d می‌گیرد. اگر منیفولد قابل جهت‌گیری باشد، دست‌پری چارچوب توسط انتقال‌دهنده‌های لوی-چویتا حفظ می‌شود، به این معنی که آن‌ها تضمین شده‌اند که با $\mathrm{SO}(d)$-ساختارهای چارچوب‌های متعامد راست‌گرد در M سازگار باشند. تمام شبکه‌های کانولوشنی در مرور ادبیات ما در بخش ؟؟ که بر اساس $\mathrm{SO}(d)$-ساختارها هستند، ویژگی‌ها را از طریق انتقال‌دهنده‌های لوی-چویتا جمع‌آوری می‌کنند.

اگر یک G-ساختار داده شده با اتصال لوی-چویتا ناسازگار باشد، باید یک اتصال جایگزین و سازگار با G را برای انتقال بردارهای ویژگی تعریف کرد. برجسته‌ترین مثال در مرور ادبیات ما، اتصال‌های بدیهی در $\{e\}$-ساختارها است. یک اتصال بدیهی با خاصیت مستقل از مسیر بودن انتقال آن مشخص می‌شود [؟]. هر $\{e\}$-ساختار یک اتصال بدیهی منحصر به فرد را نتیجه می‌دهد، که بردارهای مماس را به گونه‌ای منتقل می‌کند که زاویه

[19] به طور معادل، فرم ۱-اتصال اتصال، که نسبت به چارچوب‌های GM بیان می‌شود، ملزم به داشتن مقدار $\mathfrak{g}$ است، که در آن $\mathfrak{g}$ نشان‌دهنده جبر لی G است. به طور انتزاعی‌تر، ما به اتصال‌های اصلی اهرسمن در بندل اصلی G (GM) علاقه‌مندیم.



شکل ۱۵: بصری‌سازی پیش‌برنده آزاد از مختصات بردارهای مماس و بیان مختصاتی آن نسبت به چارچوب‌های مرجع داده شده در موقعیت منبع و هدف. پیش‌برنده آزاد از مختصات $\phi_*|_p : T_pM \to T_{\phi(p)}M$ بردارهای مماس $v \in T_pM$ را به $\phi_*|_p(v) \in T_{\phi(p)}M$ (نارنجی) منتقل می‌کند. فرض کنید $\psi_p^{\widetilde{A}}$ گیج در p که متناظر با چارچوب مرجع قرمز است و $\psi_{\phi(p)}^{A}$ گیج در $\phi(p)$ که متناظر با چارچوب مرجع سبز است. آن‌ها بردارها را قبل و بعد از پیش‌برنده توسط ضرایب عددی $\psi_p^{\widetilde{A}}(v) = (۱,۱)^\top$ و $\psi_{\phi(p)}^{A}\big(\phi_*|_p(v)\big) = (۰,-\sqrt{۲})^\top$ توضیح می‌دهند. این تبدیل ضرایب بردار توسط تبدیل گیج القا شده توسط ایزومتری $g_\phi^{A\widetilde{A}}(p) \in \mathrm{GL}(d)$ توصیف می‌شود، یعنی $\psi_{\phi(p)}^{A}\big(\phi_*|_p(v)\big) = g_\phi^{A\widetilde{A}}(p)\cdot\psi_p^{\widetilde{A}}(v)$. ضرایب بردارهای ویژگی به طور مشابه طبق $\rho\big(g_\phi^{A\widetilde{A}}(p)\big)$ تبدیل می‌شوند اگر $g_\phi^{A\widetilde{A}}(p) \in G$.

یکسانی را با چارچوب‌های مرجع $\{e\}$-ساختار حفظ کنند. این امر دلالت دارد بر اینکه $g_\gamma^{A\widetilde{A}} = e$ یعنی آن‌ها بردارهای ضریب را در $\mathbb{R}^c$ (نسبت به چارچوب‌های $\{e\}$-ساختار) بدون تغییر مقادیر عددی آن‌ها منتقل می‌کنند. چنین انتقال‌دهنده‌هایی در شبکه‌های کانولوشنی استفاده می‌شوند که انتقال‌دهنده‌های غیربدیهی را به صراحت مدل نمی‌کنند که در مورد تمام شبکه‌های دارای $G = \{e\}$ در جدول ?? صادق است، به ویژه آن‌هایی که در بخش‌های ?? و ?? هستند. توجه داشته باشید که اتصال بدیهی تنها اتصالی است که با یک $\{e\}$-ساختار سازگار است.

همانطور که در بالا ذکر شد، هر شبکه کانولوشنی برخی انتخاب از G-ساختار و اتصال سازگار را فرض می‌کند، اغلب اتصال‌های لوی-چویتا یا اتصال‌های بدیهی.

بخش ?? به طور مفصل در مورد سازگاری انتقال‌دهنده‌ها و G-ساختارها از دیدگاه مستقل از مختصات توضیح می‌دهد.

۳.۴ عمل ایزومتری‌ها و تبدیل‌های گیج القا شده

تا کنون بحث ما منحصراً بر تقارن‌های گیج محلی در مختصات‌بندی فضاهای مماس متمرکز بوده است. یک منیفولد ممکن است، با این حال، خود دارای تقارن‌های غیربدیهی باشد، که در مورد یک منیفولد ریمانی M گروه ایزومتری آن $\mathrm{Isom}(M)$ را تشکیل می‌دهند. این بخش ایزومتری‌ها و عمل آن‌ها بر منیفلدها، بردارهای مماس، چارچوب‌های مرجع و میدان‌های ویژگی را به طور خلاصه مورد بحث قرار می‌دهد، نتایجی را خلاصه می‌کند که به طور دقیق‌تری در بخش ?? استخراج شده‌اند. ما بدین ترتیب معادلی بین عمل‌های فعال ایزومتری و تفسیر منفعل آن‌ها بر حسب تبدیل‌های گیج القا شده توسط ایزومتری را برجسته خواهیم کرد. این معادلی بعداً امکان توصیف تناوب‌پذیری ایزومتری GM-کانولوشن‌ها را فراهم خواهد کرد.

ایزومتری‌ها به عنوان تقارن‌های منیفلدهای ریمانی تعریف می‌شوند، یعنی آن نگاشت‌ها (دیفئومورفیسم‌ها)

$$\phi : M \to M, \tag{۳۷}$$

که متریک و در نتیجه فواصل روی M را حفظ می‌کنند. مجموعه همه ایزومتری‌های یک منیفولد ریمانی M گروه ایزومتری آن را تشکیل می‌دهد، که ما آن را با $\mathrm{Isom}(M)$ نشان می‌دهیم. به عنوان مثال، گروه اقلیدسی $\mathrm{E}(d)$ گروه ایزومتری فضاهای اقلیدسی $\mathbb{E}_d$ است. این گروه شامل انتقال‌ها، چرخش‌ها و بازتاب‌ها است، که همه آن‌ها متریک استاندارد $\mathbb{E}_d$ را حفظ می‌کنند. گروه ایزومتری ۲-کره $S^۲$ توسط گروه متعامد ۳ داده می‌شود، که شامل چرخش‌ها و بازتاب‌ها است. شکل ۱۵ منیفولد تخم‌مرغی شکلی را نشان می‌دهد که ایزومتری‌های آن چرخش‌ها و بازتاب‌ها در ۲ حول محور عمودی هستند.

۱.۴.۳ پیش‌برنده بردارهای مماس

هر ایزومتری $\phi \in \mathrm{Isom}(M)$ از طریق پیش‌برنده (یا دیفرانسیل) آن

$$\phi_*|_p :\; T_pM \to T_{\phi(p)}M\,, \tag{۳۸}$$

به طور طبیعی بر بردارهای مماس عمل می‌کند. همانطور که در شکل ?? (وسط) بصری‌سازی شده است، پیش‌برنده را می‌توان به طور شهودی به عنوان حمل بردارهای مماس همراه با عمل ایزومتری بر منیفولد زیرین M در نظر گرفت. تعریف رسمی پیش‌برنده بر روی TM در پیوست ?? آورده شده است، با این حال، شهود ارائه شده برای هدف ما کافی است. از آنجا که پیش‌برنده یک نگاشت خطی آزاد از مختصات بین فضاهای مماس است، عمل آن در مختصات توسط ماتریس $d \times d$ نمایش داده می‌شود. با فرض گیج‌های $\psi_p^{\widetilde{A}}$ و $\psi_{\phi(p)}^{A}$ در موقعیت منبع و هدف، به ترتیب، این ماتریس توسط

$$g_\phi^{A\widetilde{A}}(p) \;:=\; \psi_{\phi(p)}^{A} \circ \phi_*|_p \circ \big(\psi_p^{\widetilde{A}}\big)^{-۱} \;\in\; \mathrm{GL}(d)\,. \tag{۳۹}$$

داده می‌شود. این ماتریس تبدیل از ضرایب عددی یک بردار اصلی $v \in T_pM$ در گیج منبع و پیش‌برنده آن $\phi_*|_p(v) \in T_{\phi(p)}M$ در گیج هدف را توضیح می‌دهد، یعنی $\psi^A_{\phi(p)}\big(\phi_*|_p(v)\big) = g_\phi^{A\widetilde{A}}(p)\cdot\psi_p^{\widetilde{A}}(v)$. نمودار جابجایی

$$\begin{tikzcd}
\mathbb{R}^d \arrow[rrrr, "g_\phi^{A\widetilde{A}}(p)\cdot"] \arrow[dd, "g_p^{\widetilde{B}\widetilde{A}}\cdot"'] & & & & \mathbb{R}^d \arrow[dd, "g_{\phi(p)}^{BA}\cdot"] \\
& T_pM \arrow[ul, "\psi_p^{\widetilde{A}}"'] \arrow[dl, "\psi_p^{\widetilde{B}}"] \arrow[rr, "\phi_*|_p"] & & T_{\phi(p)}M \arrow[ur, "\psi^A_{\phi(p)}"'] \arrow[dr, "\psi^B_{\phi(p)}"] & \\
\mathbb{R}^d \arrow[rrrr, "g_\phi^{B\widetilde{B}}(p)\cdot"'] & & & & \mathbb{R}^d
\end{tikzcd} \quad , \tag{۴۰}$$

که از نظر مفهومی شبیه آن در معادله (۳۴) است، تعریف بیان مختصاتی پیش‌برنده بردار مماس را بصری‌سازی می‌کند. علاوه بر این دلالت دارد بر اینکه تبدیل‌های گیج بین مختصات‌بندی‌های مختلف توسط

$$g_\phi^{B\widetilde{B}} \;=\; g_{\phi(p)}^{BA}\; g_\phi^{A\widetilde{A}}\big(g_p^{\widetilde{B}\widetilde{A}}\big)^{-1}\,, \tag{۴۱}$$

داده می‌شوند، که معادل مفهومی معادله (۳۵) است.

۲.۴.۳ پیش‌برنده چارچوب‌های مرجع و تقارن‌های □-ساختار

از آنجا که چارچوب‌های مرجع فقط d-تاپل‌هایی از بردارهای چارچوب خطی مستقل هستند، پیش‌برنده بردارهای مماس پیش‌برنده‌ای از چارچوب‌های مرجع را با پیش‌بردن محورهای چارچوب منفرد القا می‌کند. به طور خاص، پیش‌برنده یک چارچوب $[e_i]_{i=1}^d$ در p به عنوان چارچوب $\big[\phi_*|_p(e_i)\big]_{i=1}^d$ در $\phi(p)$ تعریف می‌شود.

این پیش‌برنده چارچوب‌ها همیشه به خوبی تعریف شده است، با این حال، ممکن است با G-ساختار سازگار نباشد، یعنی به طور کلی تضمینی وجود ندارد که چارچوب‌ها در GM هنگام پیش‌برده شدن در GM باقی بمانند. به عنوان مثال $\{e\}$-ساختار در شکل ؟؟ (بالا چپ) را در نظر بگیرید، که توسط انتقال‌های افقی حفظ می‌شود اما توسط انتقال‌های عمودی یا هر ایزومتری دیگر از $\mathbb{R}^2$ حفظ نمی‌شود. به طور مشابه، $\mathcal{R}$-ساختار در شکل ؟؟ (پایین چپ) توسط انتقال‌ها و بازتاب‌های افقی حفظ می‌شود، اما توسط چرخش‌ها نه. بنابراین ما زیرگروه

$$\mathrm{Isom}_{GM} \;:=\; \Big\{\phi\in\mathrm{Isom}(M)\;\Big|\;\big[\phi_*(e_i)\big]_{i=1}^d \in GM \quad \forall\,[e_i]_{i=1}^d\in GM\Big\} \;\leq\; \mathrm{Isom}(M) \tag{۴۲}$$

ایزومتری‌هایی که تقارن‌های G-ساختار هستند را در نظر می‌گیریم، یعنی آن‌هایی که تضمین شده‌اند هر چارچوب در GM را به چارچوب دیگری که نیز در GM موجود است، نگاشت کنند.[۲۰] توجه داشته باشید که Isom_{GM} به طور کلی به انتخاب خاص G-ساختار GM بستگی دارد، نه فقط به گروه ساختار G. برای حالت خاص که $G \geq d$، تضمین شده است که $\mathrm{Isom}_{GM} = \mathrm{Isom}(M)$ مطابقت داشته باشند زیرا ایزومتری‌ها تضمین شده‌اند که چارچوب‌های متعامد را به چارچوب‌های متعامد نگاشت کنند. ما علاقه‌مند به زیرگروه Isom_{GM} هستیم زیرا تنها آن ایزومتری‌ها پیش‌برنده به خوبی تعریف شده‌ای از بردارهای ویژگی مستقل از مختصات GM را القا خواهند کرد، همانطور که در بخش بعدی بیشتر مورد بحث قرار گرفته است.

پیش از ادامه به عمل ایزومتری بر بردارهای ویژگی، ما آنچه را که تبدیل‌های گیج القا شده توسط ایزومتری می‌نامیم، مورد بحث قرار می‌دهیم. برای این منظور، فرض کنید $\big[e_i^{\widetilde{A}}\big]_{i=1}^d$ آن چارچوب در p باشد که متناظر با گیج منبع $\psi_p^{\widetilde{A}}$ است و فرض کنید $\big[e_i^A\big]_{i=1}^d$ آن چارچوب در $\phi(p)$ باشد که متناظر با گیج هدف $\psi^A_{\phi(p)}$ است، همانطور که در شکل ۱۵ به ترتیب در قرمز (چپ) و سبز (راست) نشان داده شده است. پیش‌برنده $\big[\phi_*|_p(e_i^{\widetilde{A}})\big]_{i=1}^d$ چارچوب منبع از p به $\phi(p)$ (قرمز شفاف، راست) به طور کلی با چارچوب هدف مطابقت ندارد. با این حال، همانطور که در بخش ؟؟ اثبات شده است، دو چارچوب توسط تبدیل گیج القا شده توسط ایزومتری مرتبط هستند:

$$\big[\phi_*|_p(e_i^{\widetilde{A}})\big]_{i=1}^d \;=\; \big[e_i^A\big]_{i=1}^d \lhd g_\phi^{A\widetilde{A}}(p)\,, \tag{۴۳}$$

که در آن $g_\phi^{A\widetilde{A}}(p)$ عنصر گروهی از معادله (۳۹) و $\lhd$ عمل راست از معادله (۱۱) است. اصطلاح «تبدیل گیج القا شده توسط ایزومتری» تا آنجا معنا دارد که هندسه‌های اطراف p و $\phi(p)$ غیرقابل تشخیص هستند زیرا ϕ یک ایزومتری، یعنی یک تقارن از M است. با شناسایی دو نقطه با یکدیگر،

[۲۰] به طور رسمی‌تری بیان شود، چنین ایزومتری‌هایی اتومورفیسم‌های بندل اصلی G-ساختار هستند (یا القا می‌کنند).

بنابراین می‌توان عمل فعال ϕ بر یک کمیت هندسی را به عنوان یک تبدیل گیج منفعل، یعنی یک تغییر القا شده از چارچوب منبع به چارچوب هدف، تفسیر مجدد کرد.

قضیه ؟؟ در بخش ؟؟ اثبات می‌کند که ایزومتری‌های حفظ‌کننده G–ساختار در Isom_{GM} و تبدیل‌های گیج القا شده با مقدار G یکدیگر را نتیجه می‌دهند، یعنی

$$\phi \in \mathrm{Isom}_{GM} \quad \Longleftrightarrow \quad g_\phi^{A\widetilde{A}}(p) \in G \ \ \forall\, p \in M \tag{۴۴}$$

برای گیج‌های دلخواه $\psi_p^{\widetilde{A}}$ و $\psi_{\phi(p)}^{A}$ از G–اطلس برقرار است. خواننده باید این ادعاها را در مثال‌های ما در شکل ؟؟ تأیید کند.

۳.۴.۳ پیش‌برنده بردارهای ویژگی

اگر (و تنها اگر) یک ایزومتری تقارنی از G–ساختار باشد، منجر به پیش‌برنده بردارهای ویژگی می‌شود. به طور شهودی، این پیش‌برنده بردارهای ویژگی را از نقاط p به $\phi(p)$ منتقل می‌کند. هنگامی که نسبت به دو چارچوب مرجع در p و $\phi(p)$ بیان شود، توسط تبدیل گیج القا شده

$$\rho\big(g_\phi^{A\widetilde{A}}(p)\big)\,. \tag{۴۵}$$

داده می‌شود. توجه داشته باشید که این تبدیل برای هر $\phi \in \mathrm{Isom}_{GM}$ به خوبی تعریف شده است، زیرا تبدیل‌های گیج القا شده $g_\phi^{A\widetilde{A}}(p)$ در این حالت مقادیری در G خواهند گرفت و ρ یک G–نمایش است. در مقابل، اگر ϕ تقارنی از G–ساختار نباشد، تعریف پیش‌برنده بردار ویژگی متناظر غیرممکن است. این گزاره با این واقعیت مرتبط است که ویژگی‌های های□□□ معمولی هیچ رفتار تبدیل مشخصی تحت چرخش‌ها یا بازتاب‌ها در گروه اقلیدسی $\mathrm{E}(d)$ ندارند.

پیش‌برنده بردارهای ویژگی منفرد عملی بر کل میدان ویژگی f را دلالت می‌کند، که ما آن را با $\phi \rhd f$ نشان می‌دهیم. نسبت به مختصات، این عمل به صورت

$$\big[\phi \rhd f\big]^A(\phi(p)) \;=\; \rho\big(g_\phi^{A\widetilde{A}}(p)\big)\, f^{\widetilde{A}}(p)\,. \tag{۴۶}$$

بیان می‌شود. ما بعداً اثبات خواهیم کرد که های□□□ مستقل از مختصات نسبت به عمل ایزومتری‌ها در Isom_{GM} بر میدان‌های ویژگی تناوب‌پذیر هستند؛ به شکل ۱۹ مراجعه کنید. این ویژگی بر این واقعیت تکیه دارد که عمل فعال ایزومتری بر میدان‌های ویژگی می‌تواند توسط معادله (۴۶) به عنوان یک تبدیل گیج منفعل صرف ضرایب بردار ویژگی درک شود.

۴ شبکه‌های مستقل از مختصات و □□–کانولوشن‌ها

شبکه‌های عصبی داده‌ها را با اعمال مجموعه‌ای از نگاشت‌های پارامتری (لایه‌ها) به یک سیگنال ورودی پردازش می‌کنند □ در مورد ما به مجموعه‌ای از میدان‌های ویژگی بر روی یک منیفولد ریمانی. اصل کوواریانس بدین ترتیب ایجاب می‌کند که لایه‌های منفرد شبکه باید عملیات مستقل از مختصات باشند. نمایش‌های مختصاتی چنین لایه‌هایی بنابراین باید طوری تبدیل شوند که قوانین تبدیل میدان‌های ویژگی ورودی و خروجی خود را رعایت GM کنند. به جز این الزام سازگاری، لایه‌های مستقل از مختصات عمومی بدون محدودیت باقی می‌مانند.

یک اصل طراحی رایج شبکه‌های عصبی که بر روی سیگنال‌های فضایی (میدان‌های ویژگی) عمل می‌کنند این است که آن‌ها به معنای تعمیم یافته‌ای کانولوشنی هستند. ویژگی اصلی که اکثر تعمیم‌های عملیات کانولوشن در آن مشترک هستند این است که استنتاج آن‌ها مستقل از موقعیت است. این امر با اشتراک توابع قالب، به عنوان مثال کرنل‌های کانولوشن یا بایاس‌ها، بین مکان‌های مختلف حاصل می‌شود. هر زمان که گروه ساختار G غیربدیهی باشد، فرآیند اشتراک وزن مبهم است زیرا توابع قالب می‌توانند نسبت به چارچوب‌های مرجع مختلف به اشتراک گذاشته شوند. همانطور که در ادامه استدلال خواهیم کرد، این ابهام با طراحی توابع قالب مشترک برای تناوب‌پذیر بودن تحت تبدیل‌های گیج با مقدار G حل می‌شود. توابع قالب تناوب‌پذیر گیج نسبت به چارچوب مرجع خاصی که در آن اعمال می‌شوند بی‌تفاوت خواهند بود و بنابراین امکان اشتراک وزن مستقل از مختصات را فراهم می‌کنند.

در این بخش ما لایه‌های شبکه‌ای را در نظر خواهیم گرفت که میدان‌های $f_{□□}$ از نوع $\rho_{□□}$ را به عنوان ورودی می‌گیرند و میدان $f_{□□□}$ از نوع $\rho_{□□□}$ را به عنوان خروجی تولید می‌کنند. بخش ۱.۴ حالت خاص لایه‌هایی را که به صورت نقطه‌ای عمل می‌کنند مورد بحث قرار می‌دهد، یعنی آن‌هایی که خروجی آن‌ها $f_{□□□}(p)$ در هر $p \in M$ تنها به بردار ویژگی ورودی منفرد $f_{□□}(p)$ در همان مکان بستگی دارد. مثال‌های عملی مرتبط که در اینجا در نظر گرفته شده‌اند عبارتند از ۱×۱□□□□□□□□□□□– تناوب‌پذیر گیج در بخش ۱.۱.۴، جمع بایاس در بخش ۲.۱.۴ و غیرخطی‌ها در بخش ۳.۱.۴. حالت پیچیده‌تر کانولوشن‌ها با کرنل‌های گسترده فضایی در بخش ۲.۴ بررسی شده است. به عنوان آماده‌سازی، بخش ۱.۲.۴ میدان‌های ویژگی را همانطور که از دیدگاه مشاهده‌گران محلی (چارچوب‌های مرجع) دیده می‌شوند مورد بحث قرار می‌دهد، که کرنل‌های (کانولوشن) نسبت به آن‌ها اعمال خواهند شد. چنین مشاهداتی به عنوان پس‌کشی میدان ویژگی به فضای مماس یک مشاهده‌گر رسمی‌سازی می‌شوند؛ به شکل ۱۶ مراجعه کنید. بخش ۲.۲.۴ به اصطلاح تبدیل‌های میدان کرنل را معرفی می‌کند، که شبیه کانولوشن‌ها هستند اما اشتراک وزن فضایی را فرض نمی‌کنند و بنابراین توسط یک میدان

کرنل (به صورت هموار متغیر) بر روی M پارامتری می‌شوند. GM-کانولوشن‌های واقعی در بخش ۳.۲.۴ به عنوان آن تبدیل‌های میدان کرنل تعریف می‌شوند که توسط یک کرنل قالب منفرد مشترک پارامتری می‌شوند. به منظور تضمین استقلال از مختصات فرآیند اشتراک وزن، کرنل‌های کانولوشن مطالبه می‌شوند که G-استیریبل باشند، یعنی قید تناوب‌پذیری گیج را برآورده کنند. بخش ۳.۴ نشان می‌دهد که GM-کانولوشن‌ها به طور خودکار تحت آن ایزومتری‌هایی که تقارن‌های G-ساختار هستند (Isom_{GM}-تناوب‌پذیر) تناوب‌پذیر هستند. این بدان معنی است که GM-کانولوشن‌ها با عمل ایزومتری‌ها بر میدان‌های ویژگی همانطور که در شکل ۱۹ بصری‌سازی شده است، جابجا می‌شوند.

۱.۴ عملیات نقطه‌ای تناوب‌پذیر گیج

برای شروع، ما برخی عملیات شبکه عصبی را در نظر می‌گیریم که محدودیت‌های ناشی از استقلال از مختصات مورد نیاز و اشتراک وزن به ویژه ساده قابل استخراج هستند. همه این عملیات این ویژگی مشترک را دارند که به صورت نقطه‌ای بر بردارهای ویژگی عمل می‌کنند، یعنی بردارهای ویژگی خروجی $f_{\square\square\square}(p)$ در $p \in M$ را صرفاً بر اساس بردارهای ویژگی ورودی $f_{\square\square}(p)$ در همان مکان محاسبه می‌کنند. به منظور ارضای اصل کوواریانس، مختصات‌بندی‌های این عملیات همگی مطالبه می‌شوند که مطابق پیش‌ترکیب با $\rho_{\square\square}$ و پس‌ترکیب با $\rho_{\square\square\square}$ تبدیل شوند. هنگام تقاضای اینکه عملیات بر حسب وزن‌های مشترک تعیین شوند، این قوانین تبدیل الزامی برای تناوب‌پذیری گیج (یا ناوردایی) عملیات را دلالت می‌کنند.

استخراج‌ها برای عملیات نقطه‌ای مختلف در بخش‌های بعدی ۱.۱.۴، ۲.۱.۴ و ۳.۱.۴ در گام‌های اول عمدتاً مشابه هستند و به محدودیت‌های کوواریانس و تناوب‌پذیری اساساً یکسانی بر توابع قالب منجر می‌شوند. بنابراین آن‌ها می‌توانند با هم بررسی شوند، عملیات خاص (یا تابع قالب) را انتزاعی نگه داشته. با این حال، از آنجا که مفاهیم محدودیت‌های حاصل برای نمونه‌سازی‌های خاص متفاوت است، و از آنجا که می‌خواهیم بحث را نزدیک به کاربرد نگه داریم، چنین فرمول‌بندی انتزاعی را حذف خواهیم کرد و مستقیماً نمونه‌سازی‌های خاص را در نظر خواهیم گرفت.

۱.۱.۴ تناوب‌پذیر گیج1

—

به عنوان اولین مثال از عملیات نقطه‌ای، ما عمل خانواده‌ای از نگاشت‌های خطی $\mathcal{C}_p$ را در نظر می‌گیریم، که بردار ویژگی ورودی $f_{\square\square}(p)$ را در هر $p \in M$ به بردار ویژگی خروجی

$$f_{\square\square\square}(p) := \mathcal{C}_p\, f_{\square\square}(p)\,. \tag{۴۷}$$

ارسال می‌کنند. اگر فرض اشتراک وزن فضایی را اضافه کنیم، نگاشت‌های خطی $\mathcal{C}_p$ و $\mathcal{C}_q$ در مکان‌های مختلف p و q جفت خواهند شد، و عملیات را می‌توان به عنوان کانولوشنی با کرنل دلتای دیراک با مقدار عملگر خطی دید. این عملیات در بینایی کامپیوتر بسیار رایج است، جایی که معمولاً به عنوان ۱×۱□□□□□□□□□□□□-نشان داده می‌شود، زیرا گسسته‌سازی فضایی یک کرنل دیراک خطی که بر روی تصاویر دو بعدی عمل می‌کند توسط کرنل (با مقدار ماتریس) با گستره فضایی 1×1 پیکسل داده می‌شود. ما در ادامه استخراج خواهیم کرد که تقاضای اشتراک وزن فضایی منجر به محدودیتی خواهد شد که کرنل‌های قالب با مقدار ماتریس را مجبور می‌کند که درهم‌تنیده باشند، یعنی ماتریس‌های تناوب‌پذیر گیج.

پیش از فرض اشتراک وزن، بیان‌های مختصاتی نگاشت‌های خطی $\mathcal{C}_p$ و تبدیل‌های گیج بین آن‌ها بسیار شبیه آن‌هایی از نگاشت‌های خطی بر روی T_pM رفتار می‌کنند، که در بخش ۲.۱.۳ مورد بحث قرار دادیم. از آنجا که بردارهای ویژگی ورودی و خروجی در مختصات توسط بردارهای ضریب $f_{\square\square}^A(p) \in \mathbb{R}^{c_{\square\square}}$ و $f_{\square\square\square}^A(p) \in \mathbb{R}^{c_{\square\square\square}}$ نمایش داده می‌شوند، نگاشت خطی به طور طبیعی توسط آن ماتریس $\mathcal{C}_p^A \in \mathbb{R}^{c_{\square\square\square} \times c_{\square\square}}$ نمایش داده می‌شود که

$$f_{\square\square\square}^A(p) = \mathcal{C}_p^A \cdot f_{\square\square}^A(p)\,. \tag{۴۸}$$

را برآورده می‌کند. این رابطه البته برای مختصات‌بندی‌های دلخواه برقرار است، به طوری که برای هر گیج دیگری که با B برچسب گذاری شده است، $f_{\square\square\square}^B(p) = \mathcal{C}_p^B \cdot f_{\square\square}^B(p)$ داریم. قانون تبدیلی که $\mathcal{C}_p^B$ را به $\mathcal{C}_p^A$ مرتبط می‌کند از اصل کوواریانس از قوانین تبدیل ویژگی‌های ورودی و خروجی پیروی می‌کند. از آنجا که این‌ها توسط $f_{\square\square\square}^B(p) = \rho_{\square\square\square}\big(g_p^{BA}\big) f_{\square\square\square}^A(p)$ و $f_{\square\square}^B(p) = \rho_{\square\square}\big(g_p^{BA}\big) f_{\square\square}^A(p)$ داده می‌شوند، داریم:

$$\begin{aligned}
& & f_{\square\square\square}^B(p) \;&=\; \mathcal{C}_p^B \cdot f_{\square\square}^B(p) \\
\Leftrightarrow & & \rho_{\square\square\square}\big(g_p^{BA}\big)\, f_{\square\square\square}^A(p) \;&=\; \mathcal{C}_p^B\, \rho_{\square\square}\big(g_p^{BA}\big)\, f_{\square\square}^A(p) \\
\Leftrightarrow & & f_{\square\square\square}^A(p) \;&=\; \rho_{\square\square\square}\big(g_p^{BA}\big)^{-1}\, \mathcal{C}_p^B\, \rho_{\square\square}\big(g_p^{BA}\big)\, f_{\square\square}^A(p)\,.
\end{aligned} \tag{۴۹}$$

مقایسه با معادله (۴۸) دلالت دارد بر اینکه دو بیان مختصاتی $\mathcal{C}_p$ لزوماً با

$$\mathcal{C}_p^B \;=\; \rho_{\square\square\square}\big(g_p^{BA}\big)\, \mathcal{C}_p^A\, \rho_{\square\square}\big(g_p^{BA}\big)^{-1} \tag{۵۰}$$

مرتبط هستند اگر قرار باشد قوانین تبدیل بردارهای ویژگی را رعایت کنند. همانطور که معمول است، این ملاحظات به طور مختصر توسط نمودار جابجایی ضبط می‌شوند:

$$\begin{tikzcd} \mathbb{R}^{c_{□□}} \arrow[r, "\mathcal{C}_p^A \cdot"] \arrow[d, "\rho_{□□}(g_p^{BA})\cdot"'] & \mathbb{R}^{c_{□□□}} \arrow[d, "\rho_{□□□}(g_p^{BA})\cdot"] \\ \mathbb{R}^{c_{□□}} \arrow[r, "\mathcal{C}_p^B \cdot"'] & \mathbb{R}^{c_{□□□}} \end{tikzcd} \tag{۵۱}$$

مفهوم عملی مهم این نتیجه تا کنون این است که نگاشت خطی $\mathcal{C}_p$ به هیچ وجه محدود نیست. به بیان دیگر: تا زمانی که بیان‌های مختصاتی در گیج‌های مختلف توسط معادله (۵۰) مرتبط باشند، آزاد هستیم که $\mathcal{C}_p$ را در یک گیج دلخواه ثابت A توسط ماتریس بدون محدودیت $\mathcal{C}_p^A$ پارامتری کنیم. همانطور که خواهیم دید، وضعیت زمانی تغییر می‌کند که نگاشت‌های خطی اشتراک وزن داشته باشند.

حال حالتی را در نظر بگیرید که نگاشت‌های خطی $\mathcal{C}_p$ و $\mathcal{C}_q$ وزن‌ها را به اشتراک می‌گذارند. این بدان معنی است که فرض می‌کنیم آن‌ها توسط مجموعه مشترکی از پارامترها پارامتری شوند، که توسط کرنل قالب ۱×۱□□□□□□□□□□□– $K_{۱\times۱} \in \mathbb{R}^{c_{□□□}\times c_{□□}}$ داده می‌شود. سؤال باز این است که دقیقاً چگونه نگاشت‌های آزاد از مختصات باید بر حسب این کرنل قالب پارامتری شوند. الزام ما برای استقلال از مختصات GM تقاضا می‌کند که هیچ چارچوب مرجع خاصی را در فرآیند اشتراک وزن ترجیح ندهیم، یعنی همه مختصات‌بندی‌ها را به همان شیوه درمان کنیم. بنابراین لازم است که کرنل قالب را با همه مختصات‌بندی‌ها به طور همزمان به اشتراک بگذاریم، یعنی

$$\mathcal{C}_p^X = K_{۱\times۱} \quad \text{برای هر گیج } \left(U^X, \psi^X\right) \in \mathscr{A}^G \text{ با } p \in U^X \,, \tag{۵۲}$$

تنظیم کنیم، که در آن $\mathscr{A}^G$ (حداکثر) G–اطلس متناظر با G–ساختار در نظر گرفته شده است؛ به معادله (۲۵) مراجعه کنید. از آنجا که محدودیت کوواریانس در معادله (۵۰) نیاز دارد که برای گیج‌های دلخواه مرتبط با G برقرار باشد، و مختصات‌بندی‌های $\mathcal{C}_p^A = \mathcal{C}_p^B = K_{۱\times۱}$ نگاشت‌های خطی همگی مطابقت دارند، تقاضای مشترک برای اشتراک وزن و استقلال از مختصات GM منجر به محدودیت

$$K_{۱\times۱} \;=\; \rho_{□□□}(g)\, K_{۱\times۱}\, \rho_{□□}(g)^{-۱} \qquad \forall\, g \in G \tag{۵۳}$$

بر کرنل قالب می‌شود. تطبیق متناظر نمودار جابجایی در معادله (۵۱) با اشتراک وزن توسط:

$$\begin{tikzcd} \mathbb{R}^{c_{□□}} \arrow[r, "K_{۱\times۱}\cdot"] \arrow[d, "\rho_{□□}(g_p^{BA})\cdot"'] & \mathbb{R}^{c_{□□□}} \arrow[d, "\rho_{□□□}(g_p^{BA})\cdot"] \\ \mathbb{R}^{c_{□□}} \arrow[r, "K_{۱\times۱}\cdot"'] & \mathbb{R}^{c_{□□□}} \end{tikzcd} \tag{۵۴}$$

داده می‌شود.

نتیجه‌گیری این تحلیل این است که کرنل‌های قالبی که می‌توانند بدون ابهام به اشتراک گذاشته شوند دقیقاً آن‌هایی هستند که تحت عمل گیج ناوردا هستند. فضای برداری چنین کرنل‌های ۱×۱□□□□□□□□□□□– ناوردا گیج ساده فضای نگاشت‌های درهم‌تنیده بین نمایش‌های $\rho_{□□}$ و $\rho_{□□□}$ است، یعنی

$$\mathrm{Hom}_G(\rho_{□□}, \rho_{□□□}) \;:=\; \left\{ K_{۱\times۱} \in \mathbb{R}^{c_{□□□}\times c_{□□}} \;\middle|\; K_{۱\times۱} = \rho_{□□□}(g)\, K_{۱\times۱}\, \rho_{□□}(g)^{-۱} \;\; \forall g \in G \right\} \;\subseteq\; \mathbb{R}^{c_{□□□}\times c_{□□}} \,. \tag{۵۵}$$

توجه داشته باشید که، طبق لم شور [؟]، الزام بر $K_{۱\times۱}$ برای درهم‌تنیده بودن مانع از نگاشت بین میدان‌هایی که تحت نمایش‌های کاهش‌ناپذیر غیرهم‌ریخت تبدیل می‌شوند از طریق ۱×۱□□□□□□□□□□□– می‌شود. این محدودیت شدید با کرنل‌های ۱×۱□□□□□□□□□□□– اجتناب‌ناپذیر است اما بعداً هنگام اجازه دادن به کرنل‌های گسترده فضایی حل خواهد شد.

در این نقطه می‌خواهیم اشاره کنیم که اصطلاحات «تابع قالب تناوب‌پذیر گیج» و «تابع قالب ناوردا گیج» را به طور متقابل استفاده می‌کنیم. این با مشاهده توجیه می‌شود که محدودیت ناوردایی در معادله (۵۳) می‌تواند به عنوان محدودیت تناوب‌پذیری $K_{۱\times۱}\, \rho_{□□}(g) = \rho_{□□□}(g)\, K_{۱\times۱} \;\; \forall g \in G$ نوشته شود. به طور کلی امکان مشاهده توابعی که نسبت به عمل گروهی در دامنه و دامنه مشترک آن‌ها تناوب‌پذیر هستند به عنوان ناوردای عمل متناظر بر خود تابع وجود دارد. در کاربرد ما، دیدگاه تناوب‌پذیری برجسته می‌کند که تبدیل میدان ورودی منجر به تبدیل متناظر میدان خروجی خواهد شد، که تضمین می‌کند همه کمیت‌های درگیر به طور کوواریانت با یکدیگر تبدیل شوند. از طرف دیگر، دیدگاه ناوردایی تأکید می‌کند که تابع قالب می‌تواند در گیج دلخواه به اشتراک گذاشته شود.

۲.۱.۴ جمع بایاس تناوب‌پذیر گیج

پس از اعمال عملیات کانولوشن، رایج است که بردار بایاس (مشترک) به بردارهای ویژگی منفرد جمع شود. همراه با الزام استقلال از مختصات، اشتراک وزن دوباره منجر به محدودیت خطی خواهد شد. این محدودیت تنها اجازه جمع شدن بایاس‌ها به زیرفضاهای ناوردا عمل گیج بر میدان ویژگی ورودی را خواهد داد.

همانند قبل، ابتدا جمع بایاس را بدون اشتراک وزن در نظر می‌گیریم. بنابراین بایاس‌های $\mathscr{b}_p$ داریم، که به موقعیت p بر روی منیفولد بستگی دارند، که به بردار ویژگی ورودی جمع می‌شوند تا بردار ویژگی خروجی

$$f_{\square\square\square}(p) = f_{\square\square}(p) + \mathscr{b}_p\,. \tag{۵۶}$$

تولید کنند. نسبت به گیج‌های ψ_p^A و ψ_p^B، بایاس توسط آن بردارهای ضریب $\mathscr{b}_p^A$ و $\mathscr{b}_p^B$ در $\mathbb{R}^c$ نمایش داده می‌شود که $f_{\square\square\square}^A(p) = f_{\square\square}^A(p) + \mathscr{b}_p^A$ و $f_{\square\square\square}^B(p) = f_{\square\square}^B(p) + \mathscr{b}_p^B$ را برآورده می‌کنند. از آنجا که جمع بردارها اجازه تغییر قوانین تبدیل آن‌ها را نمی‌دهد، نمایش‌های گروهی مرتبط با ویژگی‌های ورودی و خروجی لزوماً مطابقت دارند، یعنی

$$\rho_{\square\square} = \rho_{\square\square\square} =: \rho\,. \tag{۵۷}$$

همراه با الزام برای استقلال از مختصات، این دلالت دارد بر اینکه نمودار

$$\begin{tikzcd}
\mathbb{R}^c \arrow[r, "+\mathscr{b}_p^A"] \arrow[d, "\rho(g_p^{BA})\cdot"'] & \mathbb{R}^c \arrow[d, "\rho(g_p^{BA})\cdot"] \\
\mathbb{R}^c \arrow[r, "+\mathscr{b}_p^B"'] & \mathbb{R}^c
\end{tikzcd}, \tag{۵۸}$$

که معادل آن در معادله (۵۱) است، نیاز دارد که جابجا شود. به صورت معادله نوشته شده، این رابطه $\rho\big(g_p^{BA}\big)\big(f_p^A + \mathscr{b}_p^A\big) = \rho\big(g_p^{BA}\big) f_p^A + \mathscr{b}_p^B$ را تقاضا می‌کند که برقرار باشد. از آنجا که خطی بودن $\rho(g)$ امکان بازنویسی سمت راست را به عنوان $\rho\big(g_p^{BA}\big) f_p^A + \rho\big(g_p^{BA}\big)\mathscr{b}_p^A$ می‌دهد، تفریق بردار ویژگی ورودی منجر به

$$\mathscr{b}_p^B \;=\; \rho\big(g_p^{BA}\big)\,\mathscr{b}_p^A\,. \tag{۵۹}$$

می‌شود. بردارهای ضریبی که یک بایاس مستقل از مختصات را نسبت به گیج‌های مختلف نمایش می‌دهند بنابراین نیاز دارند که دقیقاً مانند بردارهای ویژگی که به آن‌ها جمع می‌شوند تبدیل شوند. همانند مورد ۱×۱□□□□□□□□□□□□–، استقلال از مختصات بایاس $\mathscr{b}_p$ را به هیچ وجه محدود نمی‌کند، بلکه فقط مختصات‌بندی‌های مختلف همان بایاس را مطالبه می‌کند که با یکدیگر سازگار باشند. یک پیاده‌سازی بنابراین می‌تواند گیج دلخواه انتخاب کند و بایاس را در آن گیج آزادانه توسط پارامترهایی در $\mathbb{R}^{c_{\square\square}}$ پارامتری کند.

وضعیت دوباره زمانی تغییر می‌کند که اشتراک وزن فضایی تقاضا شود. فرض کنید $b \in \mathbb{R}^{c_{\square\square}}$ بردار بایاس قالبی باشد که بر روی منیفولد به اشتراک گذاشته شود. از آنجا که تنها راه انجام این کار بدون ترجیح دلخواه هر مختصات‌بندی، اشتراک بردار بایاس در همه گیج‌ها به طور همزمان است، باید

$$\mathscr{b}_p^X = b \quad \text{برای هر گیج} \;\big(U^X, \psi^X\big) \in \mathscr{A}^G \;\text{ با }\; p \in U^X\,. \tag{۶۰}$$

را در قیاس با معادله (۵۲) مطالبه کنیم. ترکیب محدودیت کوواریانس در معادله (۵۹) با این اشتراک وزن مستقل از گیج سپس منجر به محدودیت ناوردایی

$$b \;=\; \rho(g)\,b \qquad \forall\, g \in G \tag{۶۱}$$

بر قالب بردار بایاس می‌شود. برای تکمیل قیاس با مورد ۱×۱□□□□□□□□□□□□–، نسخه تطبیق یافته نمودار جابجایی در معادله (۵۸) با وزن‌های مشترک را نشان می‌دهیم:

$$\begin{tikzcd}
\mathbb{R}^c \arrow[r, "+b"] \arrow[d, "\rho(g_p^{BA})\cdot"'] & \mathbb{R}^c \arrow[d, "\rho(g_p^{BA})\cdot"] \\
\mathbb{R}^c \arrow[r, "+b"'] & \mathbb{R}^c
\end{tikzcd} \tag{۶۲}$$

برای بصیرت در مفاهیم محدودیت ناوردایی در معادله (۶۱)، فرض کنید برای بردار قالب داده شده b برآورده شود. به دلیل خطی بودن محدودیت، هر بردار مقیاس شده $\alpha\cdot b$ برای $\alpha \in \mathbb{R}$ نیز آن را برآورده خواهد کرد، یعنی هر حل زیرفضای یک بعدی از $\mathbb{R}^c$ را پوشش می‌دهد که تحت عمل ρ

ناوردا است. چنین زیرفضای ناوردا به عنوان زیرنمایش ρ نشان داده می‌شود. از آنجا که زیرفضاهای در نظر گرفته شده یک بعدی هستند، خودشان هیچ زیرفضای مناسب ندارند و بنابراین زیرنمایش‌های کاهش‌ناپذیر بدیهی هستند. نتیجه می‌شود که فضای برداری

$$\mathcal{B}^G_\rho \;:=\; \left\{ b \in \mathbb{R}^c \;\middle|\; b = \rho(g)\, b \;\; \forall g \in G \right\} \tag{۶۳}$$

بایاس‌های تناوب‌پذیر گیج با (زیرفضاهای) زیرنمایش‌های بدیهی ρ مطابقت دارد. ابعاد $\mathcal{B}^G_\rho$ □ و بنابراین تعداد پارامترهای قابل یادگیری □ با کثرت زیرنمایش‌های بدیهی موجود در ρ مطابقت دارد. برای گروه‌های فشرده G، روابط متعامد شور دلالت دارند بر اینکه این ابعاد توسط $\dim\big(\mathcal{B}^G_\rho\big) = \int_G \operatorname{tr}\big(\rho(g)\big)dg$ داده می‌شود. این گزاره موارد مهم عملی گروه‌های متعامد $G = d$ و همه زیرگروه‌های آن را پوشش می‌دهد.

دو مثال ساده از میدان‌های ویژگی که ممکن است بخواهیم بایاس‌های مشترک به آن‌ها جمع کنیم، میدان‌های اسکالر و میدان‌های بردار مماس هستند. بر اساس تعریف، میدان ضریب یک میدان اسکالر تحت تبدیل‌های گیج ناوردا است، یعنی طبق نمایش بدیهی $\rho(g) = ۱ \;\; \forall g \in G$ تبدیل می‌شود. بنابراین می‌توان بایاس (اسکالر) $b \in \mathbb{R}$ به آن‌ها جمع کرد. در مقابل، میدان ضریب یک میدان بردار مماس طبق نمایش گروهی غیربدیهی، کاهش‌ناپذیر $\rho(g) = g$ تبدیل می‌شود. از آنجا که این نمایش هیچ زیرنمایش بدیهی ندارد، جمع بردار بایاس مشترک به میدان‌های بردار مماس با حفظ استقلال از مختصات غیرممکن است. به عنوان مثال سوم، نمایش‌های منظم گروه‌های فشرده را در نظر بگیرید، که به عنوان مثال میدان‌های ویژگی شبکه‌های کانولوشنی گروهی را توصیف می‌کنند. توسط قضیه پیتر–ویل، شناخته شده است که نمایش‌های منظم دقیقاً یک زیرنمایش بدیهی دارند [؟ ؟]. بنابراین بایاس که به میدان‌های ویژگی منظم جمع می‌شود توسط یک پارامتر منفرد توصیف می‌شود.

۳.۱.۴ غیرخطی‌های تناوب‌پذیر گیج

به جز عملیات خطی (کانولوشن) و جمع بایاس، اساسی‌ترین عملیات استفاده شده در هر شبکه عصبی غیرخطی‌ها هستند. ما در اینجا حالت معمول غیرخطی‌های σ_p را در نظر خواهیم گرفت که به روشی محلی فضایی عمل می‌کنند، یعنی بردارهای ویژگی خروجی را به عنوان $f_{\square\square\square}(p) = \sigma_p\big(f_{\square\square}(p)\big)$ محاسبه می‌کنند. یک غیرخطی مشترک دوباره مطالبه می‌شود که تناوب‌پذیر گیج باشد. از آنجا که استدلالی که به این نتیجه‌گیری منجر می‌شود شبیه آن در موارد قبلی است، تنها به طور خلاصه آن را خلاصه خواهیم کرد. به دلیل عمومیت نگاشت‌های غیرخطی استخراج فضاهای حل خطی همانند معادلات (۵۵) و (۶۳) غیرممکن است، با این حال، برخی مثال‌های خاص را مورد بحث قرار خواهیم داد.

مشابه قبل، هر غیرخطی آزاد از مختصات σ_p نسبت به گیج‌های A و B توسط بیان‌های مختصاتی $\sigma^A_p : \mathbb{R}^{c_{\square\square}} \to \mathbb{R}^{c_{\square\square\square}}$ و $\sigma^B_p : \mathbb{R}^{c_{\square\square}} \to \mathbb{R}^{c_{\square\square\square}}$ داده می‌شود، که توسط تقاضای استقلال از مختصات مطالبه می‌شوند که با $\sigma^B_p = \rho_{\square\square\square}\big(g^{BA}_p\big) \circ \sigma^A_p \circ \rho_{\square\square}\big(g^{BA}_p\big)^{-۱}$ مرتبط باشند. یک تابع قالب غیرخطی $\mathscr{s} : \mathbb{R}^{c_{\square\square}} \to \mathbb{R}^{c_{\square\square\square}}$ تنها زمانی می‌تواند به روشی مستقل از مختصات به اشتراک گذاشته شود که به طور همزمان با همه گیج‌ها به اشتراک گذاشته شود. این محدودیت کوواریانس را به محدودیت ناوردایی $\mathscr{s} = \rho_{\square\square\square}(g) \circ \mathscr{s} \circ \rho_{\square\square}(g)^{-۱} \;\; \forall g \in G$ بر تابع قالب تبدیل می‌کند، یا، به طور معادل، به محدودیت تناوب‌پذیری متناظر

$$\rho_{\square\square\square}(g) \circ \mathscr{s} = \mathscr{s} \circ \rho_{\square\square}(g)^{-۱} \qquad \forall\, g \in G\,. \tag{۶۴}$$

به دلیل غیرخطی بودن این محدودیت مجبور هستیم آن را مورد به مورد بررسی کنیم □ بنابراین بحث خود را به چند مثال خاص محدود خواهیم کرد. بحث‌برانگیزترین ساده‌ترین حالت، غیرخطی‌هایی است که بین میدان‌های اسکالر نگاشت می‌کنند، یعنی برای آن‌ها $\rho_{\square\square}(g) = \rho_{\square\square\square}(g) = ۱$ برای هر $g \in G$ ناوردا هستند. در این حالت محدودیت تناوب‌پذیری در معادله (۶۴) به $\mathscr{s} = \mathscr{s}$ تبدیل می‌شود، که به طور بدیهی برای هر غیرخطی $\mathscr{s} : \mathbb{R} \to \mathbb{R}$ برآورده می‌شود. مثال جالب‌تری نمایش یکانی $\rho_{\square\square}$ است. یک غیرخطی ممکن برای این حالت توسط نرم بردارهای ویژگی داده می‌شود. از آنجا که $\|\rho_{\square\square}(g) f^A_{\square\square}(p)\| = \|f^A_{\square\square}(p)\|$ به دلیل یکانی بودن $\rho_{\square\square}$ ناوردا است، $\rho_{\square\square\square}$ نمایش بدیهی خواهد بود. گرفتن نرم بنابراین به عنوان عملیات غیرخطی، تناوب‌پذیر گیج دیده می‌شود که هر میدان یکانی را به میدان اسکالر نگاشت می‌کند. یک نگاشت غیرخطی که نوع میدان را حفظ می‌کند، یعنی $\rho_{\square\square} = \rho_{\square\square\square}$ را برآورده می‌کند، می‌تواند به طور مشابه به عنوان $f^A_{\square\square}(p) \mapsto \|f^A_{\square\square}(p)\| \cdot f^A_{\square\square}(p)$ تعریف شود. گزینه دیگر، که ممکن است در یادگیری تعاملات فیزیکی نقش داشته باشد و در [؟ ؟ ؟ ؟] مورد بررسی قرار گرفته است، غیرخطی‌های حاصل‌ضرب تانسوری هستند. با داشتن دو میدان $f^A_{\square\square,۱}(p)$ و $f^A_{\square\square,۲}(p)$، که به ترتیب طبق $\rho_{\square\square,۱}$ و $\rho_{\square\square,۲}$ تبدیل می‌شوند، چنین غیرخطی‌هایی ویژگی حاصل‌ضرب تانسوری $f^A_{\square\square\square} = f^A_{\square\square,۱}(p) \otimes f^A_{\square\square,۲}(p)$ را محاسبه می‌کنند، که به طور تناوب‌پذیر طبق نمایش حاصل‌ضرب تانسوری $\rho_{\square\square\square} = \rho_{\square\square,۱} \otimes \rho_{\square\square,۲}$ تبدیل می‌شود.

همه این مثال‌ها محدودیت تناوب‌پذیری گیج در معادله (۶۴) را برآورده می‌کنند. اینکه کدام غیرخطی خاص در عمل خوب کار می‌کند، با این حال، هنوز سؤال تحقیقاتی بازی است که نیازمند بررسی تجربی بسیار بیشتری قبل از پاسخ دادن است. اولین تلاش در این جهت در [؟] انجام شده است.

۲.۴ تبدیلات میدان کرنل و کانولوشن‌های □□

عملیات اصلی شبکه‌های کانولوشنی، عملگر کانولوشن است که الگوهای مشخصه ویژگی‌ها را از یک همسایگی محلی در اطراف هر نقطه $p \in M$ به صورت خطی در یک بردار ویژگی جدید $f_{\square\square\square}(p)$ جمع می‌کند. یک کرنل کانولوشن گسترده فضایی، جزئیات این انباشت را تعیین می‌کند. اصل

کوواریانس نیازمند استقلال از مختصات و در نتیجه یک قانون تبدیل خاص برای کرنل‌ها تحت تبدیلات پیمانه است. همانند مثال‌های قبلی، یک تقاضای اضافی برای اشتراک‌گذاری وزن منجر به الزامی برای کرنل الگو می‌شود تا نسبت به پیمانه هموردا (G-راهبر) باشد.

مطابق با بخش قبل، ما به وضوح بین الزامات استقلال از مختصات و اشتراک‌گذاری وزن تمایز قائل می‌شویم. بنابراین، بخش ۲.۲.۴ با بحث در مورد میدان‌های کرنل و قوانین تبدیل آن‌ها بدون الزام به اینکه کرنل‌ها در موقعیت‌های جداگانه به هم مرتبط باشند، شروع می‌شود. چنین میدان‌های کرنل نامحدودی منجر به تبدیلات میدان کرنل می‌شوند که تبدیلات انتگرالی هستند و می‌توان آن‌ها را به عنوان پیش‌درآمدهای کانولوشن‌ها در نظر گرفت. کانولوشن‌های واقعی GM که توسط یک کرنل الگوی مشترک و لزوماً همورداي پیمانه‌ای پارامتری می‌شوند، در بخش ۳.۲.۴ تعریف شده‌اند. به عنوان یک آماده‌سازی، در بخش بعدی ۱.۲.۴، نمایش‌های محلی میدان‌های ویژگی را بر روی فضاهای مماس توصیف خواهیم کرد، جایی که آن‌ها با کرنل‌های کانولوشن تطبیق داده می‌شوند.

۱.۲.۴ دیدگاه یک ناظر محلی به میدان‌های ویژگی

برخلاف فضاهای اقلیدسی یا فضاهای همگن کلی‌تر مانند کره، هندسه محلی یک خمینه ریمانی کلی از نقطه‌ای به نقطه دیگر متفاوت است. بنابراین، بلافاصله مشخص نیست که یک کرنل کانولوشن چگونه باید روی M تعریف شود و چگونه می‌توان آن را بین مکان‌های مختلف به اشتراک گذاشت. یک راه‌حل رایج این است که کرنل را طبق معمول روی یک فضای برداری مسطح و اقلیدسی $\mathbb{R}^d$ تعریف کرده و آن را به جای خود خمینه، روی فضاهای مماس به اشتراک بگذاریم؛ به بخش‌های ۲.۲.۴ و ۳.۲.۴ یا کارهای قبلی [؟ ؟ ؟ ؟ ؟ ؟ ؟ ؟ ؟] مراجعه کنید. سپس، کرنل را می‌توان از طریق نگاشت نمایی ریمانی به روی خمینه نگاشت. می‌توان آن را به گونه‌ای تصور کرد که توسط یک ناظر محلی اعمال می‌شود که ویژگی‌ها را در محیط اطراف خود نسبت به چارچوب مرجع محلی خود اندازه‌گیری می‌کند. در این بخش به طور خلاصه به چگونگی درک میدان‌های ویژگی از دیدگاه ناظران محلی مختلف خواهیم پرداخت. از نظر ریاضی، این موضوع به عنوان پس‌کش (□□□□□□□□) و انتقال موازی میدان ویژگی به فضاهای مماس فرمول‌بندی می‌شود؛ برای تجسم به شکل ۱۶ مراجعه کنید.

برای نگاشت بین فضاهای مماس و خمینه، ما نگاشت نمایی ریمانی (مرتبط با اتصال لوی-چیویتا) را در نظر می‌گیریم.[۲۱] با فرض اینکه خمینه برای سادگی کامل از نظر ژئودزیکی باشد[۲۲]، نگاشت نمایی در یک نقطه خاص $p \in M$ نگاشتی است به صورت

$$\exp_p : T_pM \to M \,. \tag{۶۵}$$

این نگاشت، بردارهای $v \in T_pM$ را با نقاطی $\exp_p(v) \in M$ یکی می‌داند که با دنبال کردن ژئودزیک گذرنده از p با سرعت اولیه v برای یک واحد زمان به دست می‌آیند. نگاشت نمایی با اینکه فواصل شعاعی را حفظ می‌کند، به طور کلی زوایا را تغییر می‌دهد و یک‌به‌یک نیست. به عنوان مثال، اگر خمینه یک کره باشد، نگاشت‌های نمایی فضای مماس مربوطه را بی‌نهایت بار دور آن می‌پیچند. با این حال، تضمین می‌شود که نگاشت نمایی یک دیفئومورفیسم محلی است اگر دامنه آن به فواصلی کوتاه‌تر از فاصله تا مکان بُرش (جایی که یک‌به‌یک بودن از دست می‌رود) محدود شود.

با داشتن نگاشت‌های نمایی، می‌توان میدان‌های ویژگی روی خمینه را به فضاهای مماس پس‌کش کرد. به طور خاص، اگر f یک میدان ویژگی روی M باشد، پس‌کش $\exp_p^* f := f \circ \exp_p$ به عنوان نگاشتی تعریف می‌شود که بردار ویژگی $f(\exp_p(v))$ از $\exp_p(v)$ را به $v \in T_pM$ نسبت می‌دهد. توجه داشته باشید که به دلیل عدم یک‌به‌یک بودن نگاشت نمایی، هر بردار ویژگی ممکن است به چندین بردار مماس v_1 و v_2 اختصاص یابد اگر $\exp_p(v_1) = \exp_p(v_2)$ باشد □ این پدیده تا حدودی شبیه به اثرات عدسی گرانشی در فیزیک است. در صورتی که نگاشت نمایی یک‌به‌یک باشد، یا زمانی که آن را به شعاع یک‌به‌یک بودن آن محدود کنیم، پس‌کش معادل بیان میدان‌های ویژگی در مختصات نرمال ژئودزیکی است [؟].

به یاد بیاورید که هدف از پس‌کش بردارهای ویژگی به فضاهای مماس این است که بتوان آن‌ها را توسط یک کرنل کانولوشن جمع‌آوری کرد. متأسفانه، این کار بلافاصله امکان‌پذیر نیست زیرا بردارهای ویژگی در مکان‌های مختلف در فضاهای برداری متفاوتی قرار دارند و نسبت به پیمانه‌های متفاوتی بیان می‌شوند.[۲۳] بنابراین لازم است تمام بردارهای ویژگی $[\exp_p^* f](v)$ در یک فضای برداری و نسبت به یک پیمانه یکسان بیان شوند. یک ایده طبیعی که توسط ؟] پیشنهاد شد، انجام این کار با انتقال موازی بردارهای ویژگی در امتداد ژئودزیک‌هایی است که نگاشت نمایی را از v به $\exp_p(v)$ تعریف می‌کنند.[۲۴] ما این پس‌کش f را با انتقال اضافی با $\mathrm{Exp}_p^* f$ نشان می‌دهیم تا بر ارتباط نزدیک آن با پس‌کش معمولی $\exp_p^* f$ به T_pM تأکید کنیم. شکل ۱۶ یک ایده بصری از این پس‌کش انتقال‌دهنده میدان‌های ویژگی به فضای مماس و نمایش‌های آن $\left[\,\mathrm{Exp}_p^* f\right]^A$ و $\left[\,\mathrm{Exp}_p^* f\right]^B$ روی $\mathbb{R}^d$ نسبت به مختصات‌دهی‌های مختلف ارائه می‌دهد.

ما $\mathrm{Exp}_p^* f$ را با تعریف آن بر حسب عبارت مختصاتی‌اش نسبت به یک انتخاب پیمانه فرمول‌بندی می‌کنیم. برای این منظور، فرض کنید ψ_p^A یک پیمانه در p باشد که ویژگی‌های منتقل‌شده نهایتاً نسبت به آن بیان می‌شوند و $\psi_{\exp_p(v)}^{\widetilde{A}}$ یک پیمانه دلخواه در $\exp_p(v)$ باشد که بردار ویژگی در آن

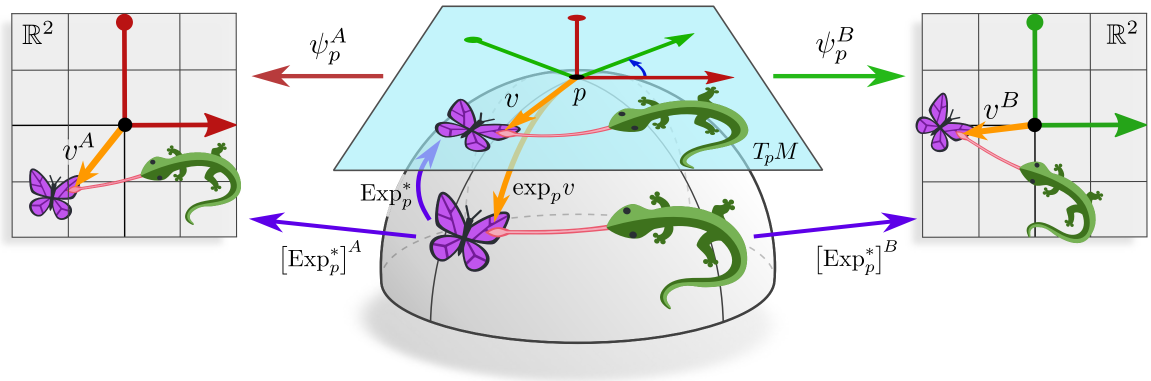
---

[۲۱] حتی مدل‌هایی که یک اتصال جایگزین (سازگار با G) برای انتقال ویژگی‌ها فرض می‌کنند، معمولاً از اتصال کانونی لوی-چیویتا برای محاسبه ژئودزیک‌ها و نگاشت‌های نمایی استفاده می‌کنند.

[۲۲] فرض کامل بودن ژئودزیکی M به این معناست که نگاشت‌های نمایی $\exp_p$ برای هر $p \in M$ روی کل فضای مماس T_pM تعریف شده‌اند. در مواردی که این فرض برقرار نباشد، می‌توان به پر کردن با صفر متوسل شد، که معمولاً در شبکه‌های کانولوشنی برای تصاویر با تکیه‌گاه متناهی استفاده می‌شود.

[۲۳] یک وضعیت بسیار مشابه، تعریف مشتقات کوواریانت را انگیزه‌مند می‌کند، که آن نیز نیاز به ترکیب اشیاء هندسی دارد که در فضاهای متفاوتی زندگی می‌کنند.

[۲۴] انتقال موازی در امتداد هر مسیر دیگری نیز به همان اندازه معتبر خواهد بود.



شکل ۱۶: یک میدان ویژگی f روی M و نمایش محلی آن $\mathrm{Exp}_p^* f$ روی T_pM از طریق پس‌کِش انتقال‌دهنده Exp_p^*. درست مانند پس‌کِش معمولی $\exp_p^* f$ از f در امتداد نگاشت نمایی $\exp_p : T_pM \to M$، پس‌کِش انتقال‌دهنده بردارهای ویژگی $f(\exp_p(v))$ را به بردارهای مماس $v \in T_pM$ نسبت می‌دهد. با این حال، از آنجایی که هدف ما انباشت ویژگی‌های پس‌کِش شده با استفاده از یک کرنل کانولوشن است، آن‌ها باید در یک فضای یکسان و نسبت به یک پیمانه یکسان در p داده شوند. بنابراین، پس‌کِش انتقال‌دهنده علاوه بر این، انتقال‌دهنده موازی (سازگار با G) را در امتداد ژئودزیک از $\exp_p(v)$ به p اعمال می‌کند. از طریق یک پیمانه ψ_p^X، پس‌کِش انتقال‌دهنده f روی T_pM را می‌توان روی $\mathbb{R}^d$ به صورت $[\mathrm{Exp}_p^* f]^X : \mathbb{R}^d \to \mathbb{R}^c$ بیان کرد □ انتخاب‌های مختلف چارچوب‌های مرجع (ناظران) در اینجا با تغییر شکل‌های خطی مختلف میدان ویژگی مطابقت دارد. تبدیلات میدان کرنل و کانولوشن‌های GM یک ویژگی خروجی $f_{\square\square\square}(p)$ را در p با تطبیق یک کرنل $\mathcal{K}_p$ روی T_pM با $\mathrm{Exp}_p^* f$ محاسبه می‌کنند (یعنی انتگرال حاصلضرب آن‌ها را روی فضای مماس می‌گیرند؛ به معادله (۷۶) مراجعه کنید).
(مارمولک‌ها و پروانه‌ها با مجوز □□□□□□□□□ □□□□□□□□ □□□□□□□□□□□□ ۰.۴ □□□□□□□□□□□□□□ از توییتر اقتباس شده‌اند.)

مکان را با یک بردار ضریب $f^{\widetilde{A}}(\exp_p(v)) \in \mathbb{R}^c$ نشان می‌دهد. فرض کنید

$$\rho\big(g^{A\widetilde{A}}_{p\leftarrow \exp_p v}\big) \tag{۶۶}$$

انتقال‌دهنده موازی سازگار با G برای ضرایب بردار ویژگی در امتداد ژئودزیک از $\exp_p(v)$ به p باشد. سپس ما پس‌کِش انتقال‌دهنده را در مختصات به صورت زیر تعریف می‌کنیم

$$\begin{aligned}\big[\mathrm{Exp}_p^* f\big]^A :\ \mathbb{R}^d \to \mathbb{R}^c, \quad v^A \mapsto\ & \big[\mathrm{Exp}_p^* f\big]^A(v^A) \\ :=\ & \rho\Big(g^{A\widetilde{A}}_{p\,\leftarrow\, \exp_p(\psi_p^A)^{-1}(v^A)}\Big)\cdot f^{\widetilde{A}}\Big(\exp_p\big(\psi_p^A\big)^{-1}(v^A)\Big)\,,\end{aligned} \tag{۶۷}$$

که در آن $v = \big(\psi_p^A\big)^{-1}(v^A) \in T_pM$ بردار مماس بدون مختصات است که توسط ضرایب v^A از طریق ψ_p^A به آن ارجاع داده می‌شود. همانطور که قبلاً ادعا شد، انتخاب پیمانه $\psi^{\widetilde{A}}_{\exp_p(v)}$ در $\exp_p(v)$ به دلیل استقلال مختصاتی تمام معادلات، بی‌ربط است و حذف می‌شود. به طور خاص، می‌توانستیم از هر پیمانه دیگری مانند $\psi^{\widetilde{B}}_{\exp_p(v)}$ در $\exp_p(v)$ استفاده کنیم، که منجر به تبدیلات پیمانه $\rho\big(g^{A\widetilde{B}}_{p\,\leftarrow\,\exp_p(v)}\big) = \rho\big(g^{A\widetilde{A}}_{p\,\leftarrow\,\exp_p(v)}\big)\rho\big(g^{\widetilde{B}\widetilde{A}}_{\exp_p(v)}\big)^{-1}$ برای انتقال‌دهنده طبق معادله (۳۵) و $f^{\widetilde{B}}\big(\exp_p(v)\big) = \rho\big(g^{\widetilde{B}\widetilde{A}}_{\exp_p(v)}\big) f^{\widetilde{A}}\big(\exp_p(v)\big)$ برای ضرایب بردار ویژگی طبق معادله (۲۸) می‌شود، که هنگام ترکیب هر دو عبارت، یکدیگر را خنثی می‌کنند.

با این حال، پس‌کِش انتقال‌دهنده $[\mathrm{Exp}_p^* f]^A$ همچنان به پیمانه در p بستگی دارد و بنابراین تحت تبدیلات پیمانه g_p^{BA} در p تبدیل می‌شود. مانند هر تابع مختصاتی، قانون تبدیل آن توسط تبدیلات پیمانه بر روی دامنه $\mathbb{R}^d$ و هم‌دامنه‌اش $\mathbb{R}^c$ تعیین می‌شود. بنابراین به صورت زیر داده می‌شود:

$$\big[\mathrm{Exp}_p^* f\big]^B \ =\ \rho\big(g_p^{BA}\big) \circ \big[\mathrm{Exp}_p^* f\big]^A \circ \big(g_p^{BA}\big)^{-1}, \tag{۶۸}$$

که توسط نمودار جابجایی زیر خلاصه می‌شود:

$$\begin{tikzcd}[column sep=huge, row sep=huge] \mathbb{R}^d \arrow[r, "{[\mathrm{Exp}_p^* f]^A}"] \arrow[d, "g_p^{BA}\cdot"'] & \mathbb{R}^c \arrow[d, "\rho(g_p^{BA})\cdot"] \\ \mathbb{R}^d \arrow[r, "{[\mathrm{Exp}_p^* f]^B}"'] & \mathbb{R}^c \end{tikzcd} \tag{۶۹}$$

همانطور که در شکل ۱۶ به تصویر کشیده شده است، $\left[\mathrm{Exp}_p^* f\right]^A$ و $\left[\mathrm{Exp}_p^* f\right]^B$ باید به عنوان دیدگاه ناظران محلی (چارچوب‌های مرجع) مختلف به میدان ویژگی در نظر گرفته شوند.

در اصل، می‌توان ساختارهای جایگزینی را برای پس‌کش میدان‌های ویژگی از M به T_pM در نظر گرفت. تعریف ما از تبدیلات میدان کرنل و کانولوشن‌های GM در بخش‌های ۲.۲.۴ و ۳.۲.۴ در ادامه، مستقل از این انتخاب خاص است.

۲.۲.۴ کرنل‌های مستقل از مختصات و تبدیلات میدان کرنل

کانولوشن‌های GM عملیات مستقلی از مختصات هستند که کرنل مشترک و یکسانی را در هر نقطه از خمینه اعمال می‌کنند. برای تفکیک واضح فرضیات، ابتدا تبدیلات میدان کرنل عمومی‌تری را مورد بحث قرار می‌دهیم که عملیات مستقلی از مختصات هستند اما الزام اشتراک‌گذاری وزن را کنار می‌گذارند. بنابراین، آن‌ها شبیه به کانولوشن‌های GM هستند اما در هر نقطه p از خمینه، کرنل بالقوه متفاوتی $\mathcal{K}_p$ اعمال می‌کنند. به منظور رعایت اصل کوواریانس، عبارات مختصاتی این کرنل‌ها باید به شیوه‌ای اصولی تبدیل شوند، با این حال، خود کرنل‌ها بدون محدودیت باقی می‌مانند.

کرنل‌های مستقل از مختصات: از آنجایی که کانولوشن‌ها در یادگیری عمیق بین میدان‌های بردارهای ویژگی با ابعاد $\mathbb{R}^{c_{\mathrm{in}}}$ و $\mathbb{R}^{c_{\mathrm{out}}}$ نگاشت برقرار می‌کنند، کرنل‌های کانولوشن ماتریس–مقدار $c_{\mathrm{out}} \times c_{\mathrm{in}}$ هستند. پیاده‌سازی‌های گسسته کانولوشن‌های d-بعدی روی فضاهای اقلیدسی معمولاً چنین کرنل‌هایی را به صورت آرایه‌هایی با شکل $(s_1, \dots, s_d,\ c_{\mathrm{out}}, c_{\mathrm{in}})$ نمایش می‌دهند. d محور اول یک شبکه فضایی از $s_1 \times \dots \times s_d$ پیکسل را نشان می‌دهند، که به هر کدام یک ماتریس $c_{\mathrm{out}} \times c_{\mathrm{in}}$ اختصاص داده شده است که در دو محور آخر کدگذاری شده است.[۲۵] در محیط پیوسته و اقلیدسی، چنین کرنل‌هایی را می‌توان به عنوان نگاشت‌هایی توصیف کرد

$$K : \mathbb{R}^d \to \mathbb{R}^{c_{\mathrm{out}} \times c_{\mathrm{in}}}\,, \tag{۷۰}$$

که یک ماتریس $c_{\mathrm{out}} \times c_{\mathrm{in}}$ را به هر نقطه از $\mathbb{R}^d$ اختصاص می‌دهد. همانطور که در بخش قبلی ۱.۲.۴ ذکر شد، ما کانولوشن‌های GM را به عنوان تطبیق پس‌کش انتقال‌دهنده $\mathrm{Exp}_p^* f_{\mathrm{in}}$ روی فضای مماس T_pM با یک کرنل $\mathcal{K}_p$ روی T_pM تعریف می‌کنیم. از آنجایی که فضاهای مماس مسطح هستند، طبیعی است که این تطبیق را مانند حالت کاملاً اقلیدسی معمول تعریف کنیم. بنابراین، ما کرنل‌های $\mathcal{K}_p$ را از طریق عبارات مختصاتی‌شان تعریف می‌کنیم که شکل معادله (۷۰) را دارند، یعنی،

$$\mathcal{K}_p^A : \mathbb{R}^d \to \mathbb{R}^{c_{\mathrm{out}} \times c_{\mathrm{in}}}\,. \tag{۷۱}$$

شکل ۱۷ یک کرنل داده‌شده بدون مختصات را روی T_pM و نمایش‌های آن را روی $\mathbb{R}^d$ نسبت به چارچوب‌های مرجع مختلف نشان می‌دهد.[۲۶]

قانون تبدیل بین نمایش‌های مختصاتی $\mathcal{K}_p^A$ و $\mathcal{K}_p^B$ یک کرنل $\mathcal{K}_p$ روی T_pM طبق معمول از قوانین تبدیل دامنه و هم‌دامنه آن‌ها پیروی می‌کند. در دامنه $\mathbb{R}^d$ قانون تبدیل توسط g_p^{BA} داده می‌شود، در حالی که قانون تبدیل $\mathbb{R}^{c_{\mathrm{out}} \times c_{\mathrm{in}}}$، همانند معادله (۵۰)، توسط یک ضرب همزمان از چپ با $\rho_{\mathrm{out}}\big(g_p^{BA}\big)$ و از راست با $\rho_{\mathrm{in}}\big(g_p^{BA}\big)^{-1}$ داده می‌شود. بنابراین، دو مختصات‌دهی کرنل $\mathcal{K}_p$ برای هر $v \in \mathbb{R}^d$ به صورت زیر با هم ارتباط دارند:

$$\mathcal{K}_p^B\big(g_p^{BA} v\big) \;=\; \rho_{\mathrm{out}}\big(g_p^{BA}\big) \cdot \mathcal{K}_p^A(v) \cdot \rho_{\mathrm{in}}\big(g_p^{BA}\big)^{-1}\,, \tag{۷۲}$$

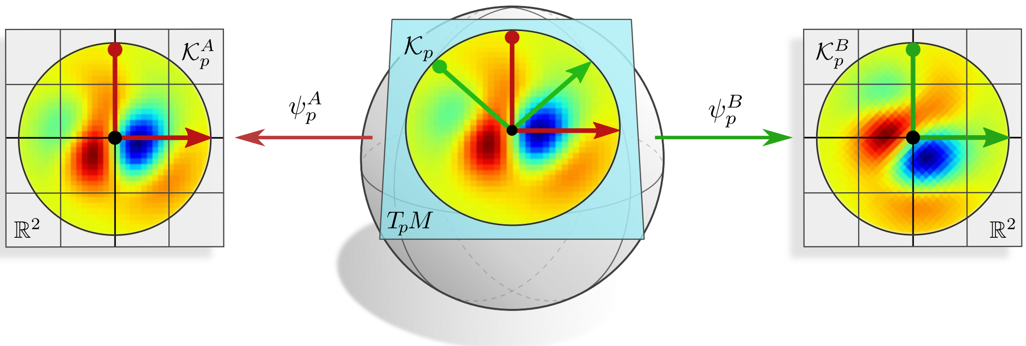
که توسط نمودار جابجایی زیر به تصویر کشیده شده است:

$$\begin{tikzcd}[column sep=huge]
\mathbb{R}^d \arrow[r, "\mathcal{K}_p^A"] \arrow[d, "g_p^{BA}\cdot"'] & \mathbb{R}^{c_{\mathrm{out}} \times c_{\mathrm{in}}} \arrow[d, "\rho_{\mathrm{out}}\left(g_p^{BA}\right)\ [\,\cdot\,]\ \rho_{\mathrm{in}}\left(g_p^{BA}\right)^{-1}"] \\
\mathbb{R}^d \arrow[r, "\mathcal{K}_p^B"'] & \mathbb{R}^{c_{\mathrm{out}} \times c_{\mathrm{in}}}
\end{tikzcd} \tag{۷۳}$$

همانند مثال‌های بخش ۱.۴، اصل کوواریانس تنها نیازمند یک رفتار تبدیل سازگار بین مختصات‌دهی‌های مختلف کرنل است اما محدودیتی بر خود کرنل اعمال نمی‌کند. بنابراین می‌توان کرنل‌های $\mathcal{K}_p$ را برای هر $p \in M$ و یک پیمانه دلخواه در p با یک کرنل ماتریس–مقدار نامحدود پارامتری کرد. ما میدان‌های هموار چنین کرنل‌هایی را به عنوان میدان‌های کرنل نشان می‌دهیم که نقش عمده‌ای در تحلیل ما از هموردایی ایزومتری کانولوشن‌های GM در بخش ؟؟ ایفا می‌کنند.

[۲۵] چیدمان واقعی حافظه به چارچوب یادگیری عمیق خاص مورد نظر بستگی دارد.

[۲۶] ما تأکید می‌کنیم که در اینجا یک کرنل بدون مختصات $\mathcal{K}_p$ را که روی T_pM داده شده است فرض می‌کنیم و عبارات مختصاتی آن $\mathcal{K}_p^X$ را روی $\mathbb{R}^d$ نسبت به چارچوب‌های مرجع X در نظر می‌گیریم. اشتراک‌گذاری وزن کانولوشنی بعداً ما را با این سؤال مواجه خواهد کرد که چگونه یک کرنل بدون مختصات $\mathcal{K}_p$ را روی T_pM با توجه به یک کرنل الگو K روی $\mathbb{R}^d$ تعریف کنیم. پیوست ؟؟ این دو مفهوم و ارتباط آنها را با G-راهبری کرنل توضیح می‌دهد.



شکل ۱۷: یک کرنل بدون مختصات $\mathcal{K}_p$ روی T_pM و عبارات مختصاتی آن $\mathcal{K}_p^X : \mathbb{R}^d \to \mathbb{R}^{c_{\mathrm{out}} \times c_{\mathrm{in}}}$ نسبت به پیمانه‌های ψ_p^X (فقط یکی از کانال‌های کرنل $c_{\mathrm{out}} \times c_{\mathrm{in}}$ نشان داده شده است). تبدیلات پیمانه‌ای که مختصات‌دهی‌های مختلف یک کرنل را به هم مرتبط می‌کنند، از قوانین تبدیل دامنه $\mathbb{R}^d$ و هم‌دامنه $\mathbb{R}^{c_{\mathrm{out}} \times c_{\mathrm{in}}}$ آن‌ها پیروی می‌کنند. بنابراین برای هر $v \in \mathbb{R}^d$ به صورت $\mathcal{K}_p^B\big(g_p^{BA} v\big) = \rho_{\mathrm{out}}\big(g_p^{BA}\big)\, \mathcal{K}_p^A(v)\, \rho_{\mathrm{in}}\big(g_p^{BA}\big)^{-۱}$ داده می‌شوند. یک میدان کرنل $\mathcal{K}$ روی M یک تخصیص هموار از کرنل‌ها بر روی فضاهای مماس است (تعریف ؟؟).
توجه داشته باشید که ما در اینجا فرض می‌کنیم کرنل روی T_pM داده شده است و سپس آن را نسبت به پیمانه‌های مختلف روی $\mathbb{R}^d$ بیان می‌کنیم. این از نظر مفهومی با وضعیت نشان داده شده در شکل‌های ۲، ۵، ۱۰ و ؟؟ متفاوت است، که در آنها فرض می‌کنیم یک کرنل الگو K روی $\mathbb{R}^d$ داده شده است و سپس $\mathcal{K}_p$ را روی T_pM از طریق اشتراک‌گذاری وزن کانولوشنی نسبت به یک چارچوب مرجع تعریف می‌کنیم. برای حفظ استقلال از مختصات در طول فرآیند اشتراک‌گذاری وزن، کرنل مشترک باید تحت تبدیلات پیمانه پایا (یا هموردا) باشد؛ به بخش ۳.۲.۴ و پیوست ؟؟ مراجعه کنید.

تبدیلات میدان کرنل مستقل از مختصات: با داشتن یک میدان کرنل هموار $\mathcal{K}$، می‌توانیم تبدیلات میدان کرنل را تعریف کنیم، که شبیه به کانولوشن‌ها هستند اما از این جهت متفاوتند که ممکن است در هر موقعیت مکانی یک کرنل متفاوت اعمال کنند. آنها یک میدان از بردارهای ویژگی خروجی $f_{\mathrm{out}}(p)$ را با انتگرال‌گیری از حاصلضرب کرنل مربوطه $\mathcal{K}_p$ و پس‌کشِ انتقال‌دهنده $\mathrm{Exp}_p^* f_{\mathrm{in}}$ از f_{in} روی T_pM محاسبه می‌کنند، یعنی،

$$f_{\mathrm{out}}(p) \;=\; \int_{T_pM} \mathcal{K}_p(v)\; \mathrm{Exp}_p^* f_{\mathrm{in}}(v)\; dv\,. \tag{۷۴}$$

برای بیان این تعریف بدون مختصات بر حسب مختصات، باید تمام کمیت‌ها را با عبارات مختصاتی‌شان جایگزین کرده و انتگرال‌گیری را از طریق پیمانه انتخاب‌شده از T_pM به $\mathbb{R}^d$ منتقل کرد. همانطور که در پیوست ؟؟ توضیح داده شده است، عنصر حجم ریمانی مناسب (پایا نسبت به پیمانه) برای یک پیمانه ψ_p^A به صورت زیر داده می‌شود:

$$\sqrt{|\eta_p^A|}\;\, dv^A, \tag{۷۵}$$

که در آن ضریب $\sqrt{|\eta_p^A|}$، تعریف شده در معادله (؟؟)، حجم (مثبت) پوشیده شده توسط چارچوب مرجع $[e_i^A]_{i=۱}^d$ در p است. بنابراین، عبارت مختصاتی تبدیل میدان کرنل به صورت زیر است:

$$f_{\mathrm{out}}^A(p) \;=\; \int_{\mathbb{R}^d} \mathcal{K}_p^A(v^A)\; \big[\mathrm{Exp}_p^* f_{\mathrm{in}}\big]^A(v^A)\; \sqrt{|\eta_p^A|}\; dv^A\,. \tag{۷۶}$$

استقلال از مختصات تبدیل میدان کرنل با بیان آن نسبت به یک پیمانه جایگزین ψ_p^B و نشان دادن اینکه میدان خروجی حاصل همانطور که انتظار می‌رود تبدیل می‌شود، تأیید می‌شود، که در واقع چنین است:

$$\begin{aligned}
f_{\mathrm{out}}^B(p) &\overset{(۱)}{=} \int_{\mathbb{R}^d} \mathcal{K}_p^B(v^B)\; \big[\mathrm{Exp}_p^* f_{\mathrm{in}}\big]^B(v^B)\; \sqrt{|\eta_p^B|}\, dv^B \\
&\overset{(۲)}{=} \int_{\mathbb{R}^d} \Big[\rho_{\mathrm{out}}\big(g_p^{BA}\big)\, \mathcal{K}_p^A\Big(\big(g_p^{BA}\big)^{-۱} v^B\Big)\, \rho_{\mathrm{in}}\big(g_p^{BA}\big)^{-۱}\Big]\; \big[\mathrm{Exp}_p^* f_{\mathrm{in}}\big]^B(v^B)\; \sqrt{|\eta_p^B|}\, dv^B \\
&\overset{(۳)}{=} \rho_{\mathrm{out}}\big(g_p^{BA}\big) \int_{\mathbb{R}^d} \mathcal{K}_p^A(v^A)\; \Big[\rho_{\mathrm{in}}\big(g_p^{BA}\big)^{-۱} \big[\mathrm{Exp}_p^* f_{\mathrm{in}}\big]^B\big(g_p^{BA} v^A\big)\Big]\; \sqrt{|\eta_p^A|}\, dv^A \\
&\overset{(۴)}{=} \rho_{\mathrm{out}}\big(g_p^{BA}\big) \int_{\mathbb{R}^d} \mathcal{K}_p^A(v^A)\; \big[\mathrm{Exp}_p^* f_{\mathrm{in}}\big]^A(v^A)\; \sqrt{|\eta_p^A|}\, dv^A \\
&\overset{(۵)}{=} \rho_{\mathrm{out}}\big(g_p^{BA}\big)\; f_{\mathrm{out}}^A(p)
\end{aligned} \tag{۷۷}$$

در اینجا ما از تعریف تبدیلات میدان کرنل و قانون تبدیل کرنل‌ها (معادله (۷۲)) در دو مرحله اول استفاده کردیم. مرحله سوم با بیرون کشیدن $\rho_{□□□}$ از انتگرال و جایگزینی v^B با $v^A = \left(g_p^{BA}\right)^{-۱} v^B$ به دست می‌آید، با استفاده از اینکه عنصر حجم $\sqrt{|\eta_p^A|}\, dv^A = \sqrt{|\eta_p^B|}\, dv^B$ به طور طراحی شده پایا نسبت به پیمانه است. دو مرحله آخر با شناسایی قانون تبدیل پس‌کِش انتقال‌دهنده میدان ویژگی در معادله (۶۸) و تعریف تبدیل میدان کرنل در پیمانه ψ_p^A به دست می‌آیند. توجه داشته باشید که استقلال از مختصات تَبدیل میدان کرنل، صحت قانون تبدیل کرنل در معادله (۷۲) را تأیید می‌کند.

یک تبدیل میدان کرنل تنها زمانی به خوبی تعریف می‌شود که انتگرال‌ها روی فضاهای مماس همگرا باشند، که به طور دقیق‌تر در بخش ؟؟ و پیوست ؟؟ مورد بحث قرار گرفته است. قضیه ؟؟ ثابت می‌کند که یک تکیه‌گاه فشرده برای کرنل‌های $\mathcal{K}_p$ برای تضمین این خوش‌تعریفی کافی است. این قضیه همچنین ثابت می‌کند که تبدیلات میدان کرنل که بر اساس میدان‌های کرنل هموار هستند، میدان‌های ویژگی ورودی هموار را به میدان‌های ویژگی خروجی هموار نگاشت می‌کنند.

۳.۲.۴ کانولوشن‌های □□ و کرنل‌های □-راهبر

آزادی تبدیلات میدان کرنل در اعمال یک کرنل متفاوت در هر مکان، به آنها اجازه تعمیم استنتاج آموخته‌شده را بر روی مکان‌های مختلف نمی‌دهد و بنابراین آنها را از نظر داده ناکارآمد می‌سازد. بنابراین، معمولاً کانولوشن‌ها در نظر گرفته می‌شوند، که می‌توان آن‌ها را به عنوان آن دسته از تبدیلات میدان کرنل خاصی دید که بر اساس میدان‌های کرنل کانولوشنی هستند، یعنی میدان‌های کرنلی که توسط یک کرنل الگوی واحد و مشترک پارامتری شده‌اند. همانند قبل، یک اشتراک‌گذاری وزن مستقل از مختصات نیازمند آن است که کرنل‌های الگو نسبت به پیمانه هموردا (G-راهبر) باشند. این هموردایی پیمانه‌ای کرنل‌های الگو تضمین می‌کند که الگوهایی که در ژست‌های هندسی مختلف و مرتبط با G ظاهر می‌شوند، تا یک تبدیل متناظر بردار ویژگی از طریق $\rho_{□□□}$ پاسخ یکسانی را برانگیزند.

اشتراک‌گذاری وزن کانولوشنی: فرض کنید $K : \mathbb{R}^d \to \mathbb{R}^{c_{□□□} \times c_{□□}}$ یک کرنل الگو باشد که باید روی تمام فضاهای مماس به اشتراک گذاشته شود. برای اینکه هیچ پیمانه خاصی را ترجیح ندهیم □ که با نیاز ما به استقلال از مختصات در تضاد است □ مجبوریم کرنل را با مختصات‌دهی‌ها در تمام پیمانه‌ها به طور همزمان به اشتراک بگذاریم. به طور ساده‌انگارانه، این به نظر می‌رسد که پیشنهاد می‌کند کرنل الگو را با قرار دادن $\mathcal{K}_p^X = K$ برای هر نقطه $p \in M$ و هر پیمانه ψ_p^X در p به اشتراک بگذاریم. در حالی که چنین تعریفی از اشتراک‌گذاری کرنل معقول به نظر می‌رسد، از اصل ما برای اشتراک‌گذاری توابع الگوی محلی به معنای دقیق پیروی نمی‌کند: به جای اشتراک‌گذاری مستقیم کرنل، مهم است که کل عملیات محلی را به اشتراک بگذاریم □ که در اینجا کل تبدیل انتگرالی در معادله (۷۶) است. از آنجایی که این عملیات بر حسب میدان کرنل $\mathcal{K}$ پارامتری می‌شود، این به طور غیرمستقیم منجر به اشتراک‌گذاری کرنل الگو می‌شود، با این حال، با نتیجه‌ای کمی متفاوت از اشتراک‌گذاری ساده‌انگارانه در نظر گرفته شده در بالا.

برای یافتن تعریف صحیح میدان‌های کرنل کانولوشنی GM طبق اصل ما برای اشتراک‌گذاری توابع الگوی محلی، ابتدا باید این عملیات محلی را شناسایی کنیم. ما این کار را با انتزاع تبدیلات میدان کرنل (در مختصات) به عنوان مجموعه‌ای از عملگرهای انتگرالی محلی به شکل زیر انجام می‌دهیم:

$$\mathscr{I}_{\mathcal{K},p}^A : C^\infty\left(\mathbb{R}^d, \mathbb{R}^c\right) \to \mathbb{R}^c, \quad F \mapsto \int_{\mathbb{R}^d} \mathcal{K}_p^A(\upsilon)\, F(\upsilon) \sqrt{|\eta_p^A|}\, d\upsilon\,, \tag{۷۸}$$

که در آن $C^\infty\left(\mathbb{R}^d, \mathbb{R}^c\right)$ فضای نگاشت‌های هموار از $\mathbb{R}^d$ به $\mathbb{R}^c$ را نشان می‌دهد. در کاربرد ما، این نگاشت‌های هموار صرفاً نمایش‌های میدان ویژگی محلی $[\mathrm{Exp}_p^* f]^A : \mathbb{R}^d \to \mathbb{R}^c$ هستند که از فضاهای مماس در p دیده می‌شوند و توسط تبدیل میدان کرنل به یک بردار ویژگی خروجی $f_{□□□}^A(p) = \mathscr{I}_{\mathcal{K},p}^A\left([\mathrm{Exp}_p^* f]^A\right)$ در p نگاشته می‌شوند. با توجه به کرنل الگوی ما $K : \mathbb{R}^d \to \mathbb{R}^{c_{□□□} \times c_{□□}}$، ما یک الگوی عملگر انتگرالی متناظر را تعریف می‌کنیم:

$$\mathfrak{I}_K : C^\infty\left(\mathbb{R}^d, \mathbb{R}^c\right) \to \mathbb{R}^c, \quad F \mapsto \int_{\mathbb{R}^d} K(\upsilon)\, F(\upsilon)\, d\upsilon\,, \tag{۷۹}$$

که نمایش میدان محلی F را در کرنل الگو K ضرب کرده و سپس حاصلضرب آنها را انتگرال می‌گیرد. توجه داشته باشید که $\mathfrak{I}_K$ به عنوان یک تابع الگو لزوماً نسبت به انتخاب‌های خاص پیمانه‌ها بی‌تفاوت است و بنابراین شامل ضریب حجم چارچوب نمی‌شود. یک طرح اشتراک‌گذاری وزن کانولوشنی مستقل از مختصات GM با الزام به اینکه این تابعک الگو با تمام عملگرهای انتگرالی منفرد در هر نقطه و در هر پیمانه موافق باشد، تحمیل می‌شود، یعنی:

$$\mathscr{I}_{\mathcal{K},p}^X = \mathfrak{I}_K \quad \text{برای هر پیمانه } \left(U^X, \psi^X\right) \in \mathscr{A}^G \text{ با } p \in U^X\,, \tag{۸۰}$$

که در آن $\mathscr{A}^G$ اطلس G (ماکسیمال) متناظر با ساختار G در نظر گرفته شده است؛ به معادله (۲۵) مراجعه کنید. این معادل اشتراک‌گذاری مستقیم کرنل الگوی محلی طبق زیر است:

$$\mathcal{K}_p^X = \frac{K}{\sqrt{|\eta_p^X|}} \quad \text{برای هر پیمانه } \left(U^X, \psi^X\right) \in \mathscr{A}^G \text{ با } p \in U^X\,, \tag{۸۱}$$

شیء	$\mathcal{K}_p^X$	K	F و $\left[\mathrm{Exp}_p^* f\right]^X$	$\sqrt{\lvert\eta^X\rvert}$	dv^X و $d\upsilon$
چگالی s	۰	۱−	۰	۱−	۱

جدول ۲: مروری بر توان‌های چگالی s برای اشیاء مختلف درگیر در تبدیلات میدان کرنل عمومی و کانولوشن‌های GM. عبارت مختصاتی یک s-چگالی با ضریب $\lvert\det g\rvert^s$ تبدیل می‌شود وقتی مختصات از طریق $g \in G$ تبدیل شوند. یک کرنل ماتریس-مقدار عمومی $\mathcal{K}_p^X$ طبق معادله (۷۲) یک ۰-چگالی است. همین امر برای میدان‌های ویژگی و پس‌کش‌های آنها نیز صادق است که قوانین تبدیل آنها در معادلات (۲۸) و (۶۸) داده شده است. کل انتگرال‌ده $\acute{\mathcal{K}}_p^X(v^X)[\mathrm{Exp}_p^* f]^X(v^X)\sqrt{\lvert\eta^X\rvert}\,dv^X$ از یک تبدیل میدان کرنل عمومی در معادله (۷۶) نیز یک ۰-چگالی دیده می‌شود □ توجه داشته باشید که این برای استقلال آن از مختصات، همانطور که در معادله (۷۷) نشان داده شده، ضروری است. از آنجایی که الگوی عملگر انتگرالی $\mathfrak{I}_K$ در معادله (۷۹) نسبت به هر انتخاب پیمانه بی‌تفاوت است، شامل ضریب حجم چارچوب $\sqrt{\lvert\eta^X\rvert}$ نمی‌شود. از آنجایی که با این وجود باید مانند عملگرهای انتگرالی $\mathscr{I}_{\mathcal{K},p}^X$ که زیربنای تبدیلات میدان کرنل هستند رفتار کند، کل انتگرال‌ده $K(\upsilon)F(\upsilon)\,d\upsilon$ از $\mathfrak{I}_K(F)$ ملزم به بودن یک ۰-چگالی است. این امر مستلزم آن است که خود کرنل‌های الگوی مشترک K مانند (۱−)-چگالی‌ها تبدیل شوند، که در قید G-راهبری در معادله (۸۳) منعکس شده است. توجه داشته باشید که این قانون تبدیل کرنل‌های الگو برای هموردایی G محلی کانولوشن‌های GM اکیداً ضروری است اگر ویژگی‌های خروجی باید مانند چگالی‌هایی با وزن ۰ تبدیل شوند؛ به معادله (۸۹) مراجعه کنید. برای دیدگاهی جایگزین، خواننده علاقه‌مند را به نتیجه ۱ در [؟] ارجاع می‌دهیم، جایی که ضریب دترمینان از اندازه‌های هار روی گروه‌های لی استخراج می‌شود.

که در آن ضریب نرمال‌سازی «چگالی کرنل» را به اندازه حجم چارچوب مرجع $\sqrt{\lvert\eta_p^X\rvert}$ کاهش می‌دهد. همانطور که در ادامه بحث می‌شود، این نرمال‌سازی برای هموردایی تحت گروه‌های تقارن غیر حافظ حجم مهم است.

ما میدان‌های کرنلی را که توسط یک کرنل مشترک K طبق معادله (۸۱) پارامتری شده‌اند، به عنوان میدان‌های کرنل کانولوشنی GM می‌نامیم. الزام همزمان برای اشتراک‌گذاری وزن و استقلال از مختصات منجر به یک قید هموردایی بر روی کرنل‌های الگو می‌شود. برای استخراج این قید، اشتراک‌گذاری کرنل در معادله (۸۱) را در قانون تبدیل کرنل در معادله (۷۲) قرار می‌دهیم، که نتیجه می‌دهد:

$$\frac{۱}{\sqrt{\lvert\eta_p^B\rvert}}\,K\big(g_p^{BA}\upsilon\big) \;=\; \frac{۱}{\sqrt{\lvert\eta_p^A\rvert}}\,\rho_{\square\square\square}\big(g_p^{BA}\big)\cdot K(\upsilon)\cdot\rho_{\square\square}\big(g_p^{BA}\big)^{-۱}\,. \qquad (۸۲)$$

از آنجایی که حجم‌های چارچوب‌های مرجع مختلف با $\sqrt{\lvert\eta_p^A\rvert} = \big\lvert\det(g_p^{BA})\big\rvert\sqrt{\lvert\eta_p^B\rvert}$ مرتبط هستند و از آنجایی که قانون تبدیل باید برای پیمانه‌های دلخواه مرتبط با G برقرار باشد، این قید G-راهبری را ایجاب می‌کند[۲۷]

$$K(g\upsilon) \;=\; \frac{۱}{\lvert\det g\rvert}\,\rho_{\square\square\square}(g)\cdot K(\upsilon)\cdot\rho_{\square\square}(g)^{-۱} \qquad \forall\ \upsilon\in\mathbb{R}^d,\ g\in G\,. \qquad (۸۳)$$

بر روی کرنل‌های الگو. همانطور که توسط ؟] ثابت شده است، این قید نیازمند آن است که کرنل‌های الگو عملگرهای نمایش [؟] باشند (تعمیم‌هایی از عملگرهای تانسوری کروی در مکانیک کوانتومی). از نظر نموداری، یک کرنل G-راهبر K باید جابجایی زیر را برآورده کند:

$$\begin{tikzcd}
\mathbb{R}^d \arrow[r, "K"] \arrow[d, "g\cdot"'] & \mathbb{R}^{c_{\square\square\square}\times c_{\square\square}} \arrow[d, "\frac{۱}{\lvert\det g\rvert}\,\rho_{\square\square\square}(g)\,[\,\cdot\,]\,\rho_{\square\square}(g)^{-۱}"] \\
\mathbb{R}^d \arrow[r, "K"'] & \mathbb{R}^{c_{\square\square\square}\times c_{\square\square}}
\end{tikzcd} \qquad (۸۴)$$

برای هر $g \in G$. توجه داشته باشید که ضریب دترمینان معکوس $\lvert\det g\rvert$ در قانون تبدیل کرنل باعث می‌شود که آن مانند یک (۱−)-چگالی ماتریس-مقدار تبدیل شود؛ برای جزئیات بیشتر به جدول ۲ مراجعه کنید. به طور شهودی، کرنل‌های G-راهبر دقیقاً آن کرنل‌هایی هستند که می‌توانند نسبت به چارچوب‌های مرجع دلخواه مرتبط با G به اشتراک گذاشته شوند بدون اینکه انتخاب خاص پیمانه بر نتیجه تأثیر بگذارد.[۲۸] بنابراین، ابهام هم‌ترازی کرنل‌ها □ که انگیزه اصلی این کار بود □ با اشتراک‌گذاری وزن اضافی بر روی تمام چارچوب‌های مرجع معادل (تمام پیمانه‌ها) در ساختار G در نظر گرفته شده GM حل می‌شود.

---

[۲۷] برخلاف کارهای قبلی [؟ ؟ ؟ ؟ ؟]، این قید شامل ضریب $\lvert\det g\rvert$ است. این ضریب در آن کارها ظاهر نشد زیرا همه آنها گروه‌های ساختاری متعامد d (یا زیرگروه‌های آن) را در نظر گرفته بودند که برای آنها ضریب دترمینان صفر می‌شود.

[۲۸] قید G-راهبری را می‌توان به صورت $\forall\,\upsilon\in\mathbb{R}^d,\ g\in G\ K(\upsilon) = \lvert\det g\rvert^{-۱}\rho_{\square\square\square}(g)\cdot K(g^{-۱}\upsilon)\cdot\rho_{\square\square}(g)^{-۱}$ بازنویسی کرد، که تأکید می‌کند کرنل‌های G-راهبر پایاها تحت عمل پیمانه در سمت راست هستند. یک کرنل G-راهبر، با پایا بودن تحت تبدیلات پیمانه، منجر به همان کرنل بدون مختصات $\mathcal{K}_p$ در p می‌شود، زمانی که نسبت به هر چارچوب مرجعی در G_pM به اشتراک گذاشته شود.

قبل از پرداختن به کانولوشن‌های GM، در مورد فضای کرنل‌های G-راهبر توضیح می‌دهیم. توجه داشته باشید که مجموعه

$$\mathscr{K} \ := \ \left\{ K \colon \mathbb{R}^d \to \mathbb{R}^{c_{\text{out}} \times c_{\text{in}}} \right\}. \tag{۸۵}$$

کرنل‌های عمومی، یعنی نه لزوماً هموردا، با مجهز شدن به جمع و ضرب اسکالر استاندارد روی $\mathbb{R}^{c_{\text{out}} \times c_{\text{in}}}$ یک فضای برداری تشکیل می‌دهد. از آنجایی که قید G-راهبری در معادله (۸۳) خطی است، فضای کرنل را به یک زیرفضای خطی محدود می‌کند:

$$\mathscr{K}^G_{\rho_{\text{in}},\rho_{\text{out}}} \ := \ \left\{ K \colon \mathbb{R}^d \to \mathbb{R}^{c_{\text{out}} \times c_{\text{in}}} \;\middle|\; K(gv) = \frac{1}{|\det g|}\, \rho_{\text{out}}(g) \cdot K(v) \cdot \rho_{\text{in}}(g)^{-1} \ \ \forall\ v \in \mathbb{R}^d,\ g \in G \right\}. \tag{۸۶}$$

بنابراین می‌توان برای یک پایه از کرنل‌های G-راهبر راه‌حلی یافت که کانولوشن GM را می‌توان بر حسب آن پارامتری کرد. در حالی که این فضا در تئوری معمولاً بی‌نهایت-بعدی است، در عمل اغلب گسسته می‌شود، به طوری که در نهایت با یک پایه متناهی $\{K_1, \dots, K_N\}$ از کرنل‌های G-راهبر مواجه می‌شویم. سپس یک مجموعه از وزن‌های با مقادیر حقیقی $\{w_1, \dots, w_N\}$ که $w_i \in \mathbb{R}$ هستند و در طول فرآیند آموزش بهینه‌سازی می‌شوند، کانولوشن را با $K = \sum_{i=1}^N w_i K_i$ پارامتری می‌کنند. توجه داشته باشید که ابعاد کاهش‌یافته (زیر)فضای کرنل‌های G-راهبر به معنای بهبود کارایی پارامتر در مقایسه با کانولوشن‌های معمولی است.

بخش ?? به بحث در مورد راه‌حل‌های تحلیلی نمونه‌ای از فضاهای کرنل هموردای بازتابی برای نمایش‌های گروهی مختلف گروه بازتاب $\mathscr{R}$ می‌پردازد. کرنل‌های حاصل که با انواع مختلفی از تقارن‌های بازتابی مشخص می‌شوند، در جدول ?? به تصویر کشیده شده‌اند.

مثال‌های بیشتری را می‌توان در ادبیات مربوط به شبکه‌های عصبی کانولوشنی راهبر یافت: یک راه‌حل تحلیلی برای قید فضای کرنل برای گروه ساختاری متعامد خاص $\mathrm{SO}(3)$ در سه بعد و نمایش‌های تحویل‌ناپذیر آن توسط ؟ [ ] ارائه شد. ؟ [ ] این رویکرد را برای پوشش نمایش‌های گروهی دلخواه تعمیم دادند و قید فضای کرنل را برای هر نمایشی از ۲ و تمام زیرگروه‌های آن $G \leq 2$ حل کردند □ یک پیاده‌سازی در `https://quva-lab.github.io/e2cnn/api/e2cnn.kernels.html` موجود است. برای گروه‌های ساختاری متناهی، قید را می‌توان به صورت عددی نیز حل کرد، همانطور که توسط ؟ [ ] توضیح داده شده است. یک استراتژی راه‌حل کلی‌تر، که برای گروه‌های ساختاری فشرده دلخواه G (و بنابراین همه موارد ذکر شده در بالا) کاربرد دارد، توسط ؟ [ ] پیشنهاد شد. این راه‌حل قضیه کلاسیک ویگنر-اکارت [؟ ؟ ؟ ؟] را به یک قضیه ویگنر-اکارت برای کرنل‌های G-راهبر تعمیم می‌دهد، که کرنل‌ها را بر حسب توابع پایه هارمونیک، ضرایب کلبش-گوردان و اندومورفیسم‌های نمایش‌ها (عناصر ماتریسی کاهش‌یافته تعمیم‌یافته) بیان می‌کند. ما برای مروری دقیق‌تر بر کارهای قبلی و مرتبط با کرنل‌های راهبر به [؟] ارجاع می‌دهیم.

کانولوشن‌های مستقل از مختصات □□: با داشتن یک کرنل الگوی G-راهبر $K \in \mathscr{K}^G_{\rho_{\text{in}},\rho_{\text{out}}}$، کانولوشن $K \star GM$ با این کرنل به عنوان تبدیل میدان کرنل با میدان کرنل کانولوشنی GM متناظر تعریف می‌شود، که برای هر نقطه $p \in M$ و هر پیمانه ψ^X_p در $\mathcal{K}^X_p = K/\sqrt{|\eta^X|}$ صدق می‌کند. با قرار دادن میدان کرنل کانولوشنی GM در معادله (۷۶)، یعنی تبدیل میدان کرنل، عبارت مختصاتی کانولوشن GM به صورت زیر ساده می‌شود:

$$f^A_{\text{out}}(p) \ = \ \left[K \star f_{\text{in}}\right]^A(p) \ := \ \int_{\mathbb{R}^d} K(v) \left[\mathrm{Exp}^*_p f_{\text{in}}\right]^A(v)\, dv \ = \ \mathfrak{I}_K\left(\left[\mathrm{Exp}^*_p f_{\text{in}}\right]^A\right). \tag{۸۷}$$

بنابراین، این به سادگی با تطبیق پس‌کش انتقال‌دهنده $[\mathrm{Exp}^*_p f_{\text{in}}]^A$ از میدان ویژگی در یک پیمانه دلخواه انتخاب شده ψ^A_p با کرنل کانولوشن مستقل از پیمانه K به دست می‌آید. بنابراین کانولوشن‌های مستقل از مختصات GM به راحتی با ۱) انتخاب چارچوب‌های مرجع دلخواه، ۲) پس‌کش (و انتقال) میدان‌های ویژگی به مختصات‌دهی‌های فضای مماس و ۳) انقباض آنها در آنجا با یک کرنل G-راهبر (قابل آموزش) پیاده‌سازی می‌شوند.

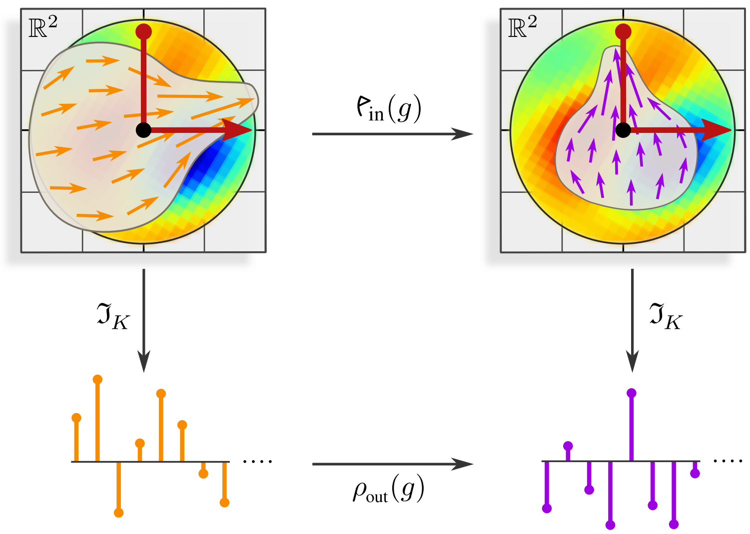
کانولوشن‌های GM چندین ویژگی تقارنی مرتبط را از خود نشان می‌دهند:

- استقلال از مختصات: به عنوان موارد خاصی از تبدیلات میدان کرنل، کانولوشن‌های GM به طور (غیرفعال) مستقل از مختصات هستند، یعنی معادله (۷۷) در مورد آنها صدق می‌کند.
- هموردایی ایزومتری سراسری: آنها تحت عمل فعال و سراسری ایزومتری‌های حافظ ساختار G در Isom_{GM} بر روی میدان‌های ویژگی هموردا هستند. بخش‌های ۳.۴ و به ویژه ?? این ویژگی را به تفصیل مورد بحث قرار می‌دهند.
- هموردایی G محلی: الگوی عملگر انتگرالی $\mathfrak{I}_K$ به دلیل G-راهبری K خود G-هموردا است. بنابراین، هر تبدیل G یک نمایش میدان ویژگی محلی روی $\mathbb{R}^d$ منجر به تبدیل متناظر بردار ویژگی حاصل خواهد شد؛ به شکل ۱۸ مراجعه کنید. بنابراین، تبدیلات G مستقل الگوهایی که در نقاط مختلف $p_i \in M$ متمرکز شده‌اند، منجر به تبدیلات ویژگی خروجی مستقل در این نقاط می‌شوند (این فقط در این نقاط برقرار است و نیازمند کرنل‌هایی با تکیه‌گاه فشرده است که کل میدان دید آنها طبق تبدیل G تبدیل شود).

برای دقیق‌تر کردن نکته آخر، ما تبدیلات G فعال نمایش‌های میدان ویژگی محلی را در $C^\infty(\mathbb{R}^d, \mathbb{R}^c)$ به صورت زیر تعریف می‌کنیم[۲۹]

$$\rhd_{\text{in}} \colon\ G \times C^\infty(\mathbb{R}^d, \mathbb{R}^c) \to C^\infty(\mathbb{R}^d, \mathbb{R}^c)\,, \quad F \ \mapsto\ \rhd_{\text{in}}(g)F \ := \ \rho_{\text{in}}(g) \circ F \circ g^{-1}\,, \tag{۸۸}$$

[۲۹] $\rhd = \mathrm{Res}^{\mathbb{R}^d \rtimes G}_G \mathrm{Ind}^{\mathbb{R}^d \rtimes G}_G \rho$ به طور رسمی با القای نمایش ρ G به یک نمایش $(\mathbb{R}^d \rtimes G)$ $\mathrm{Ind}^{\mathbb{R}^d \rtimes G}_G \rho$، و سپس محدود کردن آن به G داده می‌شود. یک توضیح شهودی از نمایش‌های القایی و محدود شده را می‌توان در پیوست □ از [؟] یافت در حالی که [؟] این موضوع را به طور رسمی‌تر بررسی می‌کند.



شکل ۱۸: همورداییِ G محلی الگوی عملگر انتگرالی مشترک $\mathfrak{I}_K$ که زیربنای یک کانولوشن $K\star GM$ است. یک تبدیل G فعال $\rhd_{\mathrm{in}}(g)$ از یک نمایش میدان محلی روی $\mathbb{R}^d$ بردارهای ویژگی را از v به $g^{-1}v$ منتقل می‌کند و آنها را علاوه بر این از طریق $\rho_{\mathrm{in}}(g)$ تبدیل می‌کند. در حالی که اولی ویژگی‌ها را به صورت فضایی جابجا می‌کند، دومی ضرایب عددی آنها را تبدیل می‌کند (که در شکل به صورت چرخش و مقیاس‌بندی بردارهای (مماس) منفرد به تصویر کشیده شده است). اعمال $\mathfrak{I}_K$ به هر دو ورودی منجر به بردارهای ویژگی خروجی متفاوتی می‌شود. با این حال، به دلیل همورداییِ G از $\mathfrak{I}_K$، تضمین می‌شود که پاسخ‌ها با $\rho_{\mathrm{out}}(g)$ مرتبط هستند؛ به معادله (۸۹) مراجعه کنید. بنابراین، یک تبدیل G فعال از یک میدان ورودی منجر به یک تبدیل G فعال متناظر از بردار ویژگی خروجی می‌شود. توجه داشته باشید که همورداییِ G از $\mathfrak{I}_K$ یک نتیجه مستقیم از G–راهبری K است.

که در آن فرض می‌کنیم F از نوع ρ_{in} است. به طور شهودی، $\rhd_{\mathrm{in}}$ با انتقال فعال بردارهای ویژگی $F(g^{-1}v) \in \mathbb{R}^c$ از $g^{-1}v$ به v و تبدیل آنها با $\rho_{\mathrm{in}}(g)$ بر روی میدان‌های F عمل می‌کند – این تعریف «معمولی» تبدیلات فعال میدان‌های ویژگی F بر روی فضاهای اقلیدسی $\mathbb{R}^d$ است. همورداییِ G ادعا شده برای $\mathfrak{I}_K$ به راحتی با اعمال آن به یک ورودی تبدیل شده، و سپس جایگزینی و استفاده از G–راهبری K دیده می‌شود:

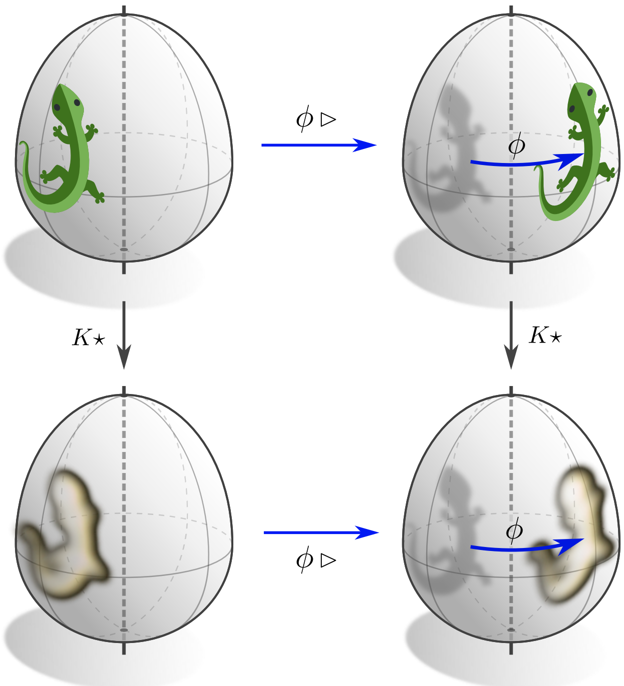
$$
\begin{aligned}
\mathfrak{I}_K\big(\rhd_{\mathrm{in}}(g)F\big) &= \mathfrak{I}_K\big(\rho_{\mathrm{in}}(g)\circ F\circ g^{-1}\big) && \big(\text{تعریف } \rhd_{\mathrm{in}}\text{، معادله (۸۸)}\big)\\
&= \int_{\mathbb{R}^d} K(v)\ \rho_{\mathrm{in}}(g)\, F\big(g^{-1}v\big)\, dv && \big(\text{تعریف } \mathfrak{I}_K\text{، معادله (۷۹)}\big)\\
&= \int_{\mathbb{R}^d} K(g\tilde{v})\ \rho_{\mathrm{in}}(g)\, F\big(\tilde{v}\big)\ |\det g|\, d\tilde{v} && \big(\text{جایگزینی } \tilde{v} = g^{-1}v\big)\\
&= \int_{\mathbb{R}^d} \rho_{\mathrm{out}}(g)\, K(\tilde{v})\ F\big(\tilde{v}\big)\, d\tilde{v} && \big(G\text{–راهبری } K\text{، معادله (۸۳)}\big)\\
&= \rho_{\mathrm{out}}(g)\ \mathfrak{I}_K(F) && \big(\text{تعریف } \mathfrak{I}_K\text{، معادله (۷۹)}\big)
\end{aligned}
\tag{۸۹}
$$

بنابراین، یک تبدیل فعال یک نمایش میدان ویژگی محلی F روی یک مختصات‌دهی فضای مماس با $\rhd_{\mathrm{in}}(g)$ تضمین می‌کند که منجر به تبدیل بردار ویژگی خروجی حاصل با $\rho_{\mathrm{out}}(g)$ می‌شود. به عبارت دیگر، ویژگی‌هایی که در ژست‌های هندسی مختلف مرتبط با G ظاهر می‌شوند، تا یک تبدیل از طریق ρ_{out} پاسخ یکسانی را برمی‌انگیزند. این به طور خلاصه در قالب یک نمودار جابجایی به صورت زیر است:

$$
\begin{tikzcd}
C^\infty(\mathbb{R}^d, \mathbb{R}^{c_{\mathrm{in}}}) \arrow[r, "\rhd_{\mathrm{in}}(g)"] \arrow[d, "\mathfrak{I}_K"'] & C^\infty(\mathbb{R}^d, \mathbb{R}^{c_{\mathrm{in}}}) \arrow[d, "\mathfrak{I}_K"] \\
\mathbb{R}^{c_{\mathrm{out}}} \arrow[r, "\rho_{\mathrm{out}}(g)"'] & \mathbb{R}^{c_{\mathrm{out}}}
\end{tikzcd}
\tag{۹۰}
$$

شکل ۱۸ یک تفسیر بصری از این ویژگی همورداییِ $\mathfrak{I}_K$ ارائه می‌دهد.

توجه داشته باشید که همورداییِ تحت تبدیلات G محلی در معادله (۸۹) دقیقاً به قید G–راهبری همانطور که در معادله (۸۳) است، یعنی به طور خاص، با ضریب دترمینان $|\det g|^{-1}$ که باعث می‌شود کرنل مانند یک (-1)–چگالی تبدیل شود، نیاز دارد. این ضریب به تعریف ما از اشتراک‌گذاری



شکل ۱۹: گفته می‌شود یک لایه شبکه تحت ایزومتری‌ها تناوب‌پذیر است زمانی که با عمل آن‌ها بر میدان‌های ویژگی جابجا می‌شود. $\phi \in \mathrm{Isom}_{GM} M - □$ که تقارن‌های G-ساختار هستند، تناوب‌پذیر باشند. در معادلات، کانولوشن $K\star$ تحت عمل ایزومتری ϕ تناوب‌پذیر است زمانی که رابطه $K \star \big(\phi \triangleright f_{□□}\big) = \phi \triangleright \big(K \star f_{□□}\big)$ را برای هر انتخابی از میدان ورودی $f_{□□}$ برآورده کند. این رابطه توسط نمودار جابجایی در معادله (۹۲) بصری‌سازی شده است که تفسیر گرافیکی آن در این شکل نشان داده شده است.

اینکه $\mathrm{Isom}_{GM} M - □$-تناوب‌پذیر هستند بر این حقایق تکیه دارد که ۱) کرنل‌ها در کل منیفولد به اشتراک گذاشته می‌شوند، ۲) ایزومتری‌ها عقب‌کشی انتقال‌دهنده میدان‌های ویژگی را حفظ می‌کنند و ۳) اینکه Isom_{GM} تبدیل‌های گیج با مقدار G را القا می‌کند که توسط G-استیریبل بودن کرنل در نظر گرفته می‌شود.

(مارمولک‌ها تحت مجوز □□□□□□□□□□□ □□□□□□□ □□□□□□□□ ۴.۰ □□□□□□□□□□□□□اقتباس شده با مجوز (.□□□□□□□

وزن کانولوشنی در معادله (۸۱) با نرمال‌سازی با حجم‌های چارچوب مرجع $\sqrt{|\eta_p^X|}$ برمی‌گردد. بنابراین، اشتراک‌گذاری وزن ساده‌انگارانه‌ای که در ابتدای این بخش ذکر شد، به رفتار تبدیل مطلوب منجر نمی‌شد. به عبارت دیگر: هر دو نسخه ساده‌انگارانه و نرمال‌شده اشتراک‌گذاری کرنل مستقل از مختصات هستند و بنابراین هر دو تحت تبدیلات پیمانه غیرفعال □ به ویژه آنهایی که حجم چارچوب را تغییر می‌دهند □ به طور سازگار رفتار می‌کنند. با این حال، در مورد اشتراک‌گذاری کرنل ساده‌انگارانه، این امر توسط پایایی عنصر حجم ریمانی $\sqrt{|\eta_p^A|}\, dv^A = \sqrt{|\eta_p^B|}\, dv^B$ مراقبت می‌شود. با حذف این ضریب در اشتراک‌گذاری وزن نرمال‌شده، سازگاری رفتار تبدیل دیگر توسط خود اندازه انتگرال‌گیری تضمین نمی‌شود □ که مستلزم آن است که خود کرنل‌های G-راهبر تغییرات حجم را از طریق ضریب دترمینان توضیح دهند. فقط دومی به تبدیلات فعال تعمیم می‌یابد، جایی که فقط میدان ویژگی تبدیل می‌شود، در حالی که اندازه انتگرال‌گیری پایا می‌ماند.

از آنجایی که تعریف ما از کانولوشن‌های GM امکان خمینه‌های ریمانی، ساختارهای □ و انواع میدان دلخواه را فراهم می‌کند، بسیار عمومی است و طیف گسترده‌ای از کارهای مرتبط را پوشش می‌دهد. ما این ادعا را در بخش ?? اثبات می‌کنیم، جایی که بسیاری از CNNها را روی فضاهای آفین اقلیدسی $\mathbb{E}_d$، کره S^2 و خمینه‌ها یا مش‌های عمومی به عنوان نمونه‌های خاصی از معادله (۸۷) توضیح می‌دهیم. برای یک نمای کلی و طبقه‌بندی این مدل‌ها، به جدول ?? مراجعه کنید.

۳.۴ تناوب‌پذیری ایزومتری

با توجه به اینکه یک منیفولد تقارن‌هایی نشان می‌دهد، معمولاً مطلوب است که شبکه‌های عصبی این تقارن‌ها را رعایت کنند، یعنی تحت عمل آن‌ها بر میدان‌های ویژگی تناوب‌پذیر باشند. $\mathrm{Isom}_{GM} M - □$-تناوب‌پذیر باشند، که به این معنی است که آن‌ها با عمل ایزومتری‌ها در Isom_{GM} بر میدان‌های ویژگی جابجا می‌شوند، همانطور که در شکل ۱۹ بصری‌سازی شده است.[۳۰] در معادلات، $□M - □\star$ تناوب‌پذیر است زمانی که رابطه

$$K \star \big(\phi \triangleright f_{□□}\big) \;=\; \phi \triangleright \big(K \star f_{□□}\big) \qquad \forall\ \phi \in \mathrm{Isom}_{GM} \tag{۹۱}$$

[۳۰] به یاد داشته باشید که عمل بر میدان‌های ویژگی مستقل از مختصات GM تنها برای ایزومتری‌های حفظ‌کننده G-ساختار در Isom_{GM} قابل تعریف است. بنابراین حتی امکان تعریف مفهوم تناوب‌پذیری ایزومتری برای ایزومتری‌هایی که تقارن‌های G-ساختار نیستند وجود ندارد. توجه داشته باشید که این بدون از دست دادن کلیت است زیرا همیشه می‌توان گروه ساختار $G = d$ را انتخاب کرد که برای آن $\mathrm{Isom}_{GM} = \mathrm{Isom}(M)$ با کل گروه ایزومتری منطبق می‌شود.

را برای هر میدان ورودی ممکن $f_{□□}$ برآورده کند، یعنی زمانی که نمودار زیر جابجا باشد:

$$\begin{tikzcd} f_{□□} \arrow[r, mapsto, "\phi \triangleright"] \arrow[d, "K\star"'] & \phi \triangleright f_{□□} \arrow[d, "K\star"] \\ f_{□□□} \arrow[r, mapsto, "\phi \triangleright"'] & \phi \triangleright f_{□□□} \end{tikzcd} \tag{۹۲}$$

به عنوان اولین قدم به سوی اثبات تناوب‌پذیری ایزومتری □ $- M$□$]$ $[\mathrm{Exp}_p^* f_{□□}]^A$ از میدان ورودی $f_{□□}$ تعریف می‌شوند. از آنجا که ایزومتری‌ها به طور تعریف هندسه ریمانی M را حفظ می‌کنند، به ویژه نگاشت نمایی ریمانی و انتقال‌دهنده‌های لوی–چیویتا را حفظ می‌کنند؛ به بخش ؟؟ و شکل ؟؟ مراجعه کنید.[۳۱] این امر دلالت دارد بر اینکه عقب‌کشی انتقال‌دهنده میدان پیش‌برده $\phi \triangleright f_{□□}$ در $\phi(p)$ تنها با تبدیل گیج القا شده توسط ایزومتری از عقب‌کشی انتقال‌دهنده میدان اصلی $f_{□□}$ در p متفاوت خواهد بود، یعنی:

$$\left[\mathrm{Exp}_{\phi(p)}^*(\phi \triangleright f_{□□})\right]^A \;=\; \rho_{□□}\big(g_\phi^{A\widetilde{A}}(p)\big) \circ \left[\mathrm{Exp}_p^* f_{□□}\right]^{\widetilde{A}} \circ g_\phi^{A\widetilde{A}}(p)^{-۱}\,; \tag{۹۳}$$

مقایسه کنید با معادله (۶۸) و، برای فرمول‌بندی مستقل از مختصات، قضیه ؟؟.

با توجه به این همانی، تناوب‌پذیری ایزومتری □ $- M$□ $-$□.(۹۳) (۸۷) $- \; GM \; :$ □$\tilde{v}$ = □v $^{A\widetilde{A}}_{\phi}(p)^{-۱}$ قدم سوم را توجیه می‌کند. در قدم چهارم G–استیریبل بودن کرنل قالب، یعنی معادله (۸۳)، اعمال می‌شود (معادله (۴۴) را به یاد بیاورید که بیان می‌کند تبدیل‌های گیج القا شده توسط Isom_{GM} دارای مقدار G هستند). نتیجه این است که بردار ویژگی خروجی حاصل توسط تبدیل گیج القا شده تبدیل می‌شود. پس از شناسایی این امر به عنوان بیان مختصاتی پیش‌برد میدان خروجی در معادله (۴۶)، گزاره پیروی می‌کند. از آنجا که همه قدم‌ها برای ایزومتری‌های دلخواه در Isom_{GM} معتبر هستند، می‌بینیم که □ $- M$□.$-\mathrm{Isom}(M)$ که ممکن است G–ساختار را رعایت نکنند تناوب‌پذیر نیستند، با این حال، تناوب‌پذیری کامل ایزومتری برای گروه‌های ساختار متعامد $G = d$ (یا ابرگروه‌های آن) تضمین شده است.

میدان‌های کرنل ناوردا: تحلیل عمیق‌تری از تناوب‌پذیری ایزومتری تبدیل‌های میدان کرنل عمومی را می‌توان در بخش‌های ؟؟ و ؟؟ یافت. نتیجه اصلی این بررسی قضیه ؟؟ است که بیان می‌کند تناوب‌پذیری ایزومتری یک تبدیل میدان کرنل، ناوردایی ایزومتری میدان کرنل آن را نتیجه می‌دهد و بالعکس. شکل ؟؟ چنین میدان کرنل ناوردایی را بصری‌سازی می‌کند که ملزم به اشتراک وزن‌ها بر روی مدارات عمل ایزومتری است. ناوردایی مورد نیاز میدان کرنل از نظر شهودی قابل فهم است زیرا تناوب‌پذیری ایزومتری مطمئناً نیازمند ثابت بودن استنتاج شبکه بر روی هر مدار است. این نتیجه انتزاعی تناوب‌پذیری ایزومتری □— $- - - GM - GM$ $\mathrm{Isom}_{GM}M$ ناوردا هستند، نتیجه می‌دهد؛ به قضیه ؟؟ و شکل ؟؟ مراجعه کنید. G–استیریبل بودن کرنل قالب بدین ترتیب ناوردایی کرنل‌ها تحت عمل زیرگروه‌های پایدار کننده گروه ایزومتری را توضیح می‌دهد.

فضاهای همگن: در حالی که تقاضای تناوب‌پذیری ایزومتری نیازمند آن است که کرنل‌ها بر روی مدارات گروه ایزومتری به اشتراک گذاشته شوند، به طور کلی نیازمند اشتراک وزن کانولوشنی در کل منیفولد نیست. یک استثنای مهم حالت منیفلدهایی است که فضاهای همگن گروه ایزومتری خود هستند، مانند $\mathbb{R}^d$ یا کره $S^۲$. به طور تعریف، عمل ایزومتری در چنین فضاهایی متعدی است، یعنی تنها یک مدار منفرد وجود دارد. در نتیجه، تنها یک کرنل مستقل وجود خواهد داشت که از طریق عمل گروه ایزومتری در کل فضا به اشتراک گذاشته می‌شود. قضیه ؟؟ در بخش ؟؟ ثابت می‌کند که تبدیل‌های میدان کرنل تناوب‌پذیر ایزومتری در فضاهای همگن لزوماً کانولوشن‌های مستقل از مختصات هستند. این مشاهده پیوند رسمی بین نظریه ما و کارهای قبلی در مورد شبکه‌های کانولوشنی در فضاهای همگن توسط ؟]؛ ؟] و ؟] برقرار می‌کند که کانولوشن‌ها را از طریق تناوب‌پذیری آن‌ها نسبت به تقارن‌های سراسری فضای زیرین تعریف می‌کنند.

تناوب‌پذیری دیفئومورفیسم: خواننده ممکن است تعجب کند که آیا امکان کاملاً دیفئومورفیسم تناوب‌پذیر کردن های□□□ مستقل از مختصات ما وجود دارد یا نه. همانطور که به راحتی می‌توان دید، عملیات نقطه‌ای از بخش ۱.۴، یعنی ۱×۱□□□□□□□□□□□□–، بایاس‌ها و غیرخطی‌ها، از قبل دیفئومورفیسم تناوب‌پذیر هستند. به طور خاص، بگذارید

$$\mathrm{Diff}_{GM} \;:=\; \Big\{\phi \in \mathrm{Diff}(M) \,\Big|\, \big[\phi_*(e_i)\big]_{i=۱}^d \in GM \quad \forall\, [e_i]_{i=۱}^d \in GM\Big\} \;\leq\; \mathrm{Diff}(M) \tag{۹۴}$$

زیرگروه دیفئومورفیسم‌های حفظ‌کننده G–ساختار باشد، یعنی آنالوگ معادله (۴۲) بدون نیاز به اینکه ϕ ایزومتری باشد. مشابه معادله (۴۴) و قضیه ؟؟، بیان‌های مختصاتی (تبدیل‌های گیج القا شده) دیفئومورفیسم‌های حفظ‌کننده G–ساختار تضمین شده‌اند که مقادیری در G بگیرند، یعنی:

$$\phi \in \mathrm{Diff}_{GM} \quad\Longleftrightarrow\quad g_\phi^{A\widetilde{A}}(p) \in G \;\; \forall\, p \in M\,. \tag{۹۵}$$

[۳۱] به طور کلی‌تر، هر زمان که یک اتصال جایگزین G–سازگار برای انتقال بردارهای ویژگی انتخاب شود، فرض می‌کنیم این اتصال تحت عمل Isom_{GM} ناوردا باشد؛ به بخش ؟؟ مراجعه کنید. این فرض برای همه مدل‌هایی که در بررسی ادبیات در بخش ؟؟ پوشش داده شده‌اند، برآورده می‌شود.