



دانشکده علوم ریاضی  
گروه ریاضی کاربردی

پایان نامه

برای دریافت درجه کارشناسی ارشد در رشته  
ریاضی کاربردی، گرایش آمار ریاضی

عنوان

# تحلیل حساسیت مدل‌های کامپیوتری با استفاده از مدل جانشین بیزی

استاد راهنما

دکتر داود شاهسونی

دانشجو

فاطمه نورمحمدزاده

بهمن ۱۳۹۲

به نام آن که جان را فکرت آموخت



دانشگاه علمی کاربردی

مدیریت تحصیلات تکمیلی

فرم شماره

(۶)

باسمه تعالی

شماره:

تاریخ:

ویرایش:

### فرم صورت جلسه دفاع از پایان نامه تحصیلی دوره کارشناسی ارشد

با تأییدات خداوند متعال و با استعانت از حضرت ولی عصر (عج) نتیجه ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خانم فاطمه نورمحمدزاده، رشته آمار گرایش آمار ریاضی تحت عنوان تحلیل حساسیت مدل‌های کامپیوتری با استفاده از مدل جانشین بیزی که در تاریخ ۹۲/۱۱/۲۸ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می‌گردد:

<input type="checkbox"/> مردود	<input type="checkbox"/> دفاع مجدد	<input checked="" type="checkbox"/> قبول (با درجه: بسیار خوب امتیاز: ۱۸/۵۰)
--------------------------------	------------------------------------	---

۱- عالی (۲۰ - ۱۹)      ۲- بسیار خوب (۱۸/۹۹ - ۱۸)

۳- خوب (۱۷/۹۹ - ۱۶)      ۴- قابل قبول (۱۵/۹۹ - ۱۴)

۵- نمره کمتر از ۱۴ غیر قابل قبول

امضاء	مرتبه علمی	نام و نام خانوادگی	عضو هیأت داوران
	استادیار	دکتر داود شاهسونی	۱- استادراهنما
	دانشیار	احمد زاکتی	۲- نماینده شورای تحصیلات تکمیلی
	دانشیار	دکتر محمد آرشی	۳- استاد ممتحن
	استادیار	دکتر حسین باغیشنی	۴- استاد ممتحن

رئیس دانشکده: دکتر احمد زیره  
 امضاء:

تقدیم بہ خانوادہ عزیزم

## سپاس‌گزاری... .

سپاس خدای را که سخنوران، در ستودن او بمانند و شمارندگان، شمردن نعمت های او ندانند و کوشندگان، حق او را گزاردن نتوانند. و سلام و مورد بر محمد و خاندان پاک او، طاهران معصوم، هم آنان که وجودمان وامدار وجودشان است؛ و نفرین پیوسته بر دشمنان ایشان تا روز رستاخیز... بدون شک جایگاه و منزلت معلم، اجل از آن است که در مقام قدردانی از زحمات بی شائبه ی او، با زبان قاصر و دست ناتوان، چیزی بنگاریم.

اما از آنجایی که تجلیل از معلم، سپاس از انسانی است که هدف و غایت آفرینش را تامین می کند و سلامت امانت‌هایی را که به دستش سپرده‌اند، تضمین؛ بر حسب وظیفه و از باب ”من لم یشکر المنعم من المخلوقین لم یشکر الله عزّ و جلّ“: از پدر و مادر عزیزم این دو معلم بزرگوارم که همواره بر کوتاهی و درستی من، قلم عفو کشیده و کریمانه از کنار غفلت هایم گذشته‌اند و در تمام عرصه‌های زندگی یار و یآوری بی چشم‌داشت برای من بوده‌اند؛ از استاد با کمالات و شایسته؛ جناب آقای دکتر داود شاهسونی که در کمال سعه صدر، با حسن خلق و فروتنی، از هیچ کمکی در این عرصه بر من دریغ ننمودند و زحمت راهنمایی این پایان‌نامه را بر عهده گرفتند و از اساتید فرزانه و دلسوز؛ جنابان آقای دکتر محمد آرشی و آقای دکتر حسین باغیشنی که زحمت داوری این پایان‌نامه را متقبل شدند؛ کمال تشکر و قدردانی را دارم.

باشد که این خردترین، بخشی از زحمات آنان را سپاس گوید.

## تعمدنامه

اینجانب فاطمه نورمحمدزاده دانشجوی کارشناسی ارشد رشته ریاضی کاربردی دانشکده علوم ریاضی دانشگاه شاهرود، نویسنده پایان‌نامه با عنوان تحلیل حساسیت مدل‌های کامپیوتری با استفاده از مدل جانشین بیزی، تحت راهنمایی دکتر داود شاهسونی متعهد می‌شوم:

- تحقیقات در این پایان‌نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های دیگر پژوهش‌گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان‌نامه، تاکنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ‌جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه شاهرود متعلق دارد، و مقالات مستخرج با نام “ دانشگاه شاهرود “ یا “ Shahrood University “ به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به‌دست آوردن نتایج اصلی پایان‌نامه تاثیرگذار بوده‌اند، در مقالات مستخرج از پایان‌نامه رعایت می‌گردد.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا بافت‌های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده) شده است، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

فاطمه نورمحمدزاده

بممن ۱۳۹۲

## مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده) متعلق به دانشگاه شاهرود می‌باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان‌نامه بدون ذکر منبع مجاز نمی‌باشد.

## چکیده

مدل‌های کامپیوتری شامل ورودی‌های زیادی هستند که مقادیر واقعی آن‌ها نامعلوم هستند. در روش‌های تحلیل حساسیت اخیر برای به‌دست آوردن اثر عدم قطعیت ورودی‌ها روی خروجی مدل کامپیوتری، از روش مونت کارلو استفاده شده است. اغلب برای مدل‌های کامپیوتری که محاسبات زمان‌بر و پیچیده‌ای دارند، گرفتن نمونه‌ای به اندازه کافی بزرگ برای انجام یک تحلیل عدم قطعیت معنی‌دار، غیرممکن است. در این پایان‌نامه با ارایه مدل کریگینگ به عنوان یک مدل جانشین که برآورد مدل کامپیوتری بر مبنای استفاده از مدل‌های فرایند تصادفی گاوسی در یک مفهوم بیزی می‌باشد، به انجام تحلیل حساسیت می‌پردازیم. تمام نتایج و شاخص‌هایی که با تحلیل حساسیت مونت کارلو مدل کامپیوتری به‌دست می‌آید با مدل جانشین نیز ولی با یک نمونه خیلی کوچکتر حاصل می‌شوند. دقت مدل کریگینگ با یک تابع تحلیلی و یک مورد مطالعاتی نشان داده می‌شود.

کلمات کلیدی: مدل کامپیوتری، تحلیل حساسیت، مدل جانشین، فرایند گاوسی، کریگینگ، رهیافت

بیزی

## پیش‌گفتار

رفتار بسیاری از پدیده‌های طبیعی یا فرآیندهای فیزیکی، در زمینه‌های مختلف علوم و مهندسی، در قالب مدل‌های ریاضی بیان می‌شود. این مدل‌ها شامل دستگاه‌های پیچیده‌ای از معادلات دیفرانسیل معمولی یا جزئی هستند که در آن‌ها برخی متغیرها به عنوان "ورودی" مدل و یک یا چند متغیر نیز نقش "خروجی" مدل را ایفا می‌کنند. در واقع، مدل مذکور، رابطه بین متغیرهای ورودی ( $x$ ) و خروجی ( $Y$ ) را تحت تابع  $f(\cdot)$  به صورت  $Y = f(x)$  که فاقد شکل بسته بوده و به صورت صریح بیان نمی‌شود، مشخص می‌کند. به برنامه یا کد کامپیوتری که قادر به حل عددی این دستگاه معادلات باشد، "مدل کامپیوتری" و اجرای مکرر مدل با ورودی‌های مختلف را "آزمایش کامپیوتری" نامند. آزمایش‌های کامپیوتری در قیاس با آزمایش‌های فیزیکی دارای این تفاوت هستند که خروجی آزمایش‌های فیزیکی آغشته به خطاهای تصادفی است، در حالی که خروجی مدل‌های کامپیوتری قطعی است به این معنا که اجرای مکرر مدل با ورودی‌های همانند، خروجی یکسانی را تولید می‌کند. از آن‌جا که تعداد متغیرهای ورودی این‌گونه مدل‌ها بسیار زیاد است، تعیین و بررسی متغیرهای ورودی موثر بر خروجی، موضوعی آماري است که تحت عنوان "تحلیل حساسیت مدل‌های کامپیوتری" دامنه وسیعی از تحقیقات آماری را دربر گرفته است. روش‌های مختلفی برای تحلیل حساسیت مدل‌ها پیشنهاد شده‌اند که در هر یک، میزان اهمیت هر ورودی و اثر متقابل آن با سایر متغیرهای ورودی، توسط اندازه‌هایی به نام شاخص حساسیت بیان می‌گردند. یکی از این روش‌ها، تحلیل حساسیت واریانس-مبنا است که در آن اندازه‌هایی به نام‌های "شاخص حساسیت اصلی" و "شاخص حساسیت کل" معرفی می‌گردد. ساختار این دو اندازه حساسیت، مبتنی بر مفاهیم امید ریاضی شرطی و واریانس شرطی است. محاسبه شاخص‌های حساسیت، نیاز به اجراهای بسیار زیاد مدل کامپیوتری دارد. فقدان رابطه صریح برای  $f(\cdot)$  و طولانی بودن زمان اجرای مدل‌های کامپیوتری که گاه به زمانی بیش از یک ساعت نیاز دارد، برآورد این شاخص‌ها با روش انتگرال‌گیری مونت کارلویی را ضرورت می‌بخشد. اگرچه اجرای مدل‌های پیچیده برای درک طبیعت موضوع فرآیند مورد بررسی، سودمند است اما هزینه محاسباتی بالا در این‌گونه مدل‌ها چالشی است که گاهی اوقات موجب استفاده محدودتر از آن‌ها می‌شود. لذا برای مقابله با موضوع هزینه زمانی تحمیل شده، انگیزه



جایگزینی تابع پیچیده ریاضی با برآورد آن مطرح می‌شود. این جایگزینی، تحت عنوان «مدل جانشین» مطرح است که در آن، رفتار تابع اصلی توسط مدل‌های پیشرفته آماری یا یادگیری ماشین، مدل‌بندی می‌شود. این پایان‌نامه که به مطالعه تحلیل حساسیت مدل‌های کامپیوتری با استفاده از روش کریگینگ می‌پردازد، در چهار فصل زیر نگاشته شده است:

- در فصل اول، مدل کامپیوتری تعیینی و آزمایش کامپیوتری را تعریف می‌کنیم و به بیان نقش آمار در آزمایش‌های کامپیوتری می‌پردازیم.
- در فصل دوم، تحلیل حساسیت واریانس-مبنا برای مدل‌های تعیینی مورد بحث قرار گرفته است. در این جا شاخص‌های حساسیت، معرفی شده و برآورد آن‌ها با روش انتگرال‌گیری مونت کارلویی مطالعه می‌شود.
- در فصل سوم، خروجی مدل‌های تعیینی را به‌عنوان مصداقی از یک فرآیند تصادفی که شامل یک مدل رگرسیونی است در نظر گرفته و روش کریگینگ را برای تحلیل آن به‌عنوان مدل جانشین بیزی شرح داده‌ایم. در این فصل مدل کریگینگ معمولی و مدل کریگینگ عام معرفی و روابط مربوط به آن‌ها اثبات می‌شوند. سپس دیدگاه بیزی مدل کریگینگ بیان می‌شود.
- در فصل چهارم، عملکرد مدل کریگینگ را به‌عنوان یک مدل جانشین، برای یک تابع تحلیلی و همچنین یک مدل کامپیوتری واقعی به نام INCA که میزان بار نیتروژن ورودی به رودخانه را شبیه‌سازی می‌کند، بررسی می‌کنیم و تحلیل حساسیت آن‌ها را با این مدل جانشین انجام داده و شاخص‌های حساسیت به‌دست آمده توسط اجرای مدل جانشین را با مقادیر واقعی این شاخص‌ها مقایسه خواهیم نمود.

# فهرست مطالب

ز	لیست تصاویر
س	لیست جداول
۱	۱ تعاریف و مفاهیم اولیه
۱	۱.۱ مقدمه
۳	۲.۱ آزمایش کامپیوتری
۳	۳.۱ نقش آمار در آزمایش‌های کامپیوتری
۵	۴.۱ انتگرال‌گیری مونت کارلویی
۵	۵.۱ روش کمترین توان‌های دوم تعمیم‌یافته
۷	۲ تحلیل حساسیت واریانس-مبنا
۷	۱.۲ مقدمه
۸	۲.۲ پیشینه ریاضی تحلیل حساسیت
۸	۱.۲.۲ تجزیه واریانس
۱۰	۲.۲.۲ تجزیه واریانس و شاخص‌های حساسیت
۱۱	۳.۲.۲ شاخص‌های حساسیت اصلی و متقابل
۱۲	۴.۲.۲ شاخص حساسیت کل
۱۴	۳.۲ برآورد شاخص حساسیت اصلی
۱۵	۴.۲ برآورد شاخص حساسیت کل

۱۶	.....	هزینه‌های محاسباتی شاخص‌های حساسیت	۵.۲
۱۷		مدل جانشین بیزی: کریگینگ	۳
۱۷	.....	مقدمه	۱.۳
۱۸	.....	پیشینه مدل‌های جانشین	۲.۳
۲۰	.....	تاریخچه کریگینگ	۳.۳
۲۱	.....	مروری بر کاربردهای کریگینگ	۴.۳
۲۲	.....	کاربرد کریگینگ در مدل‌های جانشین	۵.۳
۲۳	.....	روش کریگینگ	۶.۳
۲۵	.....	پیش‌گوی کریگینگ	۱.۶.۳
۲۶	.....	محاسبه واریانس خطا در روش کریگینگ معمولی	۲.۶.۳
۲۹	.....	محاسبه پیش‌گو در روش کریگینگ عام	۳.۶.۳
۳۲	.....	واریانس کریگینگ عام	۴.۶.۳
۳۳	.....	دیدگاه بیزی	۷.۳
۳۴	.....	توسعه مدل بیزی	۱.۷.۳
۳۵	.....	پیش‌گوی بیزی	۸.۳
۳۷		عملکرد مدل جانشین کریگینگ بر تابع تحلیلی و مورد مطالعاتی	۴
۳۷	.....	تولید داده‌های شبیه‌سازی	۱.۴
۴۰	.....	معرفی تابع تحلیلی سبل	۲.۴
۴۲	.....	روش جانشین کریگینگ برای تابع تحلیلی سبل	۳.۴
۴۶	.....	معرفی اجمالی مدل واقعی $INCA - N$	۴.۴
۴۸	.....	معرفی ورودی‌ها و خروجی‌های مدل	۵.۴
۴۹	.....	منطقه مورد مطالعه	۶.۴
۵۰	.....	تحلیل حساسیت مدل $INCA - N$ با استفاده از مدل کریگینگ	۷.۴

۵۴	نتیجه‌گیری	۸.۴
۵۵	پیشنهادات	۹.۴
۵۷	نمادها و تعریف‌ها	آ
۵۷	فرآیندهای تصادفی گاوسی	۱.آ
۵۷	درون‌یابی	۲.آ
۵۸	ناهم‌واریانسی	۳.آ
۵۸	توزیع گامای معکوس	۴.آ
۵۹	توزیع نرمال-گاما معکوس	۵.آ
۶۰	تابع چگالی نرمال چند متغیره	۶.آ
۶۱	معکوس ماتریس بلوکی	۷.آ
۶۳	تغییرنگار	۸.آ
۶۴	اثر قطعه‌ای	۹.آ
۶۵	ب کد برنامه $R$ برای اجرای روش کریگینگ و برآوردهای مونت کارلویی در فصل چهارم	
۷۹	مراجع	

## لیست تصاویر

۱۰۴	نمودار پراکنش ۲۵ نمونه تصادفی تولیدشده توسط نمونه‌گیری تصادفی ساده . . . . .	۳۹
۲۰۴	نمودار پراکنش ۲۵ نمونه $LHS(25, 2)$ . . . . .	۳۹
۳۰۴	نمودار تابع سبل در حالت یک‌متغیره . . . . .	۴۰
۴۰۴	نمودار تابع سبل در حالت دو متغیره ( $a_1 = 0, a_2 = 1$ ) . . . . .	۴۱
۵۰۴	راست: نمودار پراکنش داده‌های مدل‌ساز در برابر مقادیر پیش‌گویی شده آن‌ها، چپ:	
۴۳	نمودار پراکنش داده‌های آزمون در برابر مقادیر پیش‌گویی شده آن‌ها . . . . .	۴۳
۶۰۴	اعتبارسنجی متقابل $LOO$ برای تابع سبل . . . . .	۴۴
۷۰۴	نمودار شاخص حساسیت مرتبه اول متغیرهای تابع سبل . . . . .	۴۶
۸۰۴	نمودار شاخص حساسیت کل تابع سبل . . . . .	۴۷
۹۰۴	چرخه نیتروژن در طبیعت . . . . .	۴۸
۱۰۰۴	ساختار شماتیکی مدل $INCA - N$ . . . . .	۴۹
۱۱۰۴	ساختار انجام برآورد شاخص‌های حساسیت اصلی برای مدل $INCA - N$ به روش	
۵۱	کریگینگ . . . . .	۵۱
۱۲۰۴	نمودار مقادیر پیش‌گویی شده در مقابل مقادیر واقعی برای ۲۵۰ داده مدل‌ساز (اولین	
	نمودار از راست) و داده‌های آزمون برای حجم‌های ۲۰، ۲۰۰ و ۲۰۰۰ (به ترتیب	
۵۲	نمودارهای اول، دوم و سوم از چپ) . . . . .	۵۲
۱۳۰۴	اعتبارسنجی متقابل $LOO$ برای مدل $INCA$ . . . . .	۵۳

۱۴.۴ نمودار شاخص‌های حساسیت مرتبه اول مدل  $INCA - N$ ، توسط اجرای مدل

واقعی و روش کریگینگ ..... ۵۴

# لیست جداول

۴۱	میزان اهمیت و مقادیر حقیقی ورودی‌های تابع سبل	۱.۴
۴۲	برآورد پارامترهای مدل (۲.۳) با روش کریگینگ برای تابع سبل	۲.۴
۴۵	مقادیر واقعی و مقادیر برآوردشده اثر اصلی	۳.۴
۴۶	مقادیر واقعی و مقادیر برآوردشده اثر کل	۴.۴
۵۰	مشخصات عامل‌های ورودی مدل $INCA - N$	۵.۴
۵۱	برآورد پارامترهای مدل کریگینگ تابع $INCA$	۶.۴
۵۳	برآورد شاخص‌های حساسیت اصلی متغیرهای ورودی مدل $INCA$	۷.۴

# فصل ۱

## تعاریف و مفاهیم اولیه

### ۱.۱ مقدمه

در بسیاری از علوم، استفاده از شبیه‌سازی برای توصیف و تحلیل سیستم‌های پیچیده طبیعی، به منظور دریافت اطلاعات و آگاهی‌های مربوط به این سیستم‌ها، مرسوم است. پیچیدگی این سیستم‌ها، تحلیل روش‌های مدل‌بندی را امکان‌ناپذیر می‌سازد. گرچه یک مدل شبیه‌سازی فقط خلاصه‌ای از سیستم پیچیده واقعی است، اما شامل جزئیات و قیدهایی بسیار زیادی است که آن را بسیار مشابه با سیستم واقعی می‌سازد و این خود مفید بودن این شبیه‌سازی‌ها را نشان می‌دهد.

اغلب مدل‌های شبیه‌سازی از لحاظ محاسباتی هزینه‌بر<sup>۱</sup> هستند و چون منبع مدل‌های شبیه‌ساز پدیده‌های واقعی طبیعت اطراف ماست، به‌دست آوردن خروجی‌های شبیه‌سازی‌شده کافی برای رفع نیازهای اجرایی تحلیل حساسیت و یافتن ورودی‌های بهینه، بسیار مشکل است. بنابراین استفاده از برآورد مدل‌های شبیه‌سازی مورد نیاز است تا با استفاده از مدل برآوردشده، خروجی‌های احتمالی برای ورودی‌هایی که هنوز شبیه‌سازی نشده‌اند، برآورد گردد. از مقادیر خروجی برآوردشده، برای تحلیل حساسیت، بهینه‌سازی و تصمیم‌گیری استفاده می‌شود.

برآورد مدل شبیه‌سازی، تحت عنوان ”مدل جانشین”<sup>۲</sup> بحث و بررسی می‌شود. مدل‌بندی جانشین، عبارتست از تحلیل، ساخت و توسعه چارچوب‌ها، قوانین، ساختارها، مدل‌ها و نظریه‌هایی که برای

---

<sup>۱</sup> هزینه در این جا زمان تلقی می‌شود که گاه یک اجرای مدل به چند دقیقه یا چند ساعت زمان نیاز دارد.

<sup>۲</sup> Surrogate model



مدل‌بندی رده‌ای از مسایل از قبل تعریف‌شده، مفید هستند (لیو<sup>۱</sup>، ۲۰۰۹). مدل‌بندی جانشین نخستین بار توسط گروهی از عصب‌شناسان در اوایل قرن بیستم توسعه یافت. کلیجن<sup>۲</sup> (۱۹۹۸) چگونگی استفاده از مدل‌بندی جانشین را برای اجرای تحلیل حساسیت، بهینه‌سازی و اعتباربخشی بحث کرد.

در مبحث مدل‌بندی شبیه‌سازی‌های کامپیوتری (آزمایش‌های کامپیوتری)، کلیجن و همکاران (۲۰۰۵) یک مدل جانشین را به عنوان تقریبی از تابع ورودی و خروجی‌های واقعی، که به طور ضمنی توسط یک مدل شبیه‌سازی داده شده بود، تعریف کردند. ایده آن‌ها این بود که تحلیل مدل جانشین خیلی آسان‌تر از تحلیل مدل شبیه‌سازی است.

”کریگینگ<sup>۳</sup>” یک روش مدل‌بندی جانشین زمین‌آمار<sup>۴</sup> است که عملکرد آن به نحو مطلوبی به اثبات رسیده است. به علت ویژگی درون‌یابی دقیق کریگینگ، از این روش در برخی از حوزه‌ها استفاده شده و نتایج موفقیت‌آمیزی کسب شده‌اند. کریگینگ از نام کریگ<sup>۵</sup> یک مهندس معدن آفریقای جنوبی گرفته شده است. در دهه ۱۹۵۰ کریگ اولین بار این روش را برای پیش‌گویی تهنشست‌های طلا به‌کار برد (کرس<sup>۶</sup>، ۱۹۹۳ و کریگ، ۱۹۵۱). پیش‌گویی کریگینگ، بهترین پیش‌گویی نااریب خطی<sup>۷</sup> (BLUP) است. با شناسایی و نمایش ریاضی رابطه تابعی در نقاط معلوم مشاهده‌شده، می‌توان مقادیر نقاط نامعلوم تابع را پیش‌گویی کرد. در این پایان‌نامه، برای انجام تحلیل حساسیت، کریگینگ به عنوان یک مدل جانشین با دقت درون‌یابی بالا، برای ساختن رابطه تابعی از ورودی و خروجی‌های شبیه‌سازی‌شده، پیشنهاد شده است. مدل‌های شبیه‌سازی مد نظر این پایان‌نامه از نظر محاسباتی هزینه‌بر و تعیینی<sup>۸</sup> (قطعی) هستند، به این معنا که اجرای مکرر مدل با ورودی‌های یکسان، خروجی‌های یکسانی را ارائه می‌دهد.

---

<sup>۱</sup>Liu

<sup>۲</sup>Kleijnen

<sup>۳</sup>Kriging

<sup>۴</sup>Geostatistics

<sup>۵</sup>Krige

<sup>۶</sup>Cressie

<sup>۷</sup>Best linear unbiased predictor

<sup>۸</sup>Deterministic

## ۲.۱ آزمایش کامپیوتری

توصیف جریان هوا روی بال هواپیما، سوختن گازها در یک شعله، رفتار یک ساختار فلز تحت فشار، ایمنی یک راکتور هسته‌ای و مثال‌های متعدد دیگر، با شبیه‌سازی‌های کامپیوتری قابل انجام هستند. شبیه‌سازی سیستم‌های واقعی و پدیده‌های طبیعی، به منظور کسب اطلاع از این پدیده‌ها می‌باشد. آزمایش کامپیوتری<sup>۱</sup> یا آزمایش شبیه‌سازی، آزمایشی برای مطالعه یک شبیه‌سازی کامپیوتری می‌باشد. این آزمایش‌ها شامل محاسبات فیزیکی، شیمیایی، زیستی و سایر است. شبیه‌سازی‌های کامپیوتری پیچیدگی‌های بالایی دارند و عوامل بسیاری بر آن‌ها تاثیرگذارند. مدل ریاضی این شبیه‌سازها شامل جزئیات و محدودیت‌های بسیاری است که از واقعیت مربوطه نشات گرفته‌اند و این سبب نزدیکی هر چه بیشتر شبیه‌سازها با پدیده‌های واقعی می‌شود و به تبع آن، استفاده از شبیه‌سازها را مفید و مطلوب می‌سازد.

دیاکنیس<sup>۲</sup> (۱۹۸۸) نشان داد که در یک آزمایش کامپیوتری، مقدار تابع  $f(\cdot)$  با هر مجموعه از مقادیر ورودی  $\mathbf{x} = (x_1, \dots, x_d)$  قابل محاسبه است، ولی مکانیسمی که مدل به ورودی‌هایش پاسخ می‌گوید (ضابطه  $f(\cdot)$ ) به‌طور صریح مشخص نیست همچنین اطلاع از پیوستگی، کران‌داری، تک‌مدی بودن، محل قرارگیری بهینه‌های تابع و نیز متقارن بودن آن برای ما قابل حصول نیست. برنامه یا کد کامپیوتری که بتواند برای هر مجموعه از ورودی‌ها، خروجی مدل را تعیین کند، ”مدل کامپیوتری” و اجرای مدل با ورودی‌های متفاوت را ”آزمایش کامپیوتری” گویند. ویژگی مهم آزمایش کامپیوتری، تعیینی بودن آن است. به سبب این ویژگی‌ست که در آزمایش‌های کامپیوتری، خطای تصادفی وجود ندارد و خروجی آزمایش‌های کامپیوتری برخلاف آزمایش‌های فیزیکی، تنها تابعی از ورودی‌های مدل است.

## ۳.۱ نقش آمار در آزمایش‌های کامپیوتری

آزمایش‌های کامپیوتری تعیینی ذاتا متفاوت از آزمایش‌های فیزیکی هستند. تغییرپذیری در واحدهای آزمایش فیزیکی سبب تولید خطای تصادفی در این آزمایش‌ها می‌شود و برای کاهش اثر خطای تصادفی

<sup>۱</sup>Computer experiment

<sup>۲</sup>Diaconis

و به دست آوردن برآوردهای معتبر، اعمالی نظیر بلوک‌بندی، تکرار و تصادفی‌سازی انجام می‌شود. فقدان خطای تصادفی مستقل در آزمایش‌های کامپیوتری، تفاوتی مهم بین این دو نوع آزمایش ایجاد کرده است. در آزمایش‌های کامپیوتری، مفاهیم کلاسیکی مانند واحد آزمایشی، بلوک‌بندی، تکرار و تصادفی‌سازی موضوعیت ندارند. چون مبدا ورود آمار، خطاهای تصادفی است، شاید گمان رود که آمار در مبحث آزمایش‌های کامپیوتری نقشی ندارد. ولی در قالب موارد زیر به بیان نقش آمار می‌پردازیم:

۱. طرح آزمایش مدل‌های کامپیوتری: اگر قرار باشد آزمایشی با حداکثر کارایی انجام شود، باید شیوه علمی برای طرح آن آزمایش به کار رود. منظور از طرح آماری آزمایش‌ها، طرح‌ریزی فرایندی است که در آن با روش‌های آماری، داده‌های مناسب جمع‌آوری می‌شود و با استنتاج‌های معتبر بر روی این داده‌ها، تحلیل آماری انجام می‌گیرد. اگر بخواهیم نتایج مفید از داده‌ها استخراج کنیم، روش‌های آماری در طرح آزمایش ضروری است. گرچه در آزمایش‌های کامپیوتری، خطای تصادفی نداریم اما از آن‌جا که درک فرایند مورد مطالعه نیازمند اجزای متعدد مدل است، با عنایت به عامل هزینه‌گراف محاسباتی، یک طرح آزمایش مناسب می‌تواند درک سیستم را در زمان بسیار کوتاه میسر ساخته و کمک شایانی به مدیریت بحران نماید (شاهسونی و گریموال،<sup>۱</sup> ۲۰۰۹، سانتنر و همکاران، ۲۰۰۳).

۲. تحلیل حساسیت خروجی مدل: از آن‌جا که تعداد متغیرهای ورودی این‌گونه مدل‌ها بسیار زیاد است، تعیین و بررسی متغیرهای ورودی موثر بر خروجی، موضوعی آماری است که تحت عنوان "تحلیل حساسیت مدل‌های کامپیوتری" دامنه وسیعی از تحقیقات آماری را دربر گرفته است. (سالتلی<sup>۲</sup> و همکاران، ۲۰۰۹)

۳. مدل‌جانشین: هزینه اجرای مدل‌های کامپیوتری و عدم درک درست فرایند تحت بررسی، انگیزه‌ای است که بتوان به جای اجرای مستقیم مدل و با در دست داشتن یک طرح آزمایش مناسب، از مدل‌های رگرسیونی ناپارامتری و یادگیری ماشین<sup>۳</sup> به عنوان جانشین آن‌ها استفاده کرد.

<sup>۱</sup>Grimvall

<sup>۲</sup>Saltelli

<sup>۳</sup>Machine learning

۴. عدم قطعیت: پیش‌گویی توسط مدل‌ها و نیز برآورد شاخص‌های حساسیت، با عدم قطعیت همراه است. تحلیل عدم قطعیت، می‌تواند توسط یک مدل آماری صورت پذیرد.

از بین موارد بالا تاکید ما بیشتر بر موارد ۲ و ۳ است، که به ترتیب در بخش‌های ۲ و ۳ به آن‌ها می‌پردازیم. در واقع از مورد ۳ نیز برای تحلیل حساسیت استفاده می‌کنیم، اما به جای تحلیل  $f$  به تحلیل مدل جانشین  $\hat{f}$  می‌پردازیم. برای مورد ۱ نیز از طرح آزمایش ابرمکعب لاتین<sup>۱</sup> (LHD) استفاده می‌شود که توضیح بیشتر آن در فصل ۴ آمده است.

## ۴.۱ انتگرال‌گیری مونت کارلویی

انتگرال‌گیری مونت کارلو شکل ویژه‌ای از روش مونت کارلو است، که از اعداد تصادفی برای انتگرال‌گیری استفاده می‌کند. مساله کلی، محاسبه انتگرال

$$I = E_f[h(X)] = \int_{\mathcal{X}} h(x)f(x)dx,$$

است که در آن  $\mathcal{X}$  یک یا چند بعدی است.  $f$  یک تابع چگالی، دارای فرم بسته یا به‌طور جزئی بسته است، و  $h$  یک تابع می‌باشد. با داشتن نمونه  $x_1, \dots, x_n$  از توزیع  $f(x)$ ، تقریب مونت کارلوی  $I$  به صورت  $\hat{I} = \frac{1}{n} \sum_{i=1}^n h(x_i)$  محاسبه می‌شود، به طوری که بنابر قانون قوی اعداد بزرگ، هنگامی که  $n$  به بینهایت میل می‌کند،  $\hat{I}$  به مقدار واقعی انتگرال میل خواهد کرد.

## ۵.۱ روش کمترین توان‌های دوم تعمیم‌یافته

با در نظر گرفتن پذیره‌های زیر برای مدل  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ ، برآورد ضرایب رگرسیونی  $\boldsymbol{\beta}$  با روش کمترین توان‌های دوم تعمیم‌یافته انجام می‌گیرد.

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}, \quad \text{Cov}(\mathbf{y}) = \sigma^2 \mathbf{V},$$

که در آن  $\mathbf{X}$  دارای رتبه کامل و  $\mathbf{V}$  یک ماتریس معین مثبت معلوم است.

در قضیه زیر برآوردگرهای  $\boldsymbol{\beta}$  و  $\sigma^2$  را ارائه می‌دهیم.

<sup>۱</sup>Latin hypercube design

قضیه ۱.۵.۱. (رنچر، ۱۳۸۲) فرض کنید

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}, \quad \text{Cov}(\mathbf{y}) = \text{Cov}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{V},$$

که در آن  $\mathbf{X}$  یک ماتریس با رتبه کامل و  $\mathbf{V}$  یک ماتریس معین مثبت معلوم است. برای این مدل، نتایج زیر برقرار هستند:

(۱) بهترین برآوردگر نااریب خطی (BLUE) برای  $\boldsymbol{\beta}$  عبارت است از:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y}. \quad (1.1)$$

(۲) ماتریس کوواریانس برای  $\hat{\boldsymbol{\beta}}$  عبارت است از:

$$\text{Cov}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}.$$

(۳) یک برآوردگر نااریب  $\sigma^2$  عبارت است از:

$$s^2 = \frac{(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})}{n - k - 1} \\ = \frac{\mathbf{y}^T[\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}]\mathbf{y}}{n - k - 1},$$

که در آن  $\mathbf{X}^T$  نشان‌دهنده ترانهاد  $\mathbf{X}$  و  $\hat{\boldsymbol{\beta}}$  در (۱.۱) داده شده است.

قضیه ۲.۵.۱. (رنچر، ۱۳۸۲) اگر  $\mathbf{y}$  دارای توزیع  $N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{V})$  باشد، که در آن  $\mathbf{X}$  یک ماتریس

$(k+1) \times n$  با رتبه  $k+1$  و  $\mathbf{V}$  یک ماتریس معین مثبت معلوم است، برآوردگرهای درست‌نمایی

ماکزیم  $\boldsymbol{\beta}$  و  $\sigma^2$  عبارتند از:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y},$$

$$\hat{\sigma}^2 = \frac{1}{n}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}).$$

## فصل ۲

# تحلیل حساسیت واریانس-مبنا

### ۱.۲ مقدمه

در طراحی مهندسی، برای کسب اطلاعات بیشتر از رفتار مدل‌های پیچیده، از شاخص‌های حساسیت استفاده می‌شود تا بتوان با کمک این شاخص‌ها تصمیم‌های آگاهانه‌تری اتخاذ نمود. تحلیل حساسیت، مطالعه چگونگی تخصیص عدم قطعیت خروجی مدل به منابع مختلف عدم قطعیت ورودی‌های مدل است (سالتلی و همکاران، ۲۰۰۸). تحلیل حساسیت، روش‌های مختلفی دارد که در هر یک، میزان اهمیت هر ورودی و اثر متقابل آن با سایر متغیرهای ورودی، توسط اندازه‌هایی به نام شاخص حساسیت بیان می‌گردند. یکی از این روش‌ها به نام "تحلیل حساسیت واریانس-مبنا"، معیارهای لازم برای تحلیل مناسب را داراست. این نوع از تحلیل حساسیت، امکان مقابله با اثر شکل و مقیاس، محاسبه میانگین چندبعدی و امکان بررسی گروهی متغیرها را داراست. در تحلیل حساسیت واریانس-مبنا دو اندازه حساسیت به نام‌های "شاخص حساسیت اصلی" و "شاخص حساسیت کل" معرفی می‌گردند. ساختار این دو اندازه حساسیت، مبتنی بر مفاهیم امید ریاضی شرطی و واریانس شرطی است.

روش‌های واریانس-مبنا، تاریخچه‌ای طولانی در تحلیل حساسیت دارند. این روش‌ها با اجراهای فوریه در دهه ۷۰ شروع شدند و مرحله مهمی از کار سبل<sup>۱</sup> را به خود اختصاص داده‌اند. شاخص‌های حساسیت کل، توسط هما<sup>۲</sup> و سالتلی (۱۹۹۶) معرفی شده‌اند. از جمله روش‌های واریانس-مبنا برای

---

<sup>۱</sup>Sobol

<sup>۲</sup>Homma

تحلیل حساسیت عبارتند از: آزمون حساسیت بسامدی فوریه<sup>۱</sup> (FAST) (سالتلی و همکاران، ۱۹۹۹)، نرخ همبستگی (مک کی<sup>۲</sup>، و همکاران ۱۹۹۹)، اندازه‌های اهمیت مختلف (هما و سالتلی، ۱۹۹۶) و شاخص‌های اثر کل سبل (سبل، ۱۹۹۳). در این روش‌ها واریانس کل خروجی به مولفه‌هایی تجزیه می‌شود که منابع مختلف تغییرات ورودی را در بر دارد و سپس شاخص‌های حساسیت به صورت نسبت واریانس جزئی بر واریانس کل خروجی به دست می‌آید که واریانس جزئی در واقع تاثیر متغیرهای مورد نظر بر واریانس  $f(\mathbf{X})$  می‌باشد. با این‌که به دست آوردن خروجی‌های نمونه از لحاظ محاسباتی زمان بر هستند، با این حال این روش‌ها برای انواع توابع توسعه یافته‌اند. از آن‌جا که این روش‌ها نیازمند تعداد نمونه‌های زیاد و استفاده از روش‌های عددی هستند، به کارگیری نمونه‌گیری مونت کارلو (سبل، ۱۹۹۳) و نمونه‌گیری شبکه‌ای (مک کی و همکاران، ۱۹۹۹؛ سالتلی و همکاران، ۱۹۹۹) توصیه شده‌اند. مطالب این فصل از مراجع سالتلی (۲۰۰۲)، جین<sup>۳</sup> و همکاران (۲۰۰۴)، سالتلی و همکاران (۲۰۰۸) و سالتلی و همکاران (۲۰۱۰) استخراج شده است.

## ۲.۲ پیشینه ریاضی تحلیل حساسیت

در این بخش، ابتدا مفهوم تجزیه واریانس<sup>۴</sup> (ANOVA) را که بنیان ریاضی شاخص‌های حساسیت در تحلیل حساسیت واریانس-مبنا است، بیان می‌کنیم. سپس شاخص‌های حساسیت را معرفی می‌کنیم و یک مثال برای توضیح بیشتر تحلیل حساسیت فراموضعی<sup>۵</sup> (GSA) ارائه می‌دهیم.

### ۱.۲.۲ تجزیه واریانس

در مفهوم تحلیل حساسیت واریانس-مبنا، تابع  $f(\cdot)$  به تابع‌های<sup>۶</sup> تحلیل واریانس تجزیه می‌شود (ANOVA) (اون<sup>۷</sup>، ۱۹۹۲، سبل، ۱۹۹۳) که مرتبه عبارات آن در حال افزایش است و در واقع

<sup>۱</sup>Fourier Amplitude Sensitivity Test

<sup>۲</sup>McKay

<sup>۳</sup>Jin

<sup>۴</sup>Analysis of Variance

<sup>۵</sup>Global Sensitivity Analysis

<sup>۶</sup>functional

<sup>۷</sup>Owen

تجزیه تابع  $f(\cdot)$  بر حسب توابعی با ابعاد افزایشی<sup>۱</sup> (HDMR) است.

$$f(x_1, x_2, \dots, x_d) = f_0 + \sum_{i=1}^d \phi_i(x_i) + \sum_{i=1}^d \sum_{j=i+1}^d \phi_{ij}(x_i, x_j) + \dots + \phi_{12\dots d}(x_1, x_2, \dots, x_d) \quad (1.2)$$

فرض کنید  $x_i (i = 1, \dots, d)$  متغیرهای تصادفی مستقل با توابع چگالی احتمال  $p_i(x_i)$  باشند. عبارت  $f_0$ ، میانگین  $f(\mathbf{x})$  است، یعنی

$$f_0 = E(Y) = \int f(\mathbf{x}) \prod_{i=1}^d [p_i(x_i) dx_i].$$

باید توجه شود که فرض استقلال متغیرها که در این پایان نامه استفاده شده است، برای سهولت محاسبات و استخراج روش پیشنهادی می باشد. این فرض در خیلی از کاربردهای عملی به صورت سهیم کردن متغیرهای تصادفی چندگانه که اغلب از منابع متفاوت ولی مستقل می آیند، برقرار است. چون در عمل توزیع های احتمالی حقیقی اغلب نامعلوم هستند یا نمی توان آن ها را به طور دقیق تعیین کرد، فرض استقلال، مزایای محاسباتی مهمی را فراهم می کند.

عبارت  $\phi_i(x_i)$  مبین اثر  $x_i$  است که با متوسط گیری خروجی روی همه متغیرها به جز  $x_i$  و کاستن

ثابت  $f_0$  به دست می آید. یعنی

$$\phi_i(x_i) = \underbrace{\int f(\mathbf{x}) \prod_{j \neq i} [p_j(x_j) dx_j]}_{E(Y|X_i)} - f_0.$$

سرانجام عبارت  $\phi_{ij}(x_i, x_j)$  مبین اثر توأم متغیرهای  $x_i$  و  $x_j$  است که با متوسط گیری خروجی روی

همه متغیرها به جز این دو متغیر و کاستن اثرات تکی آن ها و ثابت  $f_0$  به دست می آید. یعنی

$$\phi_{ij}(x_i, x_j) = \underbrace{\int f(\mathbf{x}) \prod_{l \neq i, j} [p_l(x_l) dx_l]}_{E(Y|X_i, X_j)} - \phi_i(x_i) - \phi_j(x_j) - f_0.$$

در کل عبارت  $\phi_{i_1 \dots i_s}(x_{i_1}, \dots, x_{i_s})$  مبین اثر توأم  $s$  متغیر است:

$$\begin{aligned} \phi_{i_1 \dots i_s}(x_{i_1}, \dots, x_{i_s}) &= \underbrace{\int f(\mathbf{x}) \prod_{j \neq i_1, \dots, i_s} [p_j(x_j) dx_j]}_{E(Y|X_{i_1}, X_{i_2}, \dots, X_{i_s})} \\ &\quad - \sum_{l=1}^{s-1} \sum_{j_1, \dots, j_l \in (i_1, \dots, i_s)} \phi_{j_1 \dots j_l}(x_{j_1}, \dots, x_{j_l}) - f_0, \end{aligned}$$

<sup>۱</sup>High Dimensional Model Representation



که  $j_1 < j_2 < \dots < j_l$ .

## ۲.۲.۲ تجزیه واریانس و شاخص‌های حساسیت

معادله (۱.۲) را به توان دو می‌رسانیم و از طرفین انتگرال می‌گیریم. چون عبارت  $\phi_{i_1, \dots, i_s}$  دارای میانگین صفر است، واریانس  $Y = f(\mathbf{x})$  که با  $V$  نشان می‌دهیم به صورت مجموع واریانس‌های

$\phi_{i_1, \dots, i_s}$  (واریانس‌های جزئی) که با  $V_{i_1, \dots, i_s}$  نشان می‌دهیم، بیان می‌شود:

$$V = \sum_i V_i + \sum_i \sum_j V_{ij} + \dots + V_{1\dots d}, \quad (2.2)$$

که در آن

$$V = Var\{f(\mathbf{x})\} = \int f^2(\mathbf{x})p(\mathbf{x})d\mathbf{x} - f^2, \quad (3.2)$$

و

$$V_{i_1, \dots, i_s} = Var[\phi_{i_1, \dots, i_s}(X_{i_1}, \dots, X_{i_s})] = \int \phi_{i_1, \dots, i_s}^2(X_{i_1}, \dots, X_{i_s}) \prod_{j=i_1}^{i_s} [p_j(X_j)dX_j].$$

در معادله (۳.۲)، برداری از متغیرهاست. این معادله شامل انتگرال چندبعدی حاصل ضرب توان دوم تابع  $f(\cdot)$  در توابع چگالی است. معادله (۲.۲)، تغییرپذیری خروجی  $f(\cdot)$  را به بخش‌های مجزا برای هر ورودی و ورودی‌های توأم تجزیه می‌کند. شاخص حساسیت در حالت کلی به صورت نسبت واریانس جزئی بر واریانس کل تعریف می‌شود:

$$S_{i_1, \dots, i_s} = V_{i_1, \dots, i_s} / V.$$

از آنجا که  $0 \leq S_{i_1, \dots, i_s} \leq 1$  است، لذا،  $V_{i_1, \dots, i_s} \leq V$ .

## ۳.۲.۲ شاخص‌های حساسیت اصلی و متقابل

بر مبنای روابط بالا، داریم:

$$S_i = \frac{V_i}{V} = \frac{Var[\phi_i(X_i)]}{Var(Y)}$$

$$= \frac{Var \left[ \int f(\mathbf{x}) \left[ \prod_{j \neq i} p_j(x_j) dx_j \right] - \overbrace{f_0}^{E(Y)} \right]}{Var(Y)} = \frac{Var[E(Y | X_i)]}{Var(Y)},$$

و

$$S_{ij} = \frac{V_{ij}}{V} = \frac{Var[\phi_{ij}(X_i, X_j)]}{Var(Y)}$$

$$= \frac{Var \left[ \int f(\mathbf{x}) \left[ \prod_{l \neq i, j} p_l(x_l) dx_l \right] - \phi_i(X_i) - \phi_j(X_j) - f_0 \right]}{Var(Y)}$$

$$= \frac{Var[E(Y | X_i, X_j)]}{Var(Y)} - S_i - S_j.$$

که در آن،  $S_i$  را به اسامی مختلفی چون شاخص حساسیت اصلی<sup>۱</sup> (MSI)، اثر اصلی یا شاخص حساسیت مرتبه اول  $x_i$  نامند. همچنین  $S_{ij}$ ، اثر متقابل مرتبه دوم یا شاخص حساسیت متقابل<sup>۲</sup> (ISI) دو متغیر  $x_i$  و  $x_j$  نامیده می‌شود. به همین ترتیب، شاخص حساسیت متقابل متغیرهای  $x_{i_1}, x_{i_2}, \dots, x_{i_s}$  نام دارد. از معادله (۲.۲) به آسانی می‌توان دریافت که مجموع همه شاخص‌های حساسیت مساوی با ۱ است، یعنی

$$\sum_{s=1}^d \sum_{i_1 < \dots < i_s} S_{i_1 \dots i_s} = \sum S_i + \sum_{i_1 < i_2} S_{i_1 i_2} + \dots + S_{1 \dots d} = 1.$$

مجموع اثرات مرتبه اول تمام متغیرهای ورودی هر مدل همواره کمتر یا مساوی با یک است. اگر  $\sum_i S_i = 1$  در این صورت مدل موردنظر را جمع<sup>۳</sup> نامیده که در آن اثرات متقابل بین متغیرها وجود ندارد. پس اگر  $\sum_i S_i < 1$ ، نشان از وجود اثر متقابل بین متغیرهای ورودی و غیرجمع بودن مدل است.

عبارت  $1 - \sum_i S_i$ ، میزان غیرجمع بودن مدل را اندازه‌گیری می‌کند. در واقع این عبارت مبین

<sup>۱</sup>Main Sensitivity Index

<sup>۲</sup>Interaction Sensitivity Index

<sup>۳</sup>Additive

کسری از واریانس است که با حضور تکی متغیرهای ورودی، قابل بیان نیست. این کسر می‌تواند شامل اثرات متقابل مرتبه دوم و مراتب بالاتر باشد.

### ۴.۲.۲ شاخص حساسیت کل

هما و سالتلی (۱۹۹۶، ۲۰۰۲) شاخصی را تحت عنوان شاخص حساسیت کل یا اثر کل معرفی نمودند که بتواند علاوه بر کاهش هزینه محاسباتی، حاوی اطلاعات مفیدی در خصوص تقابل اثرات باشد. این شاخص با نماد  $S^t$  نشان داده می‌شود. اثر کل هر متغیر، شامل اثر اصلی آن و اثرات متقابل آن با دیگر متغیرها است. در واقع اثر کل، مجموع اثر اصلی و تمام اثراتی است که متغیر مورد نظر را درگیر می‌کند:

$$S_i^t = S_i + \widehat{S}_{i,\sim i} = 1 - \widehat{S}_{\sim i},$$

که  $\widehat{S}_{i,\sim i}$  مجموعه همه  $S_{i_1, \dots, i_s}$ ها است که زیروند آن شامل  $i$  و حداقل یک اندیس از  $(1, \dots, i-1, i+1, \dots, d)$  می‌باشد. یعنی  $\widehat{S}_{\sim i}$  مجموع همه  $S_{i_1, \dots, i_s}$ ها است که شامل زیروند  $i$  نیست. به عنوان مثال، برای یک مدل سه متغیره، اثر کل  $x_1$  یعنی  $S_1^t$  به صورت زیر می‌باشد

$$S_1^t = S_1 + S_{1,2} + S_{1,3} + S_{1,2,3} = 1 - (S_2 + S_3 + S_{2,3}).$$

به منظور ارایه یک اندازه برای اثر کل، واریانس خروجی مدل را به نوعی دیگر و به صورت تجزیه آن به شرط ثابت بودن تمام متغیرهای ورودی به جز  $x_i$ ، یعنی  $\mathbf{X}_{\sim i}$ ، به صورت زیر در نظر می‌گیریم:

$$V(Y) = V(E(Y | \mathbf{X}_{\sim i})) + E(V(Y | \mathbf{X}_{\sim i})).$$

از آنجا که  $V(E(Y | \mathbf{X}_{\sim i}))$  اثر مرتبه اول  $\mathbf{X}_{\sim i}$  است، لذا

$$E(V(Y | \mathbf{X}_{\sim i})) = V(Y) - V(E(Y | \mathbf{X}_{\sim i})),$$

دربرگیرنده تمام اثراتی است که به نوعی، متغیر  $x_i$  در آن مشارکت دارد. با تقسیم طرفین این رابطه بر  $V(Y)$ ، شاخص حساسیت کل به صورت زیر معرفی می‌گردد:

$$S_i^t = \frac{E(V(Y | \mathbf{X}_{\sim i}))}{V(Y)} = 1 - \frac{V(E(Y | \mathbf{X}_{\sim i}))}{V(Y)}.$$

مثال ۱.۲.۲. مدل  $Y = \sum_{i=1}^8 X_i$  با ۸ متغیر ورودی را در نظر بگیرید، که همگی متغیرها دارای

توزیع‌های نرمال به صورت زیر هستند:

$$X_1 \sim N(0, 1), \quad X_2 \sim N(2, 4), \quad X_3 \sim N(1, 9), \quad X_4 \sim N(1, 25),$$

$$X_5 \sim N(3, 1), \quad X_6 \sim N(4, 1), \quad X_7 \sim N(4, 4), \quad X_8 \sim N(5, 25).$$

اثرات اصلی متغیرهای ورودی را محاسبه می‌کنیم.

$$Var(Y) = \sum_{i=1}^8 Var(X_i) = 1 + 4 + 9 + 25 + 1 + 1 + 4 + 25 = 70,$$

$$E(Y | X_1) = X_1 + 2 + 1 + 1 + 3 + 4 + 4 + 5 = X_1 + 20,$$

$$Var[E(Y | X_1)] = Var[X_1 + 20] = Var(X_1) = 1,$$

$$S_1 = \frac{Var[E(Y | X_1)]}{Var(Y)} = \frac{1}{70} \cong 0.014, \quad S_2 = \frac{Var[E(Y | X_2)]}{Var(Y)} = \frac{4}{70} \cong 0.057,$$

$$S_3 = \frac{Var[E(Y | X_3)]}{Var(Y)} = \frac{9}{70} \cong 0.129, \quad S_4 = \frac{Var[E(Y | X_4)]}{Var(Y)} = \frac{25}{70} \cong 0.357,$$

$$S_5 = \frac{Var[E(Y | X_5)]}{Var(Y)} = \frac{1}{70} \cong 0.014, \quad S_6 = \frac{Var[E(Y | X_6)]}{Var(Y)} = \frac{1}{70} \cong 0.014,$$

$$S_7 = \frac{Var[E(Y | X_7)]}{Var(Y)} = \frac{4}{70} \cong 0.057, \quad S_8 = \frac{Var[E(Y | X_8)]}{Var(Y)} = \frac{25}{70} \cong 0.357.$$

سه متغیر کم اهمیت  $X_1$ ،  $X_5$  و  $X_6$  را از مدل حذف می‌کنیم و با پنج متغیر باقیمانده محاسبات را تکرار می‌کنیم. اکنون برای خروجی  $Y$ ، یک واریانس جدید داریم:

$$Var(Y) = \sum_{i=1}^5 Var(X_i) = 4 + 9 + 25 + 4 + 25 = 67.$$

اثرات اصلی به صورت زیر هستند:

$$S_1 = \frac{Var[E(Y | X_1)]}{Var(Y)} = \frac{4}{67} \cong 0.059, \quad S_2 = \frac{Var[E(Y | X_2)]}{Var(Y)} = \frac{9}{67} \cong 0.134,$$

$$S_3 = \frac{Var[E(Y | X_3)]}{Var(Y)} = \frac{25}{67} \cong 0.373, \quad S_4 = \frac{Var[E(Y | X_4)]}{Var(Y)} = \frac{4}{67} \cong 0.059,$$

$$S_5 = \frac{Var[E(Y | X_5)]}{Var(Y)} = \frac{25}{67} \cong 0.373.$$

می‌بینیم که دو عامل  $X_3$  و  $X_5$  بیشترین تاثیر را دارند و هر یک ۳۷٪ واریانس خروجی را تعیین می‌کنند. اگر این دو متغیر را در یک مقدار از دامنه تغییراتشان ثابت کنیم، آنگاه سه متغیر  $X_1$ ،  $X_2$  و

$X_4$  تغییرپذیر خواهند بود. بنابراین واریانس خروجی به مقدار زیر کاهش می‌یابد:

$$Var(Y) = 4 + 9 + 4 = 17.$$

نتیجه می‌گیریم که با ثابت نگه‌داشتن دو عامل مهم  $x_3$  و  $x_5$ ، واریانس خروجی از ۶۷ به ۱۷ کاهش می‌یابد، یعنی یک کاهش ۷۱٪ رخ می‌دهد.

## ۳.۲ برآورد شاخص حساسیت اصلی

برای اثبات روابطی که در ادامه ذکر می‌کنیم، به جانفدا (۱۳۹۰) مراجعه کنید. عبارت شاخص حساسیت مرتبه اول را می‌توان به صورت زیر بیان نمود (جانفدا، ۱۳۹۰)

$$S_j = \frac{\text{Var}(E(Y | X_j))}{\text{Var}(Y)} = \frac{U_j - E^2(Y)}{\text{Var}(Y)}, \quad j = 1, \dots, d,$$

که در آن

$$U_j = \int f(x_1, \dots, x_j, \dots, x_d) f(x'_1, \dots, x_j, \dots, x'_d) \prod_{i=1}^d p_i(x_i) dx_i \prod_{i=1, i \neq j}^d p_i(x'_i) dx'_i.$$

عبارت  $U_j$ ، انتگرالی با بعد  $2d - 1$  است که به دلایلی همچون بزرگ بودن  $d$  و نیز فقدان رابطه صریح برای  $f$ ، محاسبه آن امکان‌ناپذیر است. اما چون به ازای هر مجموعه از ورودی‌های  $\mathbf{x} = (x_1, \dots, x_d)$ ، مقدار خروجی  $y = f(x_1, \dots, x_d)$  معلوم است، می‌توانیم  $U_j$  را با روش انتگرال‌گیری مونت‌کارلویی برآورد کنیم.

در برآورد مونت‌کارلویی  $U_j$ ، از شبیه‌سازی دو ماتریس  $n \times d$  با عناصر تصادفی استفاده می‌کنیم.

این دو را ماتریس‌های  $A$  و  $B$  می‌نامیم،

$$A = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1d} \\ x_{21} & x_{22} & \dots & x_{2d} \\ \vdots & \vdots & \dots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nd} \end{bmatrix}, \quad B = \begin{bmatrix} x'_{11} & x'_{12} & \dots & x'_{1d} \\ x'_{21} & x'_{22} & \dots & x'_{2d} \\ \vdots & \vdots & \dots & \vdots \\ x'_{n1} & x'_{n2} & \dots & x'_{nd} \end{bmatrix}.$$

تابع  $f(x'_1, \dots, x_j, \dots, x'_d)$  در  $U_j$  به گونه‌ایست که فقط متغیر  $j$ ام آن از ماتریس  $A$  می‌باشد و

بقیه ورودی‌ها از عناصر ماتریس  $B$  هستند. برای ساختن مجموعه ورودی تابع  $f(x'_1, \dots, x_j, \dots, x'_d)$

ماتریس جدید  $C_j$ ،  $j = 1, \dots, d$ ، را می‌سازیم. عناصر ماتریس  $C_j$  همان عناصر ماتریس  $B$  هستند

که فقط ستون  $j$ ام آن، ستون  $j$ ام ماتریس  $A$  می‌باشد:

$$C_j = \begin{bmatrix} x'_{11} & x'_{12} & \cdots & x_{1j} & \cdots & x'_{1d} \\ x'_{21} & x'_{22} & \cdots & x_{2j} & \cdots & x'_{2d} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ x'_{n1} & x'_{n2} & \cdots & x_{nj} & \cdots & x'_{nd} \end{bmatrix}.$$

به کمک ماتریس‌های  $A$ ،  $B$  و  $C_j$ ، برآورد مونت کارلویی شاخص حساسیت مرتبه اول متغیر  $x_j$  با

رابطه زیر محاسبه می‌شود:

$$\hat{S}_j = \frac{\hat{U}_j - \hat{E}^\Psi(Y)}{\widehat{Var}(Y)},$$

که در آن

$$\begin{aligned} \hat{U}_j &= \frac{1}{n-1} \sum_{r=1}^n f(x_{r1}, \dots, x_{rd}) \times f(x'_{r1}, \dots, x'_{r(j-1)}, x_{rj}, x'_{r(j+1)}, \dots, x'_{rd}) \\ &= \frac{1}{n-1} \sum_{r=1}^n Y_A^{(r)} \cdot Y_{C_j}^{(r)}, \end{aligned}$$

که در آن

$$\hat{E}^\Psi(Y) = (\hat{E}(Y))^\Psi = \left( \frac{1}{n} \sum_{i=1}^n y_A^{(i)} \right)^\Psi$$

یا

$$\hat{E}^\Psi(Y) = \left( \frac{1}{n} \right) \sum_{i=1}^n y_A^{(i)} \times y_B^{(i)}$$

و

$$\hat{V}(Y) = \left( \frac{1}{n} \right) \sum_{i=1}^n (y_A^{(i)})^\Psi - \hat{E}^\Psi(Y)$$

که در آن  $y_A^{(i)}$ ،  $y_B^{(i)}$  و  $y_{C_j}^{(i)}$  نشان‌دهنده مقدار خروجی مدل کامپیوتری برای سطر  $i$ ام ماتریس‌های  $A$  و  $B$  و  $C_j$  است.

## ۴.۲ برآورد شاخص حساسیت کل

شاخص حساسیت کل را می‌توان به صورت زیر بازنویسی نمود (جانفدا، ۱۳۹۰):

$$S_j^t = 1 - \frac{Var(E(Y | X_{\sim j}))}{Var(Y)} = 1 - \frac{U_{\sim j} - E^\Psi(Y)}{Var(Y)},$$

که در آن

$$U_{\sim j} = \int f(x_1, \dots, x_j, \dots, x_d) f(x_1, \dots, x'_j, \dots, x_d) \prod_{i=1}^d p_i(x_i) dx_i p_j(x'_j) dx'_j.$$

برای برآورد این شاخص، فقط به برآورد  $U_{\sim j}$  نیاز داریم. ساختار  $U_{\sim j}$  نشان می‌دهد که برای برآورد آن علاوه بر ماتریس  $A$  نیاز به ماتریسی داریم که سطرهای آن ورودی‌های تابع  $f(x_1, \dots, x'_j, \dots, x_d)$  را در برداشته باشد. لذا ماتریس‌های جدید  $D_j$ ،  $j = 1, \dots, d$ ، را می‌سازیم. ماتریس  $D_j$  همان ماتریس  $A$  است که فقط ستون  $j$ ام آن با ستون  $j$ ام ماتریس  $B$  جایگزین شده است:

$$D_j = \begin{bmatrix} x_{11} & x_{12} & \dots & x'_{1j} & \dots & x_{1d} \\ x_{21} & x_{22} & \dots & x'_{2j} & \dots & x_{2d} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ x_{n1} & x_{n2} & \dots & x'_{nj} & \dots & x_{nd} \end{bmatrix}.$$

با استفاده از روش انتگرال‌گیری مونت کارلویی،  $U_{\sim j}$  به صورت زیر برآورد می‌شود:

$$\begin{aligned} \hat{U}_{\sim j} &= \frac{1}{n-1} \sum_{r=1}^n f(x_{r1}, x_{r2}, \dots, x_{rd}) \times f(x_{r1}, x_{r2}, \dots, x_{r(j-1)}, x'_{rj}, x_{r(j+1)}, \dots, x_{rd}) \\ &= \frac{1}{n-1} \sum_{r=1}^n Y_A^{(r)} \cdot Y_{D_j}^{(r)}. \end{aligned}$$

## ۵.۲ هزینه‌های محاسباتی شاخص‌های حساسیت

در برآورد اثرات اصلی، نیاز به اجرای مدل برای تمام سطرهای ماتریس‌های  $A$ ،  $B$  و  $C_j$ ها داریم. یعنی به تعداد  $2n + dn$  دفعه باید مدل کامپیوتری اجرا شود همچنین در برآورد اثرات کل نیاز به اجرای تمام سطرهای ماتریس‌های  $D_j$  داریم یعنی  $dn$  بار اجرای مدل. لذا در مجموع  $2n(1+d)$  بار اجرا برای برآورد اثرات اصلی و اثرات کل نیاز است. اگر زمان هر اجرا را برابر  $t$  فرض کنیم، آنگاه زمان اجراها برابر  $2nt(d+1)$  خواهد شد. چون زمان اجرای یک مدل کامپیوتری بالاست (چند دقیقه یا حتی چند ساعت)، لازم است  $n$  را طوری انتخاب کنیم تا ضمن بالا بودن دقت برآوردها، هزینه محاسباتی معقول و به صرفه باشد.

# فصل ۳

## مدل جانشین بیزی: کریگینگ

### ۱.۳ مقدمه

در فصل قبل با استفاده از روش مونت کارلو، برآوردی برای شاخص‌های حساسیت ارایه دادیم، که مستلزم تعداد دفعات زیاد اجرای مدل کامپیوتری است. بالا بودن تعداد اجراها موجب افزایش هزینه زمانی می‌شود و با وجود این‌که توان کامپیوترها در دهه‌های اخیر افزایش یافته است، هزینه محاسباتی بالا در مقیاس‌های بزرگ همچنان مشکلی اساسی در مطالعه مدل‌های ریاضی پیچیده به شمار می‌آید. این هزینه‌های زیاد، به علت بالا بودن بعد ورودی‌ها، تعدد اجراها و زمان‌بر بودن آن‌ها ایجاد می‌شوند. اگرچه اجرای مدل‌های پیچیده برای درک طبیعت موضوع فرآیند مورد بررسی سودمند است، اما هزینه محاسباتی بالا در این‌گونه مدل‌ها چالشی است که گاهی اوقات موجب استفاده محدودتر از آن‌ها می‌شود. لذا برای مقابله با موضوع هزینه زمانی تحمیل‌شده، انگیزه جایگزینی تابع پیچیده ریاضی، با برآوردی ساده‌تر از آن مطرح می‌شود. این جایگزینی، تحت عنوان «مدل جانشین» مطرح است که در آن رفتار تابع اصلی توسط مدل‌های پیشرفته آماری یا یادگیری ماشین، مدل‌بندی می‌شود. استفاده از مدل‌های جانشین سبب می‌شود تا هزینه زمانی در هر اجرا به نحو چشم‌گیری کاهش یابد. مدل‌های جانشین که به «مقلد»<sup>۱</sup> یا «فرا مدل»<sup>۲</sup> نیز معروف هستند، در طول دهه اخیر به‌طور گسترده‌ای تکامل یافته و طیف وسیعی از تحقیقات آماری را دربر می‌گیرند. استورلی<sup>۳</sup> و همکاران (۲۰۰۹) به منظور انجام تحلیل حساسیت،

<sup>۱</sup>Emulator

<sup>۲</sup>Meta model

<sup>۳</sup>Storlie



عملکرد روش‌های مختلف رگرسیون ناپارامتری را به عنوان مدل‌های جانشین، مقایسه نمودند. رضوی<sup>۱</sup> و همکاران (۲۰۱۲) با مرور بیش از ۱۰۰ مقاله، به دسته‌بندی و تحلیل تلاش‌های انجام شده در این زمینه پرداختند. دیدگاه بیزی براساس فرآیند گاوسی، فرامدل‌های مناسبی را تولید می‌کند. اندیشه به‌کار بردن رهیافت بیزی، به ساکس<sup>۲</sup> و همکاران (۱۹۸۹) باز می‌گردد. آن‌ها مقادیر مشاهده‌شده خروجی مدل (متغیر پاسخ) را مصداقی از فرآیند گاوسی در نظر گرفتند. پس از آن نیز تحقیقات بسیاری بر همین اساس انجام گرفت که از آن جمله می‌توان به کرگ<sup>۳</sup> و همکاران (۲۰۰۱)، کندی و اوهگن<sup>۴</sup> (۲۰۰۱)، گلداشتن و رویجر<sup>۵</sup> (۲۰۰۶)، بایاری<sup>۶</sup> و همکاران (۲۰۰۹) و گراماسی<sup>۷</sup> و همکاران (۲۰۱۲) اشاره نمود. همچنین هراری و استینبرگ<sup>۸</sup> (۲۰۱۳) ترکیب محدبی از چند فرآیند گاوسی با توابع همبستگی از خانواده‌های متفاوت را به عنوان یک مدل جانشین به‌کار بردند. روش عناصر متناهی و چگالی مشبک<sup>۹</sup> نیز توسط روی<sup>۱۰</sup> و همکاران (۲۰۱۳) برای یافتن مدل جانشین استفاده شد.

## ۲.۳ پیشینه مدل‌های جانشین

ایده مدل‌های جانشین که در دو دهه اخیر کامل‌تر گشته است، به بلانینگ<sup>۱۱</sup> (۱۹۷۵) و کلیجنن (۱۹۷۵) برمی‌گردد. «طرح و تجزیه آزمایش‌های کامپیوتری» عنوانی است که معمولاً به این مجموعه از روش‌های تحلیلی نسبت داده می‌شود و بحث‌های مرتبط با آن توسط ساکس و همکاران (۱۹۸۹)، ون گیج<sup>۱۲</sup> (۱۹۹۱)، سانتنر و همکاران (۲۰۰۳)، فنگ<sup>۱۳</sup> و همکاران (۲۰۰۶)، بایاری و همکاران (۲۰۰۹)،

<sup>۱</sup>Razavi

<sup>۲</sup>Sacks

<sup>۳</sup>Craig

<sup>۴</sup>Kennedy and O'Hagan

<sup>۵</sup>Goldstein and Rougier

<sup>۶</sup>Bayarri

<sup>۷</sup>Gramacy

<sup>۸</sup>Harari and Steinberg

<sup>۹</sup>Mesh density

<sup>۱۰</sup>Rui

<sup>۱۱</sup>Blanning

<sup>۱۲</sup>van Gigch

<sup>۱۳</sup>Fang

کوهنت و استینبرگ<sup>۱</sup> (۲۰۱۰)، لوی<sup>۲</sup> و استینبرگ (۲۰۱۰) و کلیجنن (۲۰۱۰) انجام گرفته‌اند. این جریان گسترده از تحقیقات، نشان‌دهنده روش کلاسیک مدل‌بندی جانشین است و در رابطه با مدل‌های ایستا<sup>۳</sup>، کامل گشته است (راتو<sup>۴</sup> و همکاران، ۲۰۱۲). از جمله کارهای انجام‌گرفته بر روی مدل‌های ایستا، می‌توان به اوکلی و اوگن (۲۰۰۴)، بتس<sup>۵</sup> و همکاران (۲۰۰۶)، استورلی و هلتن<sup>۶</sup> (۲۰۰۷) و راتو و پاگانو<sup>۷</sup> (۲۰۱۰) اشاره کرد.

شاهسونی و گریموال (۲۰۰۹) با تولید یک طرح آزمایش سازگار با تابع تحت بررسی در مدل کامپیوتری، مدل جانشینی را بر اساس آن طرح آزمایش و مبتنی بر چندجمله‌ای‌های موضعی درجه دوم کامل<sup>۸</sup>، ارایه دادند. شاهسونی و همکاران (۲۰۱۰) نیز مدل جانشین MARS<sup>۹</sup> را در مورد یک تابع آزمون غیرخطی و غیرجمعی به‌کار بردند. نتایج نشان داد که مدل جانشین MARS برآوردهای قابل قبولی از شاخص‌های حساسیت کل با هزینه خیلی کمتر در مقایسه با استفاده از اجراهای مدل اصلی ارایه می‌دهد. کارنوال<sup>۱۰</sup> و همکاران (۲۰۱۲) در بهینه‌سازی مدل تعیینی سه بعدی برای روابط غیرخطی متغیرهای کنترل و مشخصه‌های کیفی هوا که نیاز به هزاران اجرای مدل داشت، مدل جانشین شبکه عصبی را به‌کار گرفتند. ویلا-ویالنکس<sup>۱۱</sup> و همکاران (۲۰۱۲)، برای طرح یک مدل جانشین مشتقات نیتروژن در زمین‌های قابل کشت اروپا، هفت روش آماری (رگرسیون خطی، هموارسازی اسپلاین، کریگینگ، SVM<sup>۱۲</sup>، شبکه‌های عصبی و روش Baggin) را به‌کار گرفتند. آن‌ها با اجرای انواع متفاوت مدل‌های جانشین روی مقادیر شبیه‌سازی شده‌ای که از مدل پیچیده اصلی تولید شده بودند، به این نتیجه مهم دست یافتند که هرگونه از این مدل‌های جانشین تا چه حد توانسته‌اند جایگزین مدل واقعی باشند.

---

<sup>۱</sup>Kuhnt and Steinberg

<sup>۲</sup>Levy

<sup>۳</sup>Static

<sup>۴</sup>Ratto

<sup>۵</sup>Bates

<sup>۶</sup>Helton

<sup>۷</sup>Ratto and Pagano

<sup>۸</sup>Full quadratic local polynomials

<sup>۹</sup>Multivariate adaptive regression splines

<sup>۱۰</sup>Carnevale

<sup>۱۱</sup>Villa-Vialaneix

<sup>۱۲</sup>Support vector machine

در مبحث مدل‌بندی جانشین، انواع مدل‌های آماری مانند کریگینگ، تابع‌های بنیادی شعاعی<sup>۱</sup>، جنگل‌های تصادفی، شبکه‌های عصبی و MARS را می‌توان به‌عنوان مدل جانشین استفاده نمود (فریدمن<sup>۲</sup>، ۱۹۹۱). مدل جانشینی که در این پایان‌نامه استفاده می‌کنیم یک مدل بیزی است که «کریگینگ» نام دارد.

### ۳.۳ تاریخچه کریگینگ

ریشه کریگینگ به تحقیقات اکتشافی ولد<sup>۳</sup> (۱۹۳۸)، کلموگروف<sup>۴</sup> (۱۹۳۹ و ۱۹۴۱) و وینر<sup>۵</sup> (۱۹۴۹) برمی‌گردد، اگر چه تحقیقات آن‌ها مستقیماً به رگرسیون مربوط است تا به مدل‌بندی کریگینگ. اولین بار در سال ۱۹۵۱، کریگ، یک مهندس معدن آفریقای جنوبی، این روش را توسعه داد. او از این روش برای حل مسایل معدن طلا استفاده کرد.

کریگینگ در ابتدا توسط پیر کارلیر<sup>۶</sup>، «کریگیج»<sup>۷</sup> نامگذاری شد و سپس توسط ماترون<sup>۸</sup> (۱۹۶۳) به نام کنونی‌اش تغییر یافت. ماترون کریگینگ را به‌عنوان روشی توصیف می‌کند که می‌تواند درجه یک قطعه را توسط محاسبه متوسط وزنی نمونه‌های در دسترس، پیش‌گویی کند. او تأکید کرد که وزن‌های مناسب باید کوچکترین واریانس را بسازند. درک کریگ درباره روش مذکور این است که کریگینگ یک رگرسیون چندگانه است که بهترین متوسط وزنی خطی از درجه سنگ یک بلوک سنگی از هر اندازه‌ای را با کمک یک مجموعه اوزان از داده‌های موجود و مربوط که در درون و بیرون آن بلوک سنگ هستند، تولید می‌کند. ارد<sup>۹</sup> (۱۹۸۳) بیان می‌کند که در علوم آماری، کریگینگ روشی درون‌یابی برای فرآیندهای تصادفی فضایی است. بر طبق نظر همیاری و نوفریگر<sup>۱۰</sup> (۱۹۸۷)، کریگینگ، صورتی از

<sup>۱</sup> Radial basis functions

<sup>۲</sup> Friedman

<sup>۳</sup> Wold

<sup>۴</sup> Kolmogorov

<sup>۵</sup> Wiener

<sup>۶</sup> Pierre Carlier

<sup>۷</sup> Krigeage

<sup>۸</sup> Matheron

<sup>۹</sup> Ord

<sup>۱۰</sup> Hemyari and Nofziger

متوسط وزنی است که در آن، وزن‌ها به مکان و ساختار کوواریانس یا نیم‌تغییرنگار<sup>۱</sup> نقاط مشاهده‌شده، بستگی دارند. انتخاب وزن‌ها باید به‌گونه‌ای باشد که خطای پیش‌گویی را کمتر از هر مجموع خطی دیگر بسازد. نیم‌تغییرنگار، تابعی است که برای نشان دادن همبستگی فضایی در مشاهدات اندازه‌گیری‌شده در مکان‌های نمونه، استفاده می‌شود.

در اوایل دهه ۱۹۶۰، ماترون نظریه ریاضی منطبق با کریگینگ را توسعه داد. در اصل، کریگینگ یک ترکیب خطی بود و در تعمیم‌های بعدی در زمین‌آمار، کریگینگ به پیش‌گویی فضایی غیرخطی توسعه یافت. ماترون همچنین نشان داد که ظاهراً کریگینگ و اسپلین هم‌ارز هستند. دوبرول (۱۹۸۳) تفاوت‌های بین اسپلین و کریگینگ را بررسی کرد.

مقدمه‌ای جامع برای اصل و سرچشمه کریگینگ در کرسی (۱۹۹۰) آمده است. کرسی، کریگینگ را از منظر راهبردهای مختلف بررسی نمود و بیان کرد کریگینگ، معادل با پیش‌گویی خطی بهینه فضایی است.

## ۴.۳ مروری بر کاربردهای کریگینگ

کاربردهای کریگینگ، میدان‌های کلاسیک معدنی، زمین‌شناسی و علوم خاک، آب‌شناسی، هواشناسی و سایر علوم مختلف را پوشش می‌دهد و اخیراً برای طراحی مهندسی، برآوردهای هزینه‌بر، دریافت بی‌سیم و شبکه‌ها، درون‌یابی مدل‌های شبیه‌سازی‌شده و بهینه‌سازی تکاملی نیز به‌کار رفته است. کاربرد کریگینگ عمدتاً در زمین‌شناسی و معدن است. کریگینگ به‌طور وسیع در تهیه نقشه‌های محیطی، در علم هواشناسی، پیش‌گویی مقادیر خواص خاک در مکان‌های نمونه‌گیری‌نشده و در میدان‌های آب‌شناسی، استفاده شده است. در سال‌های اخیر، کریگینگ برای طراحی مهندسی، علم مواد، مهندسی زیست‌مکانیکی، پخش موج بی‌سیم، تحلیل حساسیت اقتصادی و برآوردهای گران، درون‌یابی مدل‌های شبیه‌سازی و بهینه‌سازی به‌کار رفته است.

<sup>۱</sup>Semi-variogram

### ۵.۳ کاربرد کریگینگ در مدل‌های جانشین

استفاده از کریگینگ در مدل‌های جانشین ابتدا در حوزه مدل‌های تعیینی رخ داد. مرجع کلاسیک آن به ساکس و همکاران (۱۹۸۹) برمی‌گردد. آن‌ها آزمایش‌های کامپیوتری را که دارای اجراهای پرهزینه و خروجی‌های تعیینی هستند، در نظر گرفتند و برای این آزمایش‌های کامپیوتری، از کریگینگ برای برازش یک پیش‌گوی موثر برای کاهش هزینه محاسباتی استفاده کردند.

میچل و موریس<sup>۱</sup> (۱۹۹۲)، کریگینگ را به عنوان جایگزینی برای روش رویه پاسخ مرسوم در آزمایش‌های کامپیوتری، بررسی کردند. آن‌ها از کریگینگ به عنوان یک روش بیزی در بررسی اهمیت عامل‌های ورودی، بهینه‌سازی یک متغیر وابسته و حل مسایل معکوس یک آزمایش کامپیوتری روی مدل جریان آب‌های زیرزمینی، استفاده کردند. گرچه بسیاری از محققان، تمرکزشان روی شبیه‌سازی‌های تعیینی بوده است، اما آن‌ها نشان دادند که چگونه می‌توان با انجام اصلاحاتی، شبیه‌سازی خروجی‌های تصادفی را نیز مدل‌بندی کرد.

تروچو<sup>۲</sup> و همکاران (۱۹۹۹) کریگینگ دوگان<sup>۳</sup> را به عنوان یک روش آماری برای مدل‌بندی رفتار مکانیکی قابل رویت آلیاژهای حافظه‌دار به‌کار بردند. آن‌ها کریگینگ دوگان را برای برقرار کردن یک معادله صریح از هر دور جزیی درون محدوده اصلی به‌کار گرفتند و بنابراین یک قانون کلی برای آلیاژهای حافظه‌دار، ارائه کردند.

مکشیمر<sup>۴</sup> و همکاران (۲۰۰۱) روش‌های موثر برای تعیین اعتبار کریگینگ و دیگر مدل‌های جانشین در مدل‌بندی شبیه‌سازی‌های تعیینی را توسعه دادند. آن‌ها روش‌های تشخیص محاسبات هزینه‌بر را برای اعتبار مدل جانشین بر مبنای اعتبارسنجی متقابل با  $k$  مشاهده خارج‌شده، بررسی کردند و راهبردهایی برای چگونگی انتخاب  $k$  ارائه دادند.

<sup>۱</sup>Morris

<sup>۲</sup>Trochu

<sup>۳</sup>Dual Kriging

<sup>۴</sup>Meckesheimer

## ۶.۳ روش کریگینگ

مساله اصلی ما در این فصل، تقریب مدل کامپیوتری توسط یک مدل جانشین است. فرآیند ساخت یک مدل جانشین به دو گام اصلی تقسیم می‌شود: در گام اول، یک مدل آماری به عنوان جایگزین مدل کامپیوتری تعیین می‌شود و در گام دوم، پس از برآورد پارامترهای مدل، پیش‌گو معین خواهد شد. بنابراین مدل جانشین را می‌توان به صورت یک تقریب آماری دید. مدل جانشین کریگینگ که در این جا به کار می‌بریم توسط ماترون (۱۹۶۳) معرفی شده است. کریگینگ روشی آماری برای درونیابی داده‌های پراکنده است و برای پیش‌گویی در آزمایش‌های کامپیوتری نیز مورد استفاده قرار می‌گیرد.

در روش کریگینگ، خروجی یک مدل تعیینی به صورت تحقق از یک فرآیند تصادفی مدل‌بندی می‌شود. فرض کنید  $y(\mathbf{x})$  خروجی معلوم مدل تعیینی و  $\mathbf{x} = (x_1, \dots, x_d)$  یک بردار از مقادیر متغیرهای ورودی با توزیع نامعلوم باشد. برای بردار تصادفی  $\mathbf{x}$ ، متغیر تصادفی  $Y(\mathbf{x})$  را تحقق از یک فرآیند تصادفی به صورت زیر در نظر بگیرید:

$$Y(\mathbf{x}) = \sum_{j=1}^k \beta_j f_j(\mathbf{x}) + Z(\mathbf{x}), \quad (1.3)$$

که در آن  $Z(\cdot)$  یک فرآیند تصادفی گاوسی ایستا با میانگین صفر است. جهت سهولت محاسبات فرض می‌شود  $Z(\cdot)$  دارای توزیع نرمال است. بخش اول عبارت (۱.۳)،  $\sum_j \beta_j f_j(\mathbf{x})$ ، که آن را روند<sup>۱</sup> گویند، یک ترکیب خطی از توابع حقیقی مقدار از قبل انتخاب شده  $f_1(\cdot), \dots, f_k(\cdot)$ ، با ضرایب  $\beta = (\beta_1, \dots, \beta_k)$  است که به طور معمول  $f_1 = 1$ . رابطه (۱.۳)، حالت کلی مدل کریگینگ است که با عنوان کریگینگ عام<sup>۲</sup> (UK) شناخته می‌شود. با ساده‌تر شدن بخش رگرسیونی این مدل، انواع متفاوتی برای کریگینگ تعریف می‌شود. کریگینگ معمولی<sup>۳</sup> (OK) که متداول‌ترین نوع کریگینگ است، به صورت

$$Y(\mathbf{x}) = \beta + Z(\mathbf{x}), \quad (2.3)$$

بیان می‌شود که در آن روند فقط شامل ضریب رگرسیونی ثابت  $\beta$  است. برای دو ورودی تصادفی  $\mathbf{u}$  و

<sup>۱</sup>Trend

<sup>۲</sup>Universal Kriging

<sup>۳</sup>Ordinary Kriging

$\mathbf{x}$ ، همبستگی بین  $Z(\mathbf{x})$  و  $Z(\mathbf{u})$  با تابع کوواریانس

$$Cov(\mathbf{u}, \mathbf{x}) = E(Z(\mathbf{u})Z(\mathbf{x})) = \sigma^2 R(\mathbf{u}, \mathbf{x}), \quad (3.3)$$

نشان داده می‌شود که  $\sigma^2$  واریانس فرآیند و  $R(\mathbf{u}, \mathbf{x})$  تابع همبستگی است که همبستگی با افزایش فاصله بین دو نقطه کاهش می‌یابد.

بر اساس کار بنیادی انجام شده در زمینه طرح و تجزیه آزمایش‌های کامپیوتری (ساکس و همکاران، ۱۹۸۹؛ ولچ<sup>۱</sup> و همکاران، ۱۹۹۲) می‌توان دریافت که انحرافات مدل پیچیده اصلی از  $\beta$  در (۲.۳) یا  $\sum_k \beta_k f_k(\mathbf{x})$  در (۱.۳)، هرچند قطعی است اما به یک فرآیند تصادفی مانند  $Z(\cdot)$  شباهت دارد. همچنین اگر رفتار مدل پیچیده اصلی، بر وجود روند دلالت کند، مدل (۱.۳) که شامل یک روند با عوامل نامعلوم است، به کار می‌رود. با این وجود ساکس و همکاران (۱۹۸۹) دریافتند که استفاده از مدل (۲.۳) نیز تاثیر نامطلوبی بر پیش‌گویی‌هایشان نداشته است.

بحرانی‌ترین قسمت این روش، انتخاب تابع همبستگی  $R(\cdot, \cdot)$  است، زیرا نشان‌دهنده میزان اطلاعاتی است که خروجی مدل در نقطه  $\mathbf{u}$  برای پیش‌گویی خروجی نقطه  $\mathbf{x}$  در اختیار ما قرار می‌دهد. برای مدل‌هایی با  $d$  متغیر ورودی، به کار بردن تابع همبستگی به صورت حاصل ضرب توابع همبستگی یک بعدی، یعنی

$$R(\mathbf{u}, \mathbf{x}) = \prod_{j=1}^d R_j(u_j, x_j), \quad (4.3)$$

از لحاظ محاسباتی مناسب است. این شکل همبستگی، متغیرها را مستقل از هم در نظر می‌گیرد. یکی از متداول‌ترین توابع همبستگی،

$$R(\mathbf{u}, \mathbf{x}) = \prod_{j=1}^d \exp(-\theta_j |u_j - x_j|^p) \quad (5.3)$$

است که در آن  $0 < p \leq 2$  (ساکس و همکاران، ۱۹۸۹). حالت  $p = 1$  فرآیندهایی را تولید می‌کند که پیوسته هستند اما خیلی هموار نیستند. حالت  $p = 2$  فرآیندهایی را حاصل می‌کند که به طور نامتناهی مشتق‌پذیر هستند (ساکس و همکاران، ۱۹۸۹). توابع همبستگی مذکور، همگی بر دو مولفه استوار هستند. اولین مولفه، تابع فاصله بین دو نقطه  $\mathbf{x}$  و  $\mathbf{u}$ ، و دومین مولفه، پارامتر مقیاس است. پارامتر مقیاس،  $\theta$ ، اندازه همواری مدل را نسبت به هر یک از متغیرهای ورودی نشان می‌دهد. به این معنا که اطلاع ما از خروجی مدل در یک نقطه از متغیر ورودی، تا چه میزان در پیش‌گویی خروجی در نقطه دیگر

<sup>۱</sup>Welch

استفاده خواهد شد. برای هر نوع کاربرد، تابع  $R(\cdot, \cdot)$  به طور منحصر به فردی تعیین می شود به طوری که

بتواند به بهترین وجه، همبستگی را پیش گویی کند. تابع همبستگی گاوسی ( $p = 2$ )

$$\text{corr}(Z(\mathbf{u}), Z(\mathbf{x})) = \exp\left(-\sum_{j=1}^d \theta_j (u_j - x_j)^2\right), \quad (6.3)$$

به طور عمده در تحلیل ها استفاده می شود (هایلاک، ۱۹۹۷).

اگر تابع کوواریانس یک فرآیند تصادفی، فقط به فاصله بین دو نقطه بستگی داشته باشد و تحت تاثیر مکان های ثابت مخصوص آن ها نباشد، فرآیند را ایستای مرتبه دوم<sup>۱</sup> گویند. زمانی که یک فرآیند تصادفی ایستا، ساختار کوواریانس یکسان در همه جهتها داشته باشد، آن فرآیند را همسان گرد<sup>۲</sup> و همگن<sup>۳</sup> می گویند.

همگنی به این معناست که اگر فاصله بین دو نقطه مساوی باشد، هر دو نقطه همبستگی متناهی یکسان با هم داشته باشند. همسان گرد بودن به این معناست که تابع کوواریانس آن فرآیند، ویژگی مستقل از جهت را داراست. برای سادگی در کل فرض می شود که توابع کوواریانس، همگن و همسان گرد هستند. البته این فرض ها پیش نیاز به کارگیری روش کریگینگ نیستند (لیو، ۲۰۰۹).

### ۱.۶.۳ پیش گوی کریگینگ

فرض کنید  $\mathbf{S} = [s_1, \dots, s_n]$  یک طرح آزمایش یا داده های مدل ساز باشد که در آن بردار  $d$  بعدی  $\mathbf{y}_S = [y(s_1), \dots, y(s_n)]^T$  بردار مقادیر پاسخ در نقاط طرح باشد.

کریگینگ یک پیش گوی نااریب خطی بهینه<sup>۴</sup> (BLUP) است که به صورت میانگین موزون نقاط طرح آزمایش بیان شده و تعیین اوزان به گونه ای صورت می گیرد که واریانس پیش گو، کمینه شود.

در کریگینگ برای پیش گویی مقدار پاسخ در نقطه آزمون نشده  $\mathbf{x}$ ، از پیش گوی خطی زیر استفاده

می شود:

$$\hat{y}(\mathbf{x}) = \mathbf{w}(\mathbf{x})^T \mathbf{y}_S, \quad (7.3)$$

<sup>۱</sup>Second order stationary

<sup>۲</sup>Isotropic

<sup>۳</sup>Homogeneous

<sup>۴</sup>Best linear unbiased predictor



که در آن  $\hat{y}(\mathbf{x})$  برآورد  $y(\mathbf{x})$  است.

با در نظر گرفتن  $\hat{Y}(\mathbf{x})$  به صورت کمیته تصادفی، بردار وزن  $\mathbf{w}(\mathbf{x})^T = [w_1(\mathbf{x}), w_2(\mathbf{x}), \dots, w_n(\mathbf{x})]$  که در آن، وزن‌ها با نقطه پیش‌گویی در رابطه هستند را به‌گونه‌ای به دست می‌آوریم تا میانگین مربع خطای پیش‌گویی<sup>۱</sup> (MSPE)، یعنی

$$MSPE[\hat{Y}(\mathbf{x})] = E[\mathbf{w}(\mathbf{x})^T \mathbf{Y}_S - Y(\mathbf{x})]^2, \quad (۸.۳)$$

با قید نااریبی برآوردگر، یعنی

$$E[\mathbf{w}(\mathbf{x})^T \mathbf{Y}_S] = E[Y(\mathbf{x})] = \mu, \quad (۹.۳)$$

کمینه شود.

### ۲.۶.۳ محاسبه واریانس خطا در روش کریگینگ معمولی

با اختصارنویسی  $w_i(\mathbf{x})$  را به صورت  $w_i$  نشان می‌دهیم. رابطه (۹.۳) را می‌توان به صورت زیر بازنویسی نمود

$$\begin{aligned} E[\mathbf{w}^T \mathbf{Y}_S] - E[Y(\mathbf{x})] &= E\left[\sum_{i=1}^n w_i Y(\mathbf{s}_i)\right] - E[Y(\mathbf{x})] \\ &= \sum_{i=1}^n w_i E[Y(\mathbf{s}_i)] - E[Y(\mathbf{x})] \\ &= \sum_{i=1}^n w_i \mu - \mu = \mu \left(\sum_{i=1}^n w_i - 1\right) = 0. \end{aligned}$$

بنابراین برای برقراری شرط نااریبی پیش‌گو، بایستی مجموع وزن‌ها برابر ۱ باشد، یعنی  $\sum_{i=1}^n w_i = 1$ .

ویژگی‌ای که کریگینگ را از سایر روش‌های برآورد خطی نااریب متمایز می‌سازد، کمینه‌سازی واریانس خطا است، که معادل با کمینه‌سازی  $MSPE(\hat{Y}(\mathbf{x}))$  است،

<sup>۱</sup>Mean square predictor error

$$\begin{aligned}
MSPE(\widehat{Y}(\mathbf{x})) &= E[\widehat{Y}(\mathbf{x}) - Y(\mathbf{x})]^2 \\
&= Var[\widehat{Y}(\mathbf{x}) - Y(\mathbf{x})] = Var[\mathbf{w}^T \mathbf{Y}_S - Y(\mathbf{x})] \\
&= Var\left[\sum_{i=1}^n w_i Y(\mathbf{s}_i) - Y(\mathbf{x})\right] \\
&= Var\left(\sum_{i=1}^n w_i Y(\mathbf{s}_i)\right) + Var(Y(\mathbf{x})) \\
&\quad - 2Cov\left(\sum_{i=1}^n w_i Y(\mathbf{s}_i), Y(\mathbf{x})\right) \\
&= \sum_{i=1}^n \sum_{j=1}^n w_i w_j Cov[Y(\mathbf{s}_i), Y(\mathbf{s}_j)] + Var(Y(\mathbf{x})) \\
&\quad - 2\sum_{i=1}^n w_i Cov[Y(\mathbf{s}_i), Y(\mathbf{x})].
\end{aligned} \tag{۱۰.۳}$$

برای کمینه‌سازی واریانس خطا، قید ناریبی را با استفاده از ضریب لاگرانژ  $\lambda(\mathbf{x})$ ، که به اختصار آن را با  $\lambda$  نشان می‌دهیم، به معادله  $MSPE$  اضافه می‌کنیم. داریم

$$\begin{aligned}
L(w_i, \lambda) &= E[\widehat{Y}(\mathbf{x}) - Y(\mathbf{x})]^2 + 2\lambda\left(\sum_{i=1}^n w_i - 1\right) \\
&= \sum_{i=1}^n \sum_{j=1}^n w_i w_j Cov[Y(\mathbf{s}_i), Y(\mathbf{s}_j)] + Var(Y(\mathbf{x})) \\
&\quad - 2\sum_{i=1}^n w_i Cov[Y(\mathbf{s}_i), Y(\mathbf{x})] + 2\lambda\left(\sum_{i=1}^n w_i - 1\right).
\end{aligned}$$

با مشتق‌گیری از معادله بالا نسبت به  $w_i$  ( $i = 1, \dots, n$ ) و  $\lambda$  و برابر صفر قرار دادن آن داریم،

$$\frac{\partial L(w_i, \lambda)}{\partial w_i} = 2\sum_{j=1}^n w_j Cov[Y(\mathbf{s}_i), Y(\mathbf{s}_j)] - 2Cov[Y(\mathbf{s}_i), Y(\mathbf{x})] + 2\lambda = 0,$$

که معادل است با

$$\sum_{j=1}^n w_j Cov[Y(\mathbf{s}_i), Y(\mathbf{s}_j)] + \lambda = Cov[Y(\mathbf{s}_i), Y(\mathbf{x})]. \tag{۱۱.۳}$$

همچنین

$$\frac{\partial L(w_i, \lambda)}{\partial \lambda} = 2\sum_{i=1}^n w_i - 2 = 0,$$

ولذا

$$\sum_{i=1}^n w_i = 1. \quad (12.3)$$

معادلات (۱۱.۳) و (۱۲.۳)، تشکیل دستگاه معادلات کریگینگ معمولی زیر را می‌دهند:

$$\begin{cases} \sum_{j=1}^n w_j Cov[Y(\mathbf{s}_i), Y(\mathbf{s}_j)] + \lambda = Cov[Y(\mathbf{s}_i), Y(\mathbf{x})], & i = 1, \dots, n, \\ \sum_{i=1}^n w_i = 1. \end{cases}$$

بیان ماتریسی این دستگاه معادلات به صورت زیر است:

$$\mathbf{C}_{(n+1) \times (n+1)} \cdot \mathbf{W}_{(n+1) \times 1} = \mathbf{D}_{(n+1) \times 1},$$

$$\begin{bmatrix} C_{11} & C_{12} & \dots & C_{1n} & 1 \\ C_{21} & C_{22} & \dots & C_{2n} & 1 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \lambda \end{bmatrix} = \begin{bmatrix} C_{1x} \\ C_{2x} \\ \vdots \\ C_{nx} \\ 1 \end{bmatrix},$$

که در آن  $C_{ij} = Cov(Y(\mathbf{s}_i), Y(\mathbf{s}_j))$  و  $C_{ix} = Cov(Y(\mathbf{s}_i), Y(\mathbf{x}))$ . ماتریس  $\mathbf{C}$  شامل کوواریانس‌های

بین نقاط طرح و یک ردیف و یک ستون ثابت‌ها است؛ بردار  $\mathbf{W}$  شامل وزن‌ها و ضریب لاگرانژ است

و بردار  $\mathbf{D}$  شامل کوواریانس‌های بین نقاط طرح و نقطه  $\mathbf{x}$  و عدد ۱ است. بنابراین

$$\mathbf{W} = \mathbf{C}^{-1} \mathbf{D}.$$

لازم به ذکر است که ماتریس  $\mathbf{C}$  فقط یک بار نیاز به محاسبه دارد ولی بردار  $\mathbf{D}$  برای هر  $\mathbf{x}$ ، باید

محاسبه شود. برای  $n$ ‌های بزرگ، معکوس کردن ماتریس  $\mathbf{C}$  محاسبات بسیار پیچیده‌ای دارد.

برای محاسبه واریانس پیش‌گویی کریگینگ معمولی در نقطه  $\mathbf{x}$ ، با استفاده از رابطه (۱۱.۳) داریم

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^n w_i w_j Cov[Y(\mathbf{s}_i), Y(\mathbf{s}_j)] &= \sum_{i=1}^n w_i \sum_{j=1}^n w_j Cov[Y(\mathbf{s}_i), Y(\mathbf{s}_j)] \\ &= \sum_{i=1}^n w_i (Cov[Y(\mathbf{s}_i), Y(\mathbf{x})] - \lambda), \end{aligned} \quad (13.3)$$

با جایگذاری (۱۳.۳) در رابطه (۱۰.۳)، واریانس پیش‌گویی کریگینگ معمولی به صورت زیر به دست

می‌آید:

$$\sigma_{OK}^2 = \sigma^2 - \sum_{i=1}^n w_i (Cov[Y(\mathbf{s}_i), Y(\mathbf{x})] + \lambda), \quad (14.3)$$

که بیان ماتریسی آن به صورت  $\sigma_{OK}^2 = \sigma^2 - \mathbf{W}^T \mathbf{D}$  است.

### ۳.۶.۳ محاسبه پیش‌گو در روش کریگینگ عام

با استفاده از مدل (۱.۳)، محاسبه پیش‌گوی کریگینگ توسط ساکس و همکاران (۱۹۸۹) به صورت زیر انجام گرفته است. در این جا  $\sigma^2$  معلوم فرض می‌شود. دوباره  $\mathbf{Y}_S = [Y(\mathbf{s}_1), \dots, Y(\mathbf{s}_n)]^T$  را به صورت یک بردار تصادفی از مقادیر پاسخ در نقاط طرح آزمایش در نظر بگیرید. به علاوه فرض کنید،

$$\mathbf{f}(\mathbf{x})^T = [f_1(\mathbf{x}), \dots, f_k(\mathbf{x})],$$

برداری از  $k$  تابع رگرسیونی است. همچنین ماتریس  $R_{n \times n}$  و بردار  $r(\mathbf{x})$  را که به ترتیب شامل مقادیر همبستگی بین نقاط طرح، و همبستگی بین نقاط طرح و نقطه  $\mathbf{x}$  هستند را به صورت زیر در نظر بگیرید:

$$\mathbf{R} = (R(\mathbf{s}_i, \mathbf{s}_j))_{1 \leq i, j \leq n},$$

$$\mathbf{r}(\mathbf{x})^T = [R(\mathbf{s}_1, \mathbf{x}), \dots, R(\mathbf{s}_n, \mathbf{x})].$$

ماتریس طرح توسعه یافته با بعد  $n \times k$  نیز به صورت زیر تعریف می‌شود:

$$\mathbf{F} = [f_l(\mathbf{s}_i)]_{1 \leq i \leq n, 1 \leq l \leq k}.$$

اکنون اگر  $\mathbf{w}(\mathbf{x})^T \mathbf{y}_S$  یک پیش‌گوی خطی برای  $y(\mathbf{x})$  باشد، میانگین مربع خطای پیش‌گویی ( $MSPE$ )

عبارت است از

$$\begin{aligned} E[\mathbf{w}(\mathbf{x})^T \mathbf{Y}_S - Y(\mathbf{x})]^2 &= E \left[ \sum_{i=1}^n w_i \left[ \sum_{j=1}^k \beta_j f_j(\mathbf{s}_i) + Z(\mathbf{s}_i) \right] - \left[ \sum_{j=1}^k \beta_j f_j(\mathbf{x}) + Z(\mathbf{x}) \right] \right]^2 \\ &= E \left[ \sum_{i=1}^n \sum_{m=1}^n w_i w_m Z(\mathbf{s}_i) Z(\mathbf{s}_m) + Z(\mathbf{x})^2 - 2 \sum_{i=1}^n w_i Z(\mathbf{s}_i) Z(\mathbf{x}) \right] \\ &= \sigma^2 [\mathbf{w}(\mathbf{x})^T \mathbf{R} \mathbf{w}(\mathbf{x}) + 1 - 2 \mathbf{w}(\mathbf{x})^T \mathbf{r}(\mathbf{x})], \end{aligned} \quad (15.3)$$

که اندیس‌های  $i$  و  $m$  برای مجموعه  $n$  نقطه طرح و اندیس  $j$  برای  $k$  تابع برآورد در نظر گرفته شده‌اند.

در حالت تعمیم یافته، قید نارایی به صورت زیر به دست می‌آید:

$$\begin{aligned}
 E[\mathbf{w}^T Y_{\mathbf{S}}] - E[Y(\mathbf{x})] &= E\left[\sum_{i=1}^n w_i Y(\mathbf{s}_i)\right] - E\left[\sum_{l=1}^k \beta_l f_l(\mathbf{x}) + Z(\mathbf{x})\right] \\
 &= E\left[\sum_i w_i \left[\sum_l \beta_l f_l(\mathbf{s}_i) + Z(\mathbf{s}_i)\right]\right] - E\left[\sum_l \beta_l f_l(\mathbf{x})\right] \quad (16.3) \\
 &= \sum_l \beta_l \sum_i w_i(\mathbf{x}) f_l(\mathbf{s}_i) - \sum_l \beta_l f_l(\mathbf{x}) \\
 &= \sum_l \beta_l \left(\sum_i w_i f_l(\mathbf{s}_i) - f_l(\mathbf{x})\right) = 0,
 \end{aligned}$$

که منجر به شرط نارایی  $\sum_i w_i f_l(\mathbf{s}_i) = f_l(\mathbf{x})$  می‌شود و بیان ماتریسی آن عبارتست از  $\mathbf{F}^T \mathbf{w}(\mathbf{x}) = \mathbf{f}(\mathbf{x})$ . برای به دست آوردن پیش‌گوی کریگینگ، باید  $MSPE$  با قید نارایی کمینه شود. این قید روی پیش‌گو باعث می‌شود که از درونیابی نقاط مورد نظر، اطمینان حاصل شود. قید مورد نظر با استفاده از ضرایب لاگرانژ  $\lambda(\mathbf{x})$  به معادله اضافه می‌شود:

$$\begin{aligned}
 L(w_i(\mathbf{x}), \lambda(\mathbf{x})) &= \sigma^2 [\mathbf{w}(\mathbf{x})^T \mathbf{R} \mathbf{w}(\mathbf{x}) + \mathbf{1} - \mathbf{2} \mathbf{w}(\mathbf{x})^T \mathbf{r}(\mathbf{x})] \\
 &\quad + \mathbf{2} \lambda(\mathbf{x})^T (\mathbf{F}^T \mathbf{w}(\mathbf{x}) - \mathbf{f}(\mathbf{x})). \quad (17.3)
 \end{aligned}$$

با گرفتن مشتقات جزئی از معادله (17.3) نسبت به  $w_i$  ( $i = 1, \dots, n$ ) و  $\lambda_l$  ( $l = 1, \dots, k$ )، و برابر صفر قرار دادن آن‌ها، دستگاه کریگینگ عام به صورت زیر تولید می‌شود:

$$\begin{cases} \sigma^2 \sum_j w_j R(\mathbf{s}_i, \mathbf{s}_j) + \sum_l \lambda_l f_l(\mathbf{s}_i) = \sigma^2 R(\mathbf{s}_i, \mathbf{x}), & i = 1, \dots, n \\ \sum_i w_i f_l(\mathbf{s}_i) = f_l(\mathbf{x}), & l = 1, \dots, k. \end{cases} \quad (18.3)$$

فرم ماتریسی این دستگاه معادلات به صورت زیر خواهد بود،

$$\begin{bmatrix} \mathbf{O} & \mathbf{F}^T \\ \mathbf{F} & \mathbf{R} \end{bmatrix} \begin{bmatrix} \lambda(\mathbf{x}) \\ \mathbf{w}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \mathbf{f}(\mathbf{x}) \\ \mathbf{r}(\mathbf{x}) \end{bmatrix}. \quad (19.3)$$

برای محاسبه ضرایب  $\mathbf{w}(\mathbf{x})$  در BLUP به صورت زیر عمل می‌کنیم،

$$\begin{bmatrix} \lambda(\mathbf{x}) \\ \mathbf{w}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \mathbf{O} & \mathbf{F}^T \\ \mathbf{F} & \mathbf{R} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{f}(\mathbf{x}) \\ \mathbf{r}(\mathbf{x}) \end{bmatrix}. \quad (20.3)$$

اما وارون ماتریس افراز شده فوق برابر است با: (هایلاک، ۱۹۹۷)

$$\Lambda = \begin{bmatrix} \mathbf{O} & \mathbf{F}^T \\ \mathbf{F} & \mathbf{R} \end{bmatrix}^{-1} = \begin{bmatrix} -(F^T R^{-1} F)^{-1} & (F^T R^{-1} F)^{-1} F^T R^{-1} \\ R^{-1} F (F^T R^{-1} F)^{-1} & R^{-1} (I - F (F^T R^{-1} F)^{-1} F^T R^{-1}) \end{bmatrix}$$

لذا رابطه (20.3) به فرم زیر حاصل می‌شود

$$\begin{bmatrix} \boldsymbol{\lambda}(\mathbf{x}) \\ \mathbf{w}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} -(F^T R^{-1} F)^{-1} & (F^T R^{-1} F)^{-1} F^T R^{-1} \\ R^{-1} F (F^T R^{-1} F)^{-1} & R^{-1} (I - F (F^T R^{-1} F)^{-1} F^T R^{-1}) \end{bmatrix} \begin{bmatrix} \mathbf{f}(\mathbf{x}) \\ \mathbf{r}(\mathbf{x}) \end{bmatrix}.$$

در نتیجه خواهیم داشت

$$\mathbf{w}(\mathbf{x}) = \begin{bmatrix} \Lambda_{\gamma_1} & \Lambda_{\gamma_2} \end{bmatrix} \begin{bmatrix} \mathbf{f}(\mathbf{x}) \\ \mathbf{r}(\mathbf{x}) \end{bmatrix}$$

لذا

$$\mathbf{w}(\mathbf{x})^T = \begin{bmatrix} \mathbf{f}(\mathbf{x})^T & \mathbf{r}(\mathbf{x})^T \end{bmatrix} \begin{bmatrix} \Lambda_{\gamma_1}^T \\ \Lambda_{\gamma_2}^T \end{bmatrix}$$

قابل ذکر است که

$$\Lambda_{\gamma_1}^T = \left( (F^T R^{-1} F)^{-1} \right)^T F^T R^{-1T} = (F^T R^{-1} F)^{-1} F^T R^{-1}$$

$$\Lambda_{\gamma_2}^T = \left( R^{-1} (I - F (F^T R^{-1} F)^{-1} F^T R^{-1}) \right)^T = R^{-1} (I - F (F^T R^{-1} F)^{-1} F^T R^{-1})$$

بنابراین پیش‌گو به صورت زیر خواهد بود

$$\hat{y}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \hat{\boldsymbol{\beta}} + \mathbf{r}(\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y}_S - \mathbf{F} \hat{\boldsymbol{\beta}}), \quad (21.3)$$

که

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{y}_S, \quad (22.3)$$

برآوردگر حداقل توان‌های دوم تعمیم‌یافته  $\beta$  در معادله (۱۰.۳) است. چون برای به دست آوردن پیش‌گوی  $\hat{y}(\mathbf{x})$ ، معکوس کردن ماتریس  $\mathbf{R}$  ضروری است، لذا تابع  $R(\cdot, \cdot)$  باید به گونه‌ای انتخاب شود تا وارون‌پذیر باشد.

برخی نتایجی که از پیش‌گوی کریگینگ به دست می‌آیند، به شرح زیر هستند:

۱. معادله (۲۱.۳) نشان می‌دهد که در روش کریگینگ باید سه مورد زیر، از قبل مشخص شود.

الف- نقاط طرح آزمایش  $s_1, \dots, s_n$ .

ب- صورت تابعی همبستگی نقاط طرح که منجر به تعیین ماتریس کوواریانس  $\mathbf{R}$  می‌شود.

ج- صورت تابعی  $f_1(\cdot), \dots, f_k(\cdot)$  در رابطه (۱۰۳).

این پیش‌نیازها به معنای ضرورت داشتن یک طرح آزمایش خوب و مناسب برای به‌دست آوردن مقادیر پاسخ مربوطه می‌باشد. پیش‌گوی کریگینگ با روش GLS برای برآورد کردن ضرایب  $\beta$  مطابقت دارد. GLS یک روش استوار است و همچنین می‌تواند ناهم‌واریانسی<sup>۱</sup> را در نظر بگیرد.

۲. پیش‌گوی کریگینگ شامل دو بخش است، نخست برآوردگر  $\mathbf{f}(\mathbf{x})^T \hat{\beta}$  بر مبنای برآوردگر حداقل مربعات تعمیم‌یافته  $\hat{\beta} = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{y}_S$  و دیگری بخش تصحیح وزنی خطاهای پیش‌گویی برای این برآوردگر است، که  $(\mathbf{y}_S - \mathbf{F} \hat{\beta})$  بردار خطای پیش‌گویی است. اگر  $\mathbf{r}(\mathbf{x}) \mathbf{R}^{-1}$  به عنوان ضریب وزنی بردار خطای پیش‌گویی عنوان شود، آنگاه همبستگی بزرگ‌تر بین یک نقطه پیش‌گویی  $(\mathbf{x})$  و یک نقطه طرح آزمایش، ضریب وزنی بزرگ‌تری را نتیجه می‌دهد.

۳. اگر نقطه پیش‌گویی از نقاط طرح دور باشد، همبستگی بین آن و نقاط طرح خیلی ضعیف است، یعنی  $\mathbf{r}(\mathbf{x})$  خیلی کوچک می‌شود که در نتیجه عبارت  $\mathbf{r}(\mathbf{x}) \mathbf{R}^{-1} (\mathbf{y}_S - \mathbf{F} \hat{\beta})$  نزدیک به صفر خواهد بود. واضح است کریگینگ، بسیار مشابه با رگرسیون عمل می‌کند، یعنی  $\mathbf{f}(\mathbf{x})^T \hat{\beta}$  مقدار  $Y(\mathbf{x})$  را برآورد می‌کند. در این مورد، کریگینگ فقط قادر به برآورد میانگین نقطه پیش‌گویی است.

### ۴.۶.۳ واریانس کریگینگ عام

با استفاده از معادلات دستگاه کریگینگ عام در (۱۸.۳)، واریانس کریگینگ عام به‌صورت زیر محاسبه

می‌شود:

$$\begin{aligned} \sigma^2[\mathbf{w}(\mathbf{x})^T \mathbf{R} \mathbf{w}(\mathbf{x})] &= \sigma^2 \sum_{i=1}^n \sum_{j=1}^n w_i w_j R(\mathbf{s}_i, \mathbf{s}_j) = \sum_{i=1}^n w_i \left( \sigma^2 \sum_{j=1}^n w_j R(\mathbf{s}_i, \mathbf{s}_j) \right) \\ &= \sum_{i=1}^n w_i \left( \sigma^2 R(\mathbf{s}_i, \mathbf{x}) - \sum_{l=1}^k \lambda_l f_l(\mathbf{s}_i) \right) \\ &= \sigma^2 \mathbf{w}(\mathbf{x})^T \mathbf{r}(\mathbf{x}) - \lambda(\mathbf{x})^T \mathbf{f}(\mathbf{x}). \end{aligned} \quad (22.3)$$

با جایگذاری (۲۳.۳) در (۱۵.۳) داریم،

$$\sigma_{UK}^2 = \sigma^2 - \sigma^2 \mathbf{w}(\mathbf{x})^T \mathbf{r}(\mathbf{x}) - \lambda(\mathbf{x})^T \mathbf{f}(\mathbf{x}). \quad (24.3)$$

<sup>۱</sup>Heteroscedasticity

## ۷.۳ دیدگاه بیزی

اکنون پیش‌گوی مدل فرآیند تصادفی را در چارچوب بیزی در نظر می‌گیریم. در اینجا فرآیند تصادفی (۱۰۳) به صورت یک فرآیند پیشین در نظر گرفته می‌شود که آگاهی ما را درباره کد کامپیوتری قبل از ساختن پیش‌گو نمایش می‌دهد. مانند حالت کلاسیک فرض بر این است که فرآیند تصادفی  $Z(\cdot)$  ایستا و دارای توزیع نرمال است. بردار پاسخ‌های  $y_S$ ، را که دارای توزیع نرمال چندمتغیره

$$y_S | \sigma^2, \beta \sim N(\mathbf{F}\beta, \mathbf{R}) \quad (25.3)$$

است، در نظر می‌گیریم. میانگین فرآیند پسین با استفاده از داده‌ها و پیشینی که به صورت پیش‌گوی نقطه‌ای برای کد کامپیوتری استفاده شده است، به دست می‌آید. چون پیشین دارای توزیع نرمال و داده‌ها نیز نرمال چندمتغیره هستند، پس پسین نیز دارای توزیع نرمال خواهد بود که به دست آوردن فرآیند پسین با استفاده از تکنیک‌های استاندارد امکان‌پذیر است. میانگین این فرآیند پسین، پیش‌گوی بیزی خواهد بود و آن را می‌توان با شرط معلوم بودن  $\sigma^2$  و  $\beta$  به صورت زیر نوشت (هایلاک، ۱۹۹۷)

$$\mathbf{f}(\mathbf{x})^T \beta + \mathbf{r}(\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y}_S - \mathbf{F}\beta). \quad (26.3)$$

اگر  $\beta$  نامعلوم باشد با قراردادن یک پیشین ناسره استاندارد روی بردار  $\beta$ ، میانگین فرآیند پسین به صورت زیر حاصل می‌شود

$$\mathbf{f}(\mathbf{x})^T \hat{\beta} + \mathbf{r}(\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y}_S - \mathbf{F}\hat{\beta}), \quad (27.3)$$

که  $\hat{\beta}$  در (۲۲.۳) آمده است. این پیش‌گو همان کریگینگ تعریف شده در تحلیل کلاسیک است. همان‌طور که دیده می‌شود تابع همبستگی باید طوری انتخاب شود که ماتریس  $R$  معکوس‌پذیر باشد، در غیر این صورت محاسبه پیش‌گوی کریگینگ یا میانگین پسین امکان‌پذیر نیست. در واقع انتخاب توابع همبستگی محدود به این است که ماتریس  $R$  دارای درمینان غیرصفر یا معین مثبت باشد. همچنین فرض شده است که واریانس فرآیند، و پارامتر هموارسازی معلوم هستند. متأسفانه این فرض معمولاً غیرواقعی است و در نتیجه تحلیل‌هایی که برآورد این پارامترها را انجام دهد، مورد توجه قرار گرفته‌اند.



## ۱.۷.۳ توسعه مدل بیزی

فرض کنید  $\eta(\cdot)$  نشان دهنده کد کامپیوتری پیچیده و  $\eta(\mathbf{X})$  و  $\eta(\mathbf{Z})$ ، خروجی‌های معلوم کد کامپیوتری به ترتیب به ازای  $\mathbf{X} = (x_1, x_2, \dots, x_d)$  و  $\mathbf{Z} = (z_1, z_2, \dots, z_d)$  باشند. در دیدگاه بیزی، برای ارزیابی یک نقطه، آگاهی پیشین درباره  $\eta(\cdot)$  به صورت یک مدل فرآیند تصادفی خواهد بود.

ابتدا فرض می‌کنیم تابع  $\eta(\cdot)$  می‌تواند به صورت یک ترکیب خطی از  $k$  تابع ساده  $h_j(\cdot)$ ،  $j = 1, \dots, k$  تقریب شود. آنگاه امید و واریانس  $\eta(\mathbf{X})$  برای همه  $\mathbf{X}$ ها در فضای مقادیر ورودی  $\mathcal{X}$  را به صورت

$$E(\eta(\mathbf{X}) | \beta, \sigma^2) = \mathbf{h}(\mathbf{X})^T \beta \quad (28.3)$$

$$\text{Var}(\eta(\mathbf{X}) | \beta, \sigma^2) = \sigma^2, \quad (29.3)$$

تعریف خواهیم کرد، که  $\beta = (\beta_1, \beta_2, \dots, \beta_k)^T$  بردار ضرایب رگرسیونی نامعلوم،  $\sigma^2$  پارامتر نامعلومی است که تغییرپذیری  $\eta(\cdot)$  حول میانگینش را کمی می‌کند و  $\mathbf{h}(\mathbf{X}) = (h_1(\mathbf{X}), h_2(\mathbf{X}), \dots, h_k(\mathbf{X}))^T$  بردار توابع رگرسیونی منتخب را نشان می‌دهد.

کوواریانس بین  $\eta(\mathbf{X})$  و  $\eta(\mathbf{Z})$  به صورت

$$\text{Cov}(\eta(\mathbf{X}), \eta(\mathbf{Z})) = \sigma^2 C(\mathbf{X}, \mathbf{Z}), \quad (30.3)$$

تعریف می‌شود که  $C(\cdot, \cdot)$  تابع همبستگی است به طوری که،  $C(\mathbf{X}, \mathbf{X}) = 1$  همچنین  $C(\mathbf{X}, \mathbf{Z})$  یک تابع اکیدا نزولی نسبت به فاصله بین  $\mathbf{X}$  و  $\mathbf{Z}$  است. این تابع کوواریانس مهمترین بخش مدل بیزی است، زیرا همواری  $\eta(\cdot)$  نسبت به  $\mathbf{X}$  را که ایده اصلی در این مدل است، نشان می‌دهد.

توزیع پیشین توام  $\eta(\cdot)$  به شرط بردار  $\beta$  و پارامتر  $\sigma^2$  در  $n$  نقطه متناهی  $\eta(\mathbf{s}_1), \eta(\mathbf{s}_2), \dots, \eta(\mathbf{s}_n)$  نرمال چندمتغیره است. لذا، توزیع پیشین  $\eta(\cdot)$  به شرط  $(\beta, \sigma^2)$  یک فرآیند گاوسی است که به صورت

$$\eta(\cdot) | \beta, \sigma^2 \sim N(\mathbf{h}(\cdot)^T \beta, \sigma^2 C(\cdot, \cdot)) \quad (31.3)$$

بیان می‌شود، که  $\mathbf{h}(\cdot)^T \beta$  میانگین فرآیند و  $\sigma^2 C(\cdot, \cdot)$  تابع کوواریانس است.

در نهایت باید توزیع پیشین را برای ابرپارامترهای  $\beta$  و  $\sigma^2$  تعیین کنیم. اگر اطلاعات پیشین در

دسترس، مناسب باشند، یک پیشین آمیخته برای  $(\beta, \sigma^2)$ ، دارای توزیع نرمال گاما معکوس

$$f(\beta, \sigma^2) \propto (\sigma^2)^{-(k+a+2)/2} \exp\left(\frac{-(\beta - \beta_0)^T \mathbf{V}^{-1}(\beta - \beta_0) + 2b}{2\sigma^2}\right) \quad (32.3)$$

خواهد بود، که در آن  $a$  و  $b$  معلوم‌اند،  $\beta_0$ ، میانگین  $\beta$  است و ماتریس  $V_{k \times k}$  شامل کوواریانس‌های  $Cov(\beta_i, \beta_j)_{1 \leq i, j \leq k}$  می‌باشد. رابطه (۳۲.۳) نشان می‌دهد که توزیع پیشین شرطی  $\sigma^2 | \beta$ ، نرمال چندمتغیره و توزیع پیشین حاشیه‌ای  $\sigma^2$  گاما معکوس است. اغلب اطلاعات پیشین محدودند و معمولا پیشین ناسره

$$f(\beta, \sigma^2) \propto \sigma^{-2} \quad (33.3)$$

جایگزین می‌شود، که می‌تواند از (۳۲.۳) با این فرض که عناصر  $\mathbf{V}$  به  $\infty$  میل کنند و  $b = 0$  و  $a = -k$  باشد به دست آید. مقایسه با بخش ۶.۳ نشان می‌دهد که (۳۱.۳) در واقع همان مدل فرآیند تصادفی تعریف شده در (۱.۳) است.

## ۸.۳ پیش‌گوی بیزی

فرض کنید کد کامپیوتری برای  $n$  نقطه طرح اجرا شود، آنگاه  $\mathbf{y}^T = [\eta(\mathbf{s}_1), \eta(\mathbf{s}_2), \dots, \eta(\mathbf{s}_n)]$  برداری شامل  $n$  خروجی  $\eta(\cdot)$  در  $n$  نقطه طرح می‌باشد. بردار  $\mathbf{y}$  برای منتج کردن توزیع پسین  $\eta(\cdot)$  استفاده خواهد شد و فرض می‌شود که این بردار به شرط  $\beta$  و  $\sigma^2$  دارای توزیع نرمال  $n$  متغیره است. به عبارتی

$$\mathbf{y} | \beta, \sigma^2 \sim N(\mathbf{H}\beta, \sigma^2 \mathbf{A}) \quad (34.3)$$

که در آن  $\mathbf{H} = [\mathbf{h}(\mathbf{s}_1), \dots, \mathbf{h}(\mathbf{s}_n)]^T$  و  $\mathbf{A}$  یک ماتریس متقارن با بعد  $n \times n$  است که  $[i, j]$  امین عنصر آن  $C(\mathbf{s}_i, \mathbf{s}_j)$  است. بنابراین توزیع پسین شرطی  $\beta | \mathbf{y}, \sigma^2$  و توزیع پسین حاشیه‌ای  $\sigma^2$ ، به صورت زیر توسط اوهگن (۲۰۰۱) به دست می‌آیند:

$$\beta | \sigma^2, \mathbf{y} \sim N(\hat{\beta}, \sigma^2 (\mathbf{H}^T \mathbf{A}^{-1} \mathbf{H})^{-1}), \quad (35.3)$$

و

$$\frac{(n - q - 2) \hat{\sigma}^2}{\sigma^2} | \mathbf{y} \sim \chi_{n-q}^2, \quad (36.3)$$

که

$$\hat{\beta} = (\mathbf{H}^T \mathbf{A}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{A}^{-1} \mathbf{y}, \quad (37.3)$$

$$\hat{\sigma}^2 = \frac{\mathbf{y}^T (\mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{H} (\mathbf{H}^T \mathbf{A}^{-1} \mathbf{H})^{-1} \mathbf{A}^{-1}) \mathbf{y}}{n - q - 2}, \quad (38.3)$$

و  $n$  طول بردار  $\mathbf{y}$  و  $q$  رتبه  $\mathbf{H}$  است.

توزیع پیش‌گوی  $\eta(\cdot)$  در سه مرحله به‌دست می‌آید. ابتدا توزیع پسین  $\eta(\cdot) | \beta, \sigma^2, \mathbf{y}$  باید از توزیع توام  $\eta(\cdot) | \beta, \sigma^2$  و  $\mathbf{y} | \beta, \sigma^2$  نتیجه شود، سپس حاصل ضرب  $\eta(\cdot) | \beta, \sigma^2, \mathbf{y}$  با  $\beta | \sigma^2, \mathbf{y}$  روی  $\beta$  انتگرال‌گیری می‌شود تا فرآیند نرمال زیر به‌دست آید،

$$\eta(\cdot) | \sigma^2, \mathbf{y} \sim N(m^*(\cdot), \sigma^2 C^*(\cdot, \cdot)), \quad (39.3)$$

که

$$m^*(\mathbf{x}) = \mathbf{h}(\mathbf{x})^T \hat{\beta} + \mathbf{t}(\mathbf{x})^T \mathbf{A}^{-1} (\mathbf{y} - \mathbf{H} \hat{\beta}), \quad (40.3)$$

$$C^*(\mathbf{x}, \mathbf{z}) = C(\mathbf{x}, \mathbf{z}) - \mathbf{t}(\mathbf{x})^T \mathbf{A}^{-1} \mathbf{t}(\mathbf{z}) + (\mathbf{h}(\mathbf{x})^T - \mathbf{t}(\mathbf{x})^T \mathbf{A}^{-1} \mathbf{H}) (\mathbf{H}^T \mathbf{A}^{-1} \mathbf{H})^{-1} (\mathbf{h}(\mathbf{z})^T - \mathbf{t}(\mathbf{z})^T \mathbf{A}^{-1} \mathbf{H})^T, \quad (41.3)$$

و

$$\mathbf{t}(\mathbf{x}) = [C(\mathbf{x}, \mathbf{s}_1), C(\mathbf{x}, \mathbf{s}_2), \dots, C(\mathbf{x}, \mathbf{s}_n)]^T. \quad (42.3)$$

در نهایت انتگرال‌گیری حاصل ضرب  $\eta(\cdot) | \sigma^2, \mathbf{y}$  و  $\sigma^2 | \mathbf{y}$  نسبت به  $\sigma^2$  یک توزیع پیش‌گو برای  $y | \eta(\cdot)$  برقرار خواهد کرد که برای یک  $\mathbf{x}$  داده شده می‌تواند به‌صورت زیر نوشته شود:

$$\frac{\eta(\mathbf{x}) - m^*(\mathbf{x})}{\hat{\sigma} \sqrt{C^*(\mathbf{x}, \mathbf{x})}} | \mathbf{y} \sim t_{n-q}. \quad (43.3)$$

این توزیع، توزیع  $t$  چندمتغیره است. پس برای مجموعه ورودی داده شده  $\mathbf{x}$ ، برآورد نقطه‌ای متغیر پاسخ توسط امید ریاضی توزیع پیش‌گو، یعنی  $m^*(\mathbf{x})$ ، انجام می‌شود و واریانس آن توسط  $\hat{\sigma}^2 C^*(\mathbf{x}, \mathbf{x})$  به‌دست می‌آید.

# فصل ۴

## عملکرد مدل جانشین کریگینگ بر تابع تحلیلی و مورد مطالعاتی

در این فصل عملکرد روش کریگینگ را به عنوان مدل جانشین برای یک تابع تحلیلی و همچنین یک مدل کامپیوتری واقعی بررسی نموده و دقت برآوردهای مونت کارلویی شاخص‌های حساسیت را مورد ارزیابی قرار می‌دهیم. کلیه برنامه‌ها توسط بسته‌های موجود در نرم‌افزار R-2.15.3 اجرا شده است که در هر قسمت به آن‌ها اشاره خواهیم نمود.

از آن‌جا که ساخت هرگونه مدل جانشین نیازمند داده‌های مدل‌ساز است، لذا در ابتدا موضوع تولید این داده‌ها که بخشی از سرفصل ”طرح آزمایش‌های مدل‌های کامپیوتری” است را مطرح می‌کنیم.

### ۱.۴ تولید داده‌های شبیه‌سازی

نمونه‌گیری ابرمکعب لاتین<sup>۱</sup> (LHS) برای تولید نمونه‌هایی از متغیرهای ورودی، انتخاب شده است. LHS توسط مک‌کی و همکاران<sup>۲</sup> (۱۹۷۹)، پیشنهاد شده است. آن‌ها نشان دادند که LHS واریانسی کوچک‌تر از نمونه‌گیری تصادفی ساده<sup>۳</sup> (SRS) دارد. ساکس و همکاران<sup>۴</sup> (۱۹۸۹) نشان دادند که LHS تعمیمی از نمونه‌گیری طبقه‌ای است که از نمونه‌برداری تمام بخش‌ها در محدوده تعریف‌شده هر متغیر

---

<sup>۱</sup>Latin Hypercube Sampling

<sup>۲</sup>McKay, Beckman and Conover

<sup>۳</sup>Simple Random Sampling

<sup>۴</sup>Sacks, Welch, Mitchell and Wynn

ورودی، اطمینان حاصل می‌کند. تولید نمونه کارا و نیز قابلیت اجرایی در فضای توام متغیرهای ورودی با بعد بالا، از مزایای اصلی LHS می‌باشند. فرض کنید می‌خواهیم یک نمونه  $N$  تایی از متغیرهای ورودی در فضای  $s$  متغیره تولید کنیم. مراحل تولید LHS براساس فنگ و سودجیان<sup>۱</sup> (۲۰۰۶)، به صورت زیر هستند:

۱. فرض کنید  $\pi_j(1), \dots, \pi_j(N)$  یک جایگشت تصادفی از اعداد  $1, \dots, N$  باشند. تعداد  $s$  جایگشت تصادفی  $(j = 1, \dots, s)$  را تولید کنید.

۲. تعداد  $s$  بردار مستقل  $N$  تایی  $U_1, U_2, \dots, U_s$  را از توزیع یکنواخت در  $(0, 1)$  تولید کنید که در آن  $U_j = (U_{1j}, \dots, U_{Nj})$  قرار دهید

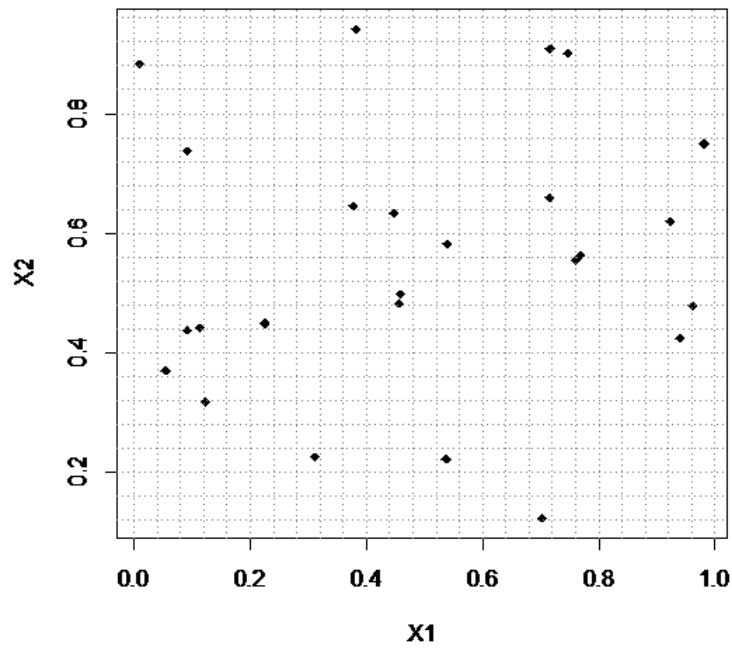
$$x_{kj} = \frac{\pi_j(k) - U_j}{N}, \quad k = 1, \dots, N; j = 1, \dots, s.$$

در این صورت  $0 \leq x_{kj} < 1$  خواهد بود. آنگاه  $D_N = \{\mathbf{X}_1, \dots, \mathbf{X}_N\}$  یک نمونه LHS است که با  $LHS(N, s)$  نشان داده می‌شود.

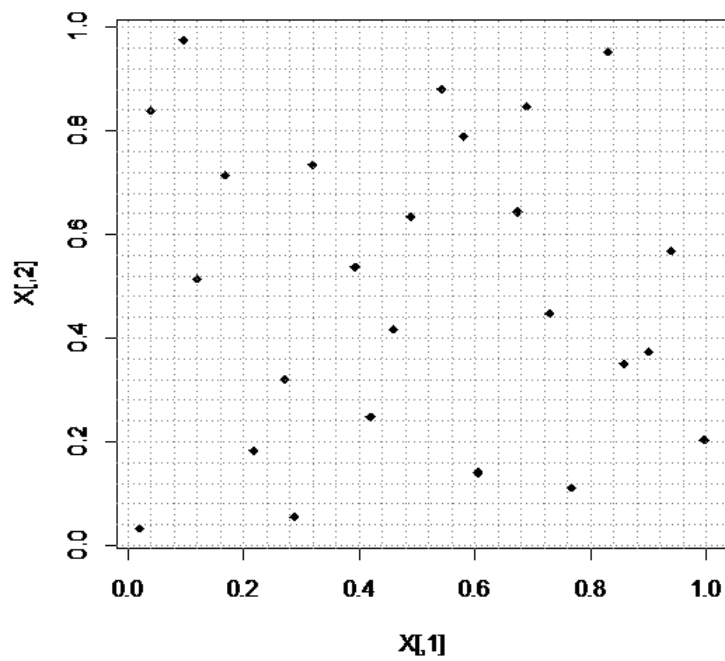
بسته lhs موجود در نرم افزار R دارای توابعی است که نمونه‌های LHS را تولید می‌کند. ما در این جا از تابع maximinLHS برای تولید نمونه‌های شبیه‌سازی استفاده کرده‌ایم. این تابع با بیشینه کردن کمترین فاصله‌های بین نقاط، نمونه را بهینه می‌کند.

در شکل ۱.۴ و ۲.۴، نمودار پراکنش تعداد ۲۵ نمونه تصادفی، به ترتیب بر اساس نمونه‌گیری تصادفی ساده و LHS، رسم شده‌اند. در شکل ۲.۴ هر ردیف و هر ستون فقط شامل یک نقطه تصادفی است، ولی در شکل ۱.۴ برخی از سطر و ستون‌ها شامل هیچ نقطه‌ای نیستند و برخی نیز شامل بیش از یک نقطه می‌باشند. واضح است که LHS از نمونه‌گیری زاید اجتناب می‌کند، شکاف‌ها را می‌پوشاند و به خوبی قادر به پوشش و کاوش در ابرفضاها است.

<sup>۱</sup>Fang and Sudjanto



شکل ۱.۴: نمودار پراکنش ۲۵ نمونه تصادفی تولیدشده توسط نمونه‌گیری تصادفی ساده



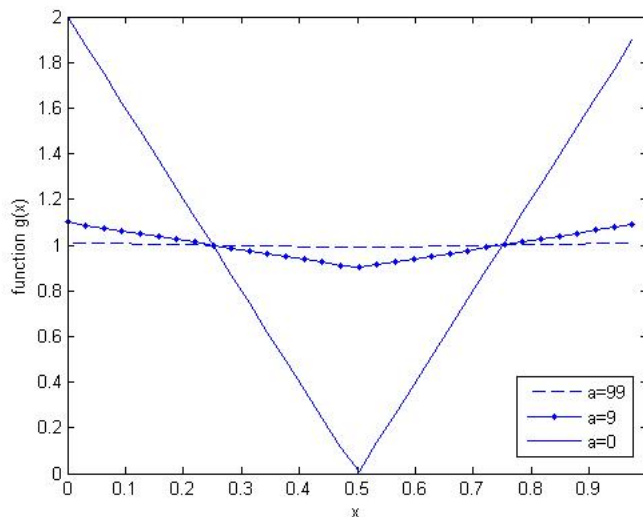
شکل ۲.۴: نمودار پراکنش ۲۵ نمونه  $LHS(25, 2)$

## ۲.۴ معرفی تابع تحلیلی سبل

تابع تحلیلی سبل، موسوم به  $g$ -تابع<sup>۱</sup>، که در اکثر مراجع مرتبط با تحلیل حساسیت و مدل‌های جانشین، مورد استفاده قرار گرفته است، دارای صورت کلی زیر می‌باشد:

$$Y = \prod_{i=1}^d g_i(X_i), \quad g_i(X_i) = \frac{|4X_i - 2| + a_i}{1 + a_i}; \quad 0 \leq X_i \leq 1,$$

که در آن  $X_i \sim U(0, 1), i = 1, \dots, d$  و  $a_i$ ها مقادیر ثابت و مثبتی هستند. نمودار این تابع به ازای مقادیر مختلف  $a_i$  در حالت یک و دو متغیره، به ترتیب در شکل‌های ۳.۴ و ۴.۴ نمایش داده شده است.



شکل ۳.۴: نمودار تابع سبل در حالت یک متغیره

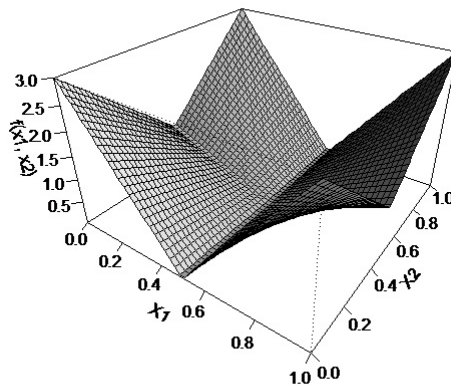
این تابع غیرخطی و غیریکنواست و امکان تغییر تعداد متغیرهای ورودی در این تابع وجود دارد. از

آن جایی که کران‌های بالا و پایین  $g_i(X_i)$  عبارتند از

$$1 - \frac{1}{1 + a_i} \leq g_i(X_i) \leq 1 + \frac{1}{1 + a_i}, \quad (1.4)$$

لذا درجه اهمیت هر یک از متغیرها توسط مقدار  $a_i$  تعیین می‌شود. هر چه متغیر مورد بررسی با اهمیت‌تر باشد، نوسانات تابع نیز در دامنه تغییر آن متغیر بیشتر خواهد بود. جدول ۱.۴ برخی از مقادیر  $a_i$  به همراه درجه اهمیت آن‌ها که به‌طور متداول در مراجع علمی مرتبط استفاده می‌شوند را نشان می‌دهد.

<sup>۱</sup>g-function



شکل ۴.۴: نمودار تابع سبل در حالت دو متغیره ( $a_1 = 0, a_2 = 1$ )

جدول ۱.۴: میزان اهمیت و مقادیر حقیقی ورودی‌های تابع سبل

شماره متغیر	$(a_i)$	میزان اهمیت	مقدار واقعی شاخص حساسیت مرتبه اول
$X_1$	۰	بسیار مهم	۰٫۷۱۶
$X_2$	۱	مهم	۰٫۱۷۹
$X_3$	۴٫۵	نسبتاً مهم	۰٫۰۲۴
$X_4$	۹	کم اهمیت	۰٫۰۰۷۲
$X_5$	۹۹	فاقد اهمیت	$۷٫۱۶۲ \times 10^{-5}$
$X_6$	۹۹	فاقد اهمیت	$۷٫۱۶۲ \times 10^{-5}$
$X_7$	۹۹	فاقد اهمیت	$۷٫۱۶۲ \times 10^{-5}$
$X_8$	۹۹	فاقد اهمیت	$۷٫۱۶۲ \times 10^{-5}$

ثابت می‌شود که مقدار واقعی شاخص حساسیت مرتبه اول متغیر  $X_i$  برابر با  $S_i = \frac{\nu_i}{V}$  و مقدار واقعی شاخص حساسیت کل  $X_i$  برابر با  $S_i^t = 1 - \frac{\sum_{j=1, j \neq i}^d \nu_j + \prod_{j=1, j \neq i}^d \nu_j}{V}$  است که در آن  $\nu_i = \frac{1}{3(1+a_i)^2}$  است (سالتلی و همکاران، ۲۰۰۸).

برای ارزیابی روش کریگینگ و مقایسه مقادیر واقعی و برآورد شده شاخص‌های حساسیت، تعداد متغیرهای ورودی تابع تحلیلی را برابر هشت ( $d = 8$ ) و میزان اهمیت متغیرها را نیز متفاوت در نظر گرفته‌ایم. جدول ۱.۴ مقادیر مفروض  $a_i$ ، میزان اهمیت و مقادیر واقعی شاخص حساسیت مرتبه اول این متغیرها را نشان می‌دهد.



### ۳.۴ روش جانشین کریگینگ برای تابع تحلیلی سبل

برای بررسی مناسبت روش کریگینگ در تقریب تابع سبل، حجم داده‌های مدل‌ساز و آزمون بر اساس طرح آزمایش LHS، به ترتیب برابر  $n_{train} = 250$  و  $n_{test} = 2000$  انتخاب شدند. با استفاده از بسته Dicekriging در نرم‌افزار R، برآوردگر درست‌نمایی ماکزیم (MLE) پارامترهای همواری  $(\theta_j)$ ها، با بیشینه ساختن تابع لگاریتم درست‌نمایی زیر حاصل می‌شود (باسبی و همکاران، ۲۰۰۷):

$$\ln L(\theta) \propto -1/2(n \ln \sigma^2 + \ln |\mathbf{R}|). \quad (2.4)$$

بیشینه ساختن (۲.۴) نیاز به بهینه‌سازی و محاسبات پرهزینه‌ای دارد که توسط برنامه‌های کامپیوتری انجام می‌گیرد. با استفاده از بسته Dicekriging در نرم‌افزار R، برآورد پارامترهای  $\beta$  و  $\sigma^2$  (رجوع شود به (۳۷.۳) و (۳۸.۳)) و پارامترهای  $\theta_j$  در جدول (۲.۴) نمایش داده شده‌اند.

جدول ۲.۴: برآورد پارامترهای مدل (۲.۳) با روش کریگینگ برای تابع سبل

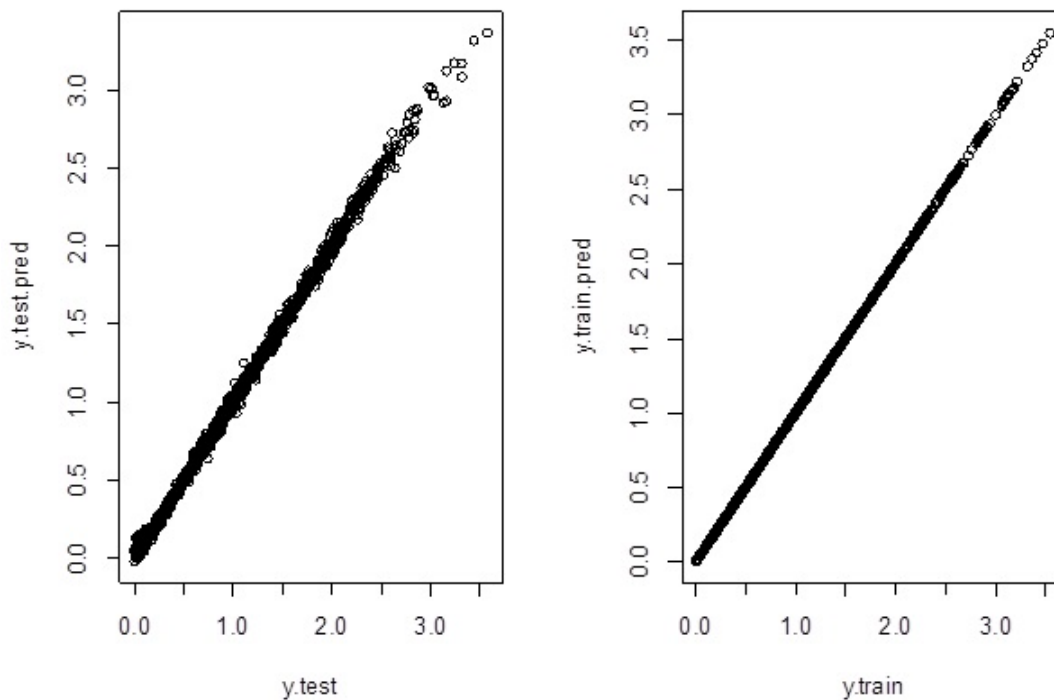
پارامتر	$\beta$	$\sigma^2$	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$	$\theta_6$	$\theta_7$	$\theta_8$
برآورد	۲٫۷	۰٫۲	۰٫۲۹	۰٫۴۵	۰٫۸۴	۱٫۱۶	۱٫۹۹	۱٫۹۹	۱٫۹۹	۱٫۹۹

همچنین با استفاده از بسته مذکور، مقادیر پیش‌بینی متغیر پاسخ برای داده‌های مدل‌ساز و آزمون نیز به دست آمد که در شکل ۵.۴ نمودار پراکنش مقایر برآورد کریگینگ در مقابل مقادیر واقعی نشان داده شده است. با توجه به ضریب تعیین مقادیر واقعی و برآورد شده داده‌های آزمون که برابر است با ۰٫۹۶۲۹، می‌توان دید مدل برآورد شده توانسته است، برای داده‌هایی که در ساخت مدل مشارکت ندارند نیز عملکرد خوبی داشته باشد.

راه دیگر برای ارزیابی عملکرد روش کریگینگ، استفاده از معیار «اعتبارسنجی متقابل با حذف هر بار یک مشاهده<sup>۱</sup> (LOOCV)» است. فرض کنید  $n$  مشاهده  $y_1, \dots, y_n$  را داریم. معیار LOOCV را به صورت زیر به دست می‌آوریم:

۱. فرض کنید تک مشاهده  $i$ ام، مجموعه آزمون را تشکیل دهد. مدل را با بقیه مشاهدات برازش

<sup>۱</sup>Leave-one-out cross-validation



شکل ۵.۴: راست: نمودار پراکنش داده‌های مدل‌ساز در برابر مقادیر پیش‌گویی شده آن‌ها، چپ: نمودار پراکنش داده‌های آزمون در برابر مقادیر پیش‌گویی شده آن‌ها

داده و خطای  $e_i^* = y_i - \hat{y}_i$  را برای مشاهده حذف شده محاسبه می‌کنیم. به این خطای پیش‌گویی نیز می‌گویند.

۲. مرحله ۱ را برای  $i = 1, \dots, n$  تکرار می‌کنیم.

۳. میانگین توان‌های دوم خطاهای پیش‌گویی<sup>۱</sup> (MSPE)  $e_1^*, \dots, e_n^*$  را محاسبه می‌کنیم که برابر است با

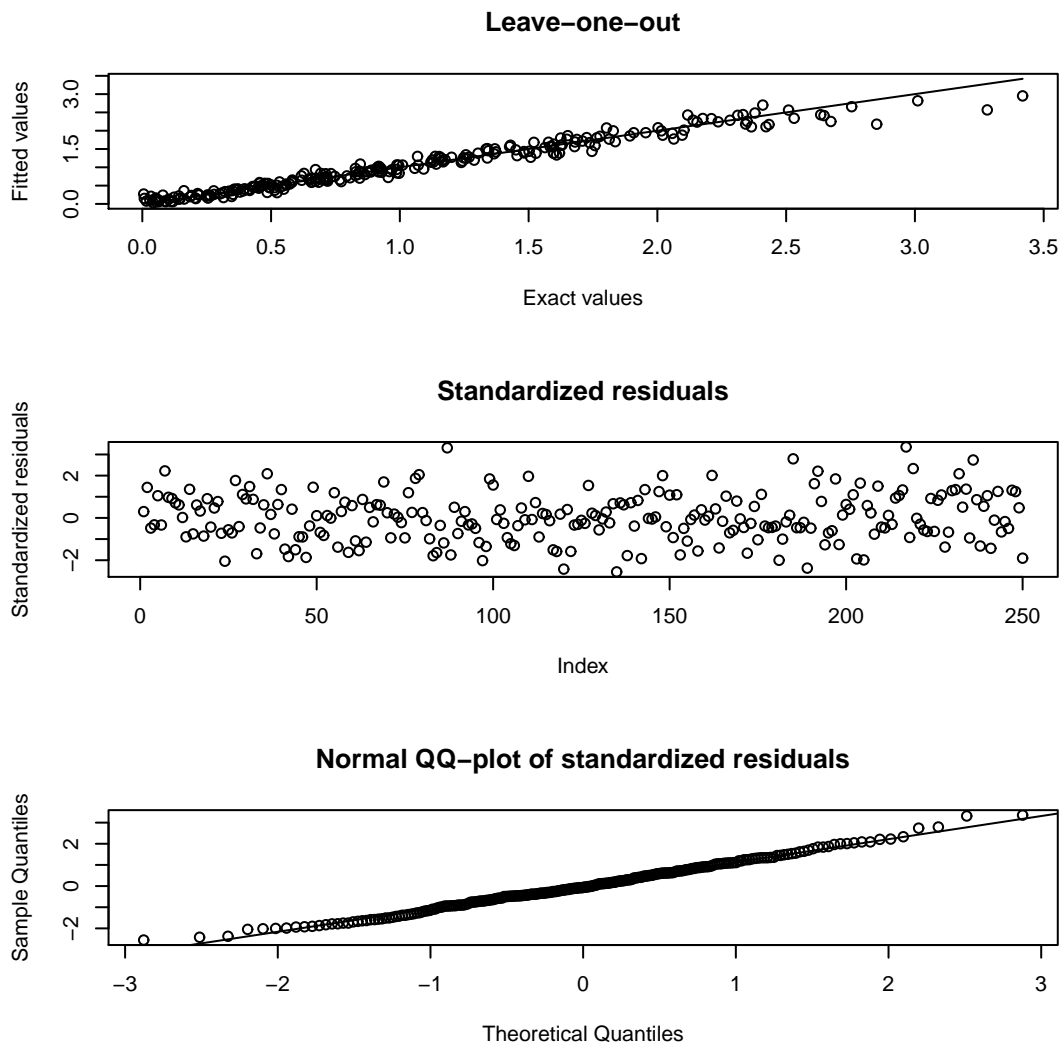
$$MSPE = \frac{1}{n} \sum_{i=1}^n e_i^{*2}.$$

هر چقدر این معیار کوچک‌تر باشد، قدرت پیش‌گویی مدل بهتر است.

معیار LOOCV را برای تابع سبل محاسبه کرده و اعتبار روش کریگینگ در تقریب این تابع در شکل

۶.۴ نشان داده شده است. با حذف هر بار یک مشاهده، مقدار  $y_i$  که در برازش مدل مرحله  $i$ ام استفاده

<sup>۱</sup>Minimom Square Predictor Error



شکل ۶.۴: اعتبارسنجی متقابل  $LOO$  برای تابع سبل

نشده است را در مقابل  $\hat{y}_i$  که از مدل برازش یافته مرحله  $i$ ام به دست آمده را برای  $i = 1, \dots, 250$  در مقابل هم رسم می‌کنیم. نمودار بالایی شکل ۶.۴ این پراکنش را نشان می‌دهد. این نمودار نشان می‌دهد که مقادیر برآورد شده به مقادیر واقعی بسیار نزدیک هستند و انحراف قابل توجهی از هم ندارند و این، مناسب بودن مدل برازش یافته به روش کریگینگ را نشان می‌دهد. باقیمانده‌های استاندارد در این ۲۵۰ نقطه، توسط رابطه‌ی  $[y(x_i) - \hat{y}(x_i)] / \{MSE[\hat{y}(x_i)]\}^{\frac{1}{2}}$  در قسمت میانی شکل ۶.۴، نشان داده شده است. این نمودار نیز حاکی از صحیح بودن مدل برازش یافته کریگینگ می‌باشد زیرا هیچ‌گونه طرح و الگوی خاصی در این نمودار مشاهده نمی‌شود. نمودار پایینی شکل ۶.۴، نمودار چندک چندک باقیمانده‌ها است که نشان می‌دهد باقیمانده‌های استاندارد شده، تقریباً نرمال استاندارد هستند. قرار گرفتن اکثر این

باقیمانده‌ها روی خط نشان از این دارد که برآورد انجام شده معتبر است.

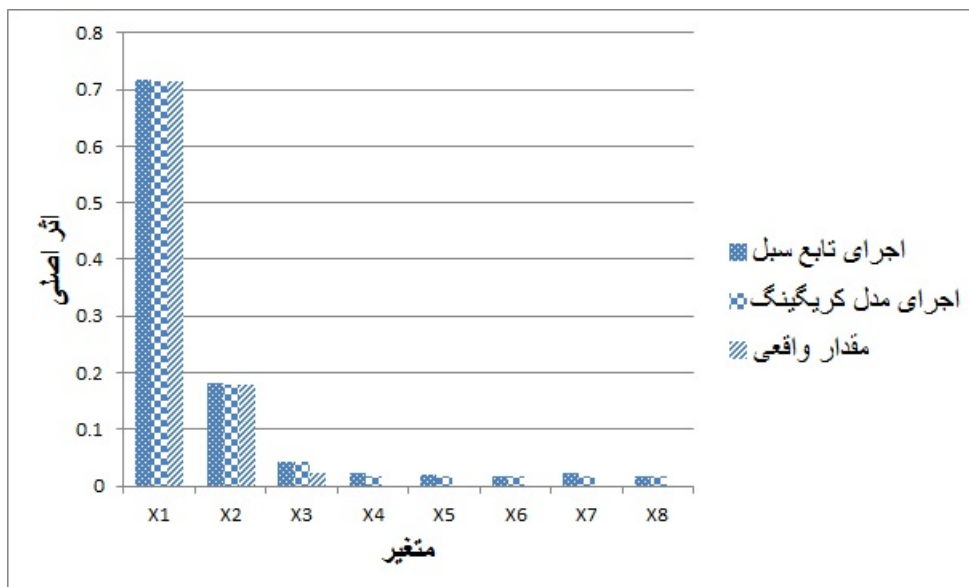
از تحلیل شکل‌های ۵.۴ و ۶.۴ درمی‌یابیم که روش کریگینگ به‌عنوان تقریب جانشین تابع سبل معتبر است. این معتبر بودن روش کریگینگ در جانشینی بسیار مفید است و آنرا پرکاربرد می‌سازد. یکی از کاربردهای این جانشینی، در تحلیل حساسیت است و سبب اجزاهای بسیار ارزان‌تر نسبت به مدل واقعی می‌شود. برای نشان دادن مفید بودن این جانشینی، تحلیل حساسیت تابع سبل را با این مدل جانشین بررسی می‌کنیم.

تقریب کریگینگ تابع سبل، با یک مجموعه ۱۰۰۰ تایی از داده‌های آزمون اجرا شده و برآوردهای شاخص حساسیت اصلی (اثر اصلی) و شاخص حساسیت کل (اثر کل) تابع سبل در دو حالت، مدل اصلی و تقریب کریگینگ، به ترتیب در جدول ۳.۴ و ۴.۴ آمده‌اند. محاسبات این بخش توسط تابع sensitivity انجام گرفته است.

جدول ۳.۴: مقادیر واقعی و مقادیر برآوردشده اثر اصلی

متغیر	برآورد اثر اصلی با تابع سبل	برآورد اثر اصلی با تقریب کریگینگ	مقدار واقعی اثر اصلی
$X_1$	۰/۷۱۶۶	۰/۷۱۶۱	۰/۷۱۶
$X_2$	۰/۱۸۲۶	۰/۱۸۰۳	۰/۱۷۹
$X_3$	۰/۰۴۴۲	۰/۰۴۲۹	۰/۰۲۴
$X_4$	۰/۰۲۳۶	۰/۰۱۹۳	۰/۰۰۷۲
$X_5$	۰/۰۲۰۵	۰/۰۱۸۵	$۷/۱۶۲ \times ۱۰^{-۵}$
$X_6$	۰/۰۱۹۰	۰/۰۱۸۷	$۷/۱۶۲ \times ۱۰^{-۵}$
$X_7$	۰/۰۲۲۴	۰/۰۱۹۳	$۷/۱۶۲ \times ۱۰^{-۵}$
$X_8$	۰/۰۱۹۰	۰/۰۱۷۹	$۷/۱۶۲ \times ۱۰^{-۵}$

نمودارهای میله‌ای ۷.۴ و ۸.۴ مقادیر شاخص‌های حساسیت گزارش شده در جداول ۳.۴ و ۴.۴ را نمایش می‌دهند. با توجه به این شواهد می‌توان گفت که کریگینگ نه تنها برای مدل‌بندی مناسب است بلکه در برآورد شاخص‌های حساسیت متغیرهای با اهمیت بالا نیز بسیار دقیق است.



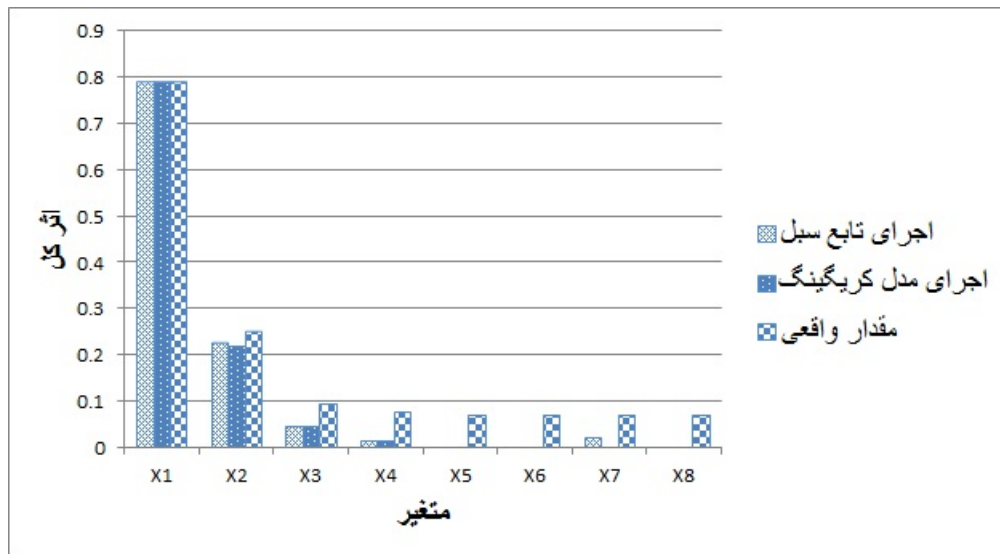
شکل ۷.۴: نمودار شاخص حساسیت مرتبه اول متغیرهای تابع سبیل

جدول ۴.۴: مقادیر واقعی و مقادیر برآوردشده اثر کل

متغیر	برآورد اثر کل با تقریب کریگینگ	برآورد اثر کل با تابع سبیل	مقدار واقعی اثر کل
$X_1$	۰/۷۸۹۳	۰/۷۹۰۱	۰/۷۸۹۹
$X_2$	۰/۲۲۳۷	۰/۲۲۷۳	۰/۲۵۲۷
$X_3$	۰/۰۴۸۰	۰/۰۵۰۲	۰/۰۹۷۳
$X_4$	۰/۰۱۸۳	۰/۰۱۷۵	۰/۰۸۰۸
$X_5$	۰/۰۰۲۰	۰/۰۰۱۲	۰/۰۷۳۷
$X_6$	۰/۰۰۰۲	۰/۰۰۰۸	۰/۰۷۳۷
$X_7$	۰/۰۰۱۵	۰/۰۲۳۱	۰/۰۷۳۷
$X_8$	۰/۰۰۱۰	۰/۰۰۰۰۷	۰/۰۷۳۷

#### ۴.۴ معرفی اجمالی مدل واقعی $INCA - N$

نیترات یکی از مشتقات نیتروژن است که در طی فرآیندهای تبدیل نیتروژن در طی چرخه این ماده، تولید می‌شود. نیترات از طریق تماس منابع آب با فاضلاب‌های شهری و صنعتی، تراوش شیرابه کود زمین‌های کشاورزی به رودخانه‌ها و از همه مهم‌تر اکسیداسیون مواد آلی ازت‌دار نظیر پروتئین‌ها وارد منابع آب می‌شود. نیترات به‌طور مستقیم و غیر مستقیم در اثر تجزیه و تغییرات زیست-شیمیایی ترکیب‌های مختلف معدنی و آلی در خاک به وجود می‌آید و به راحتی توسط آب باران به منابع آب سطحی و



شکل ۸.۴: نمودار شاخص حساسیت کل تابع سیل

زیرزمینی نفوذ می‌کند و به ترکیبات سمی چون نیتريت و نیتروسامین تبدیل می‌شود. بیماری‌هایی مانند متاهموگلوبین، سرطان، ناقص‌الخلقه بودن نوزادان، بزرگ شدن تیروئید و غیره از اثرات سوء نیترات بیش از حد در آب آشامیدنی است. با توجه به آثار سوء نیتروژن بر سلامت انسان، دانستن فرایند پیچیده چرخه نیتروژن، امری اجتناب‌ناپذیر است. شکل ۹.۴ چرخه نیتروژن و نحوه پدید آمدن نیترات را نشان می‌دهد. مدل INCA-N<sup>۱</sup> به شبیه‌سازی چرخه نیتروژن در طبیعت می‌پردازد. این مدل فرایند-مبنا<sup>۲</sup> است و با استفاده از معادلات جنبشی واکنش‌ها، شبیه‌سازی سازوکارهای اصلی عملیاتی رودخانه را انجام می‌دهد. مهمترین ویژگی‌های این الگو، پویایی، تصادفی و نیمه‌توزیعی<sup>۳</sup> است (وایتهد<sup>۴</sup> و همکاران، ۱۹۹۸). این مدل تعیینی میزان جریان آب و نیتروژن را در یک حوضه آبریز تعیین می‌کند. همچنین عوامل و فرایندهای تاثیرگذار بر میزان نیترات ( $NO_3$ ) و آمونیاک ( $NH_4$ ) ذخیره‌شده در خاک و سیستم‌های آب زیرزمینی را شبیه‌سازی نموده و خروجی را با استفاده از این سیستم در یک مدل رودخانه چندناحیه‌ای تولید می‌کند (وید<sup>۵</sup>، ۲۰۰۲). ساختار این مدل در شکل ۱۰.۴ نشان داده شده است. الگوی INCA-N برای نشان دادن تاثیر منابعی مانند شهرنشینی، تغییرات آب و هوایی، مدیریت اجرایی، سیاست‌های

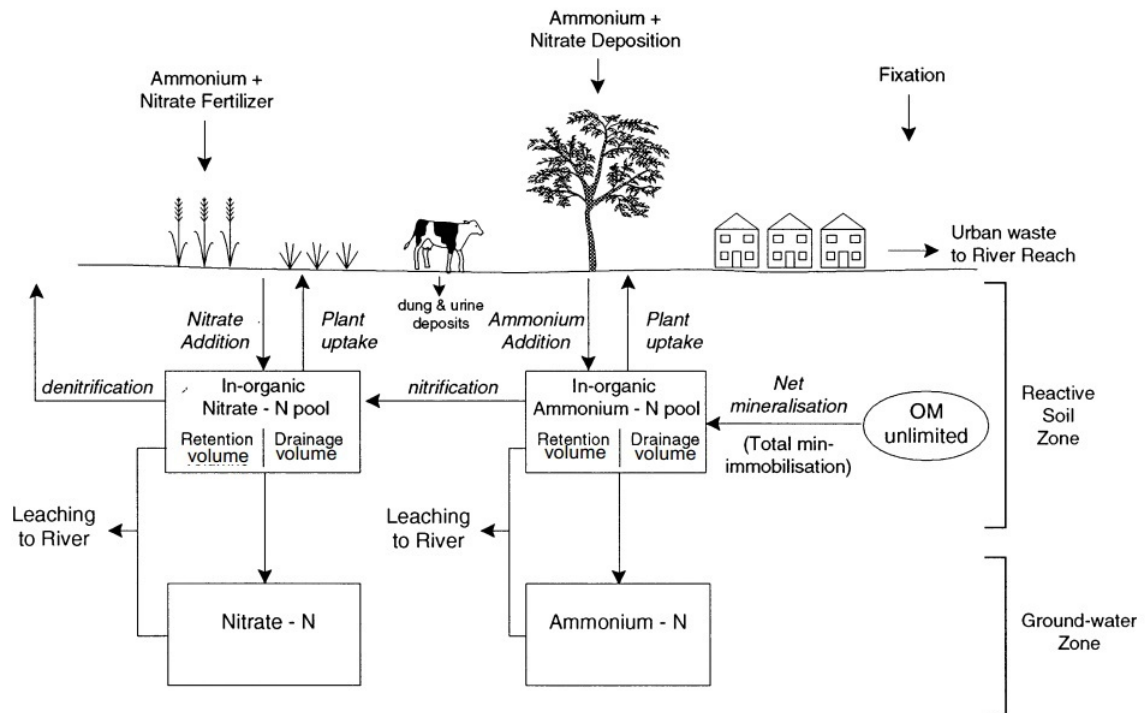
<sup>۱</sup>Integrated Nitrogen in Catchments

<sup>۲</sup>Process-based

<sup>۳</sup>Semi distributed

<sup>۴</sup>Whitehead

<sup>۵</sup>Wade



شکل ۹.۴: چرخه نیتروژن در طبیعت

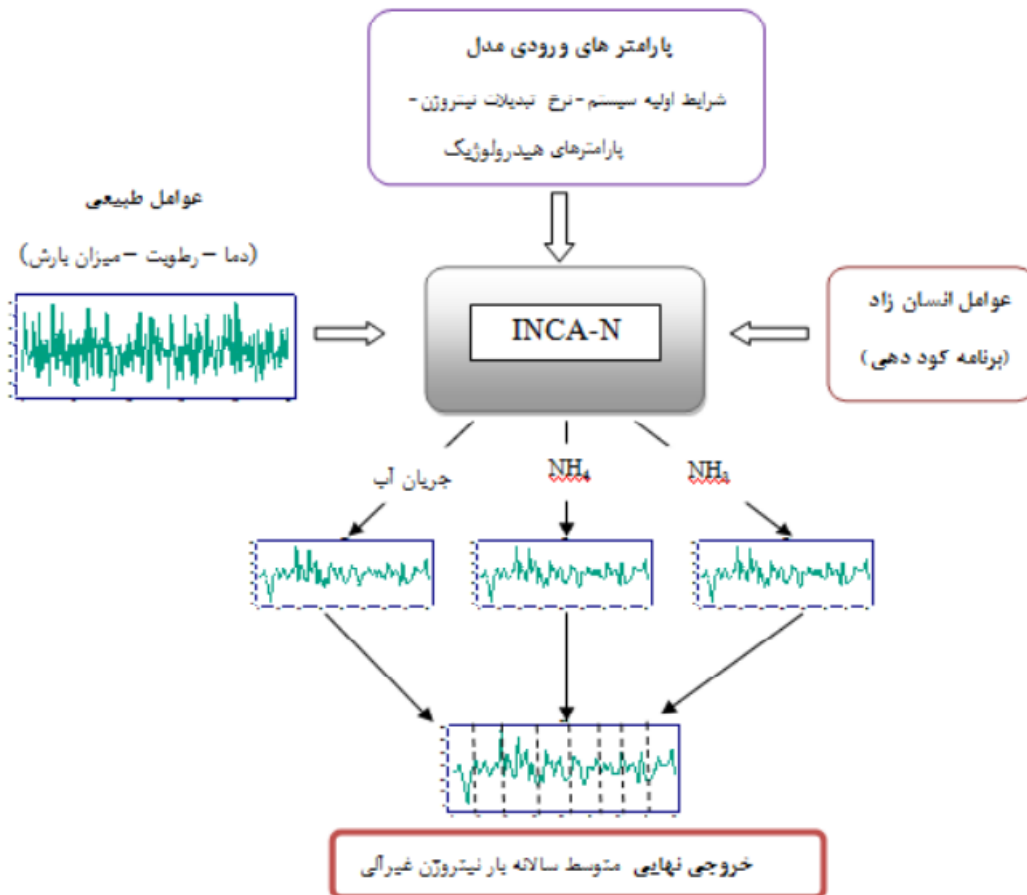
بومی مانند افزایش یا کاهش زمین مورد استفاده، تغییر در مقدار یا توزیع نیترات و آمونیاک بر میزان نیتروژن ورودی و غیره نیز به کار می‌رود. (جانفدا، ۱۳۹۰)

## ۵.۴ معرفی ورودی‌ها و خروجی‌های مدل

اطلاعات ورودی این مدل شامل عوامل طبیعی (دما، رطوبت و بارش روزانه)، عوامل دخالت انسانی (نظیر رژیم کوددهی زمین‌های کشاورزی) و پارامترهای ورودی مدل است. پارامترهای ورودی در سه دسته تقسیم‌بندی شده و مشخصات آن‌ها در جدول ۵.۴ خلاصه شده‌اند. در این جدول، علامت \* اشاره به زیرحوضه مورد مطالعه به وسعت یک کیلومتر مربع از زمین‌های کشاورزی دارد.

خروجی این مدل شامل سری‌های زمانی برآورد روزانه میزان جریان آب و غلظت‌های نیتروژن و آمونیاک در ایستگاه‌های مختلفی است که در طول رودخانه اصلی قرار دارند. این سه سری زمانی توسط رابطه زیر ترکیب شده و میزان بار نیتروژن غیرآلی روزانه را نتیجه می‌دهند:

(میزان غلظت آمونیاک + میزان غلظت نیترات)  $\times$  میزان جریان آب



شکل ۱۰.۴: ساختار شماتیکی مدل *INCA - N*

از آن جا که ترکیب این سه خروجی، خود نیز یک سری زمانی است، خروجی نهایی به صورت تک مقداری و تحت عنوان «متوسط سالانه بار ورودی نیتروژن غیرآلی به رودخانه» با معدل‌گیری از متوسط‌های سالانه محاسبه شده است.

## ۶.۴ منطقه مورد مطالعه

منطقه مورد مطالعه، حوزه آبریز توپید<sup>۱</sup> به مساحت ۴۳۹۰ کیلومتر مربع با رودخانه اصلی به طول ۱۵۶ کیلومتر واقع در کشور انگلستان است. موقعیت جغرافیایی آن از عرض شمالی ۵۵/۴۴۵° تا ۵۵/۷۷° و طول شرقی ۳/۴۹۶° تا ۱/۹۹۰۹° است. داده‌های آب و هواشناسی شامل دما، رطوبت، بارش روزانه و دبی روزانه ایستگاه‌های واقع در طول رودخانه اصلی به مدت هفت سال (۱۹۹۴-۲۰۰۰) اندازه‌گیری

<sup>۱</sup>Tweed



جدول ۵.۴: مشخصات عامل‌های ورودی مدل INCA-N

گروه	نام متغیر	واحد	دامنه تغییرات
شرایط اولیه	جریان سطحی: $x_1$	$m^3/s$	$[0, 0.01]^*$
	نیتрат سطح: $x_2$	$mgN/l$	$[0, 10]$
	آمونیاک سطح: $x_3$	$mgN/l$	$[0, 2]$
	حجم زهکشی سطحی: $x_4$	$m^3$	$[10^5, 2 \times 10^7]^*$
	جریان زیر سطحی: $x_5$	$m^3/s$	$[0, 0.01]^*$
	نیترات زیر سطحی: $x_6$	$mgN/l$	$[0, 10]$
	آمونیاک زیر سطحی: $x_7$	$mgN/l$	$[0, 2]$
	حجم زهکشی زیر سطحی: $x_8$	$m^3$	$[10^5, 2 \times 10^7]^*$
نرخ تبدیلات نیتروژن	نیترات زدایی: $x_9$	$mol/day$	$[0, 0.01]$
	ثابت نیتروژن: $x_{10}$	$kgN/ha/day$	$[0, 0.01]$
	جذب نیترات گیاه: $x_{11}$	$mol/day$	$[0, 0.05]$
	حداکثر جذب نیترات: $x_{12}$	$kgN/ha/year$	$[80, 140]$
	نرخ معدنی‌سازی: $x_{13}$	$kgN/ha/day$	$[0, 1]$
	آلی‌سازی: $x_{14}$	$mol/day$	$[0, 0.1]$
	جذب آمونیاک گیاه: $x_{15}$	$mol/day$	$[0, 0.05]$
عوامل هیدرولوژیکی	کمبود رطوبت خاک: $x_{16}$	$Mm$	$[100, 170]$
	زمان اقامت آب و خاک: $x_{17}$	$Day$	$[0.05, 5]$
	زمان اقامت آب‌های زیرزمینی: $x_{18}$	$Day$	$[10, 200]$
	حداکثر حجم احتباس خاک: $x_{19}$	$M$	$[0.01, 1]$

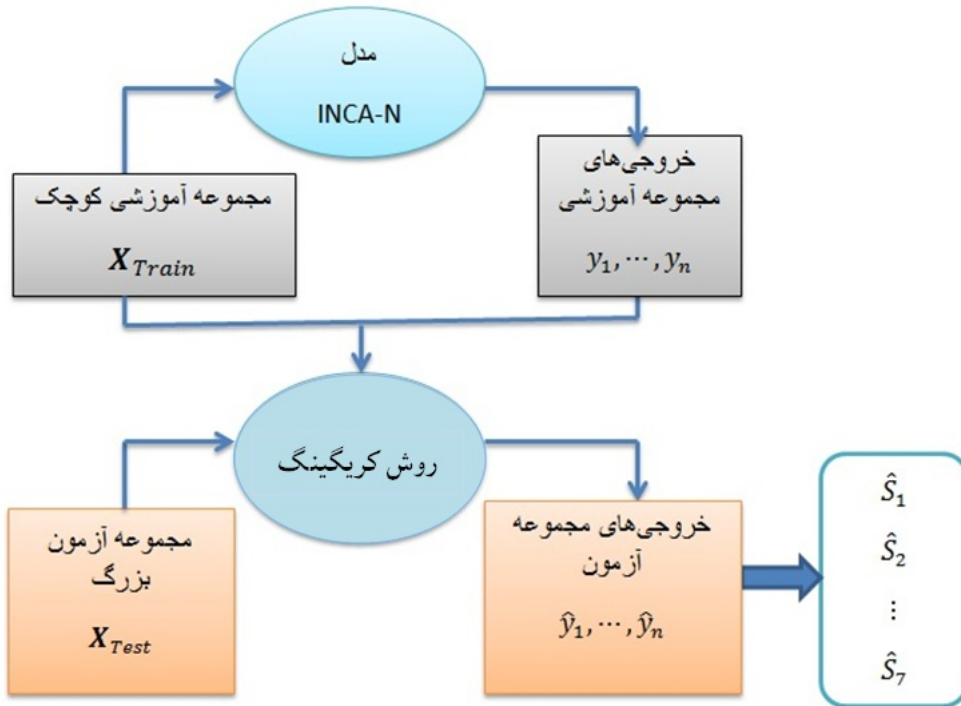
و ثبت شده‌اند (جاروی<sup>۱</sup> و همکاران، ۲۰۰۲). جهت سهولت در تحلیل نتایج، تنها یک زیرمجموعه از حوضه آبریز به وسعت یک کیلومتر مربع از زمین‌های کشاورزی مورد بررسی قرار گرفته است.

مدل INCA-N را در این مطالعه، سامانه‌ای از پارامترهای نرخ تبدیلات نیتروژن ( $x_9 - x_{15}$ ) به‌عنوان ورودی مدل در نظر گرفته‌ایم و سایر عوامل ورودی در مقادیر میانگین دامنه تغییرات آن‌ها ثابت نموده‌ایم. به عبارت دیگر مدل INCA-N را تابعی از هفت متغیر در نظر گرفته‌ایم.

## ۷.۴ تحلیل حساسیت مدل INCA-N با استفاده از مدل کریگینگ

شکل ۱۱.۴، کلیات نحوه برآورد شاخص‌های حساسیت مدل مذکور را نشان می‌دهد. برازش مدل

<sup>۱</sup>Jarvie



شکل ۱۱.۴: ساختار انجام برآورد شاخص‌های حساسیت اصلی برای مدل  $INCA - N$  به روش کریگینگ

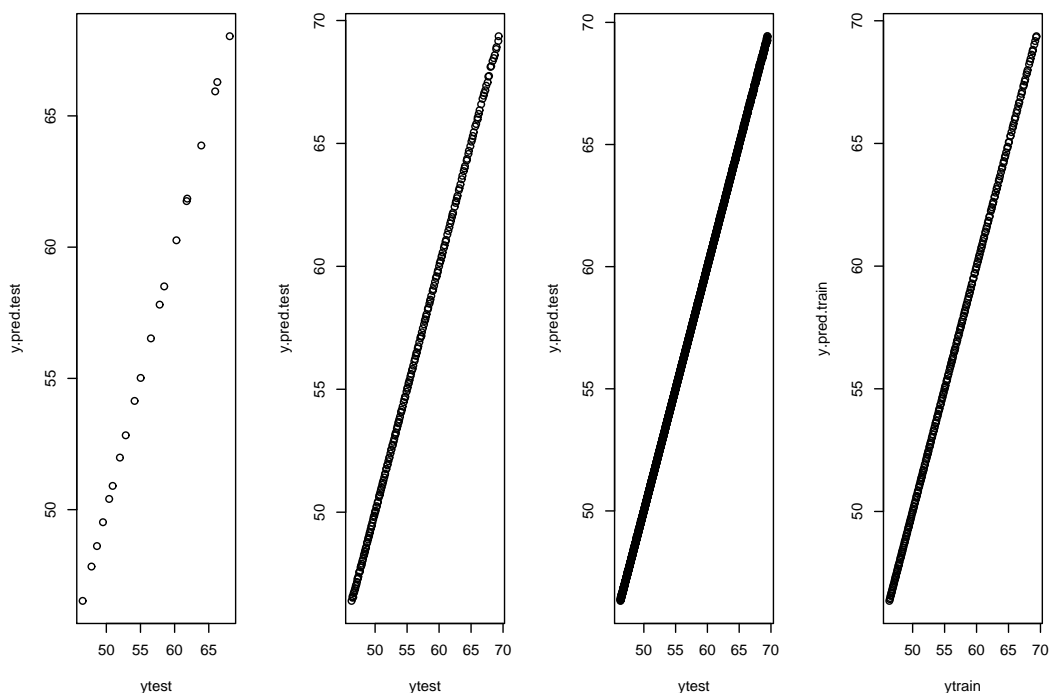
کریگینگ برای تابع واقعی  $INCA$ ، با  $۲۵^{\circ}$  داده مدل‌ساز بر اساس طرح آزمایش LHS، انجام شده است. برآوردهای درست‌نمایی ماکزیمم پارامترهای مدل، در جدول ۶.۴ آمده است.

جدول ۶.۴: برآورد پارامترهای مدل کریگینگ تابع  $INCA$

پارامتر	$\beta$	$\sigma^2$	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$	$\theta_6$	$\theta_7$
برآورد	۱۰/۱۲	۵۸/۴۸	۰/۰۰۵	۰/۰۱۹۹	۰/۰۹۹۷	۹۳/۰۵	۱/۹۹۷	۰/۱۹۹	۰/۰۹۹

نمودار پراکنش مقادیر پیش‌گویی شده در مقابل مقادیر واقعی برای داده‌های مدل‌ساز ( $n_{train} = ۲۵^{\circ}$ ) و برای داده‌های آزمون با حجم‌های  $n_{test} = ۲^{\circ}$ ،  $n_{test} = ۲۰^{\circ}$  و  $n_{test} = ۲۰۰۰^{\circ}$  در شکل ۱۲.۴ نشان داده شده‌اند. ضریب تعیین در حجم‌های متفاوت داده‌های آزمون، بدون تغییر است و بسیار نزدیک ۱ می‌باشد.

نمودارهای مربوط به اعتبارسنجی متقابل LOO برای تابع  $INCA$  در شکل ۱۳.۴ داده شده است. در همه‌ی این نمودارها مشاهده می‌کنیم که مدل کریگینگ برای تابع  $INCA$  بسیار عالی برازش یافته

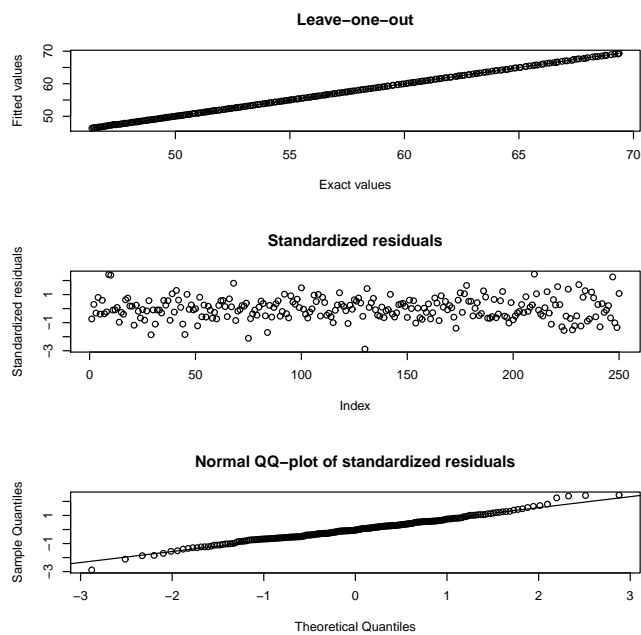


شکل ۱۲.۴: نمودار مقادیر پیش‌گویی شده در مقابل مقادیر واقعی برای  $250^\circ$  داده مدل‌ساز (اولین نمودار از راست) و داده‌های آزمون برای حجم‌های  $20^\circ$ ،  $200^\circ$  و  $2000^\circ$  (به ترتیب نمودارهای اول، دوم و سوم از چپ)

است.

با توجه به نمودارها و مقادیر  $R^2$  که در بالا به دست آمد، می‌توانیم از اعتبار تقریب کریگینگ در جانشینی تابع INCA مطمئن شویم و به محاسبه برآوردهای مونت کارلوی شاخص‌های حساسیت با کمک این مدل جانشین بپردازیم.

برای برآورد شاخص‌های حساسیت اصلی با استفاده از مدل جانشین، ابتدا تقریب کریگینگ تابع INCA را با  $250^\circ$  داده مدل‌ساز ساختیم و سپس با  $1000^\circ$  داده آزمون به برآورد این شاخص‌ها پرداختیم. در جدول ۷.۴ شاخص‌های حساسیت اصلی متغیرهای ورودی مدل INCA را که از مدل جانشین و مدل واقعی به دست آمده‌اند، مشاهده می‌کنیم. مقادیر اثر اصلی با مدل واقعی، بر اساس میانگین‌گیری از  $50^\circ$  تکرار بوت‌استرپ با نمونه‌هایی با حجم  $4000^\circ$  به دست آمده‌اند. همچنین برای سهولت در مقایسه، محتویات جدول ۷.۴ را در قالب نمودار میله‌ای ۱۴.۴ رسم کردیم.



شکل ۱۳.۴: اعتبارسنجی متقابل  $LOO$  برای مدل  $INCA$

جدول ۷.۴: برآورد شاخص‌های حساسیت اصلی متغیرهای ورودی مدل  $INCA$

متغیر	اثر اصلی با مدل جانشین	اثر اصلی با مدل واقعی
نرخ نیتراژ زدایی ( $X_1$ )	۰/۱۸۹۲۶	۰/۱۸۳۰۶
ثابت نیتروژن ( $X_2$ )	۰/۰۱۰۶۷	۰/۰۰۷۹۲
جذب نیتراژ ( $X_3$ )	۰/۴۴۵۳۵	۰/۴۳۲۱۹
حداکثر جذب نیتراژ ( $X_4$ )	۰/۰۱۰۰۶	۰/۰۲۴۷۱
نرخ معدنی سازی ( $X_5$ )	۰/۰۷۲۵۰	۰/۰۹۶۰۵
آلی سازی ( $X_6$ )	۰/۱۸۲۶۲	۰/۱۴۶۲۹
جذب آمونیاک گیاه ( $X_7$ )	۰/۰۰۳۹۷	۰/۰۰۲۱۱

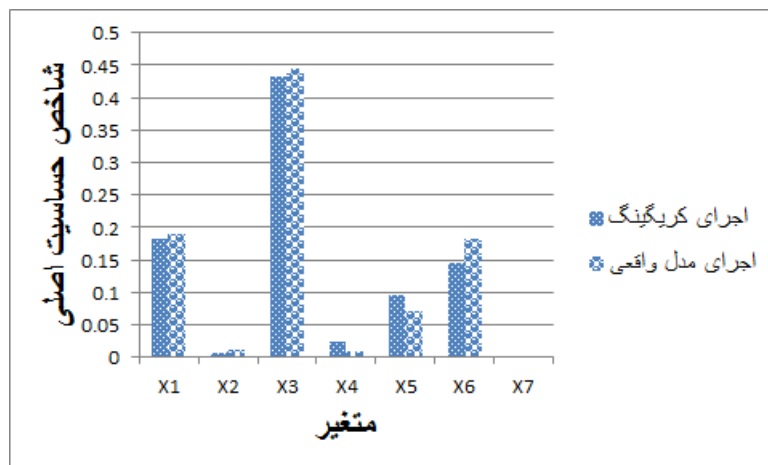
با توجه به مقادیر برآورد شده شاخص حساسیت مرتبه اول در جدول ۷.۴ و نمودار ۱۴.۴، مهمترین عوامل تاثیرگذار بر متوسط بار نیتروژن سالانه به ترتیب عبارتند از:

۱. میزان جذب نیتراژ توسط گیاه ( $X_3$ )

۲. نرخ نیتراژ زدایی ( $X_1$ )

۳. آلی سازی ( $X_6$ )

۴. نرخ معدنی سازی ( $X_5$ )



شکل ۱۴.۴: نمودار شاخص‌های حساسیت مرتبه اول مدل  $INCA - N$ ، توسط اجرای مدل واقعی و روش کریگینگ

میزان اثر سایر متغیرهای مدل بسیار ناچیز است و می‌توان از آن‌ها صرفه‌نظر نمود. در واقع، ثابت نیتروژن موجود در هوا ( $X_2$ )، نرخ جذب آمونیاک گیاه ( $X_4$ ) و حداکثر جذب نیترات ( $X_7$ )، ورودی‌های بی‌تأثیر در مدل INCA هستند.

همچنین مجموع اثرات اصلی تمام متغیرهای مدل برابر ۰/۹۱۴۴۳ است. این بدان مفهوم است که مدل INCA در این سامانه هفت‌متغیره، به‌طور کامل جمعی نبوده و اثرات متقابل بین این متغیرها وجود دارد، اما میزان آن‌ها چندان زیاد نمی‌باشد.

## ۸.۴ نتیجه‌گیری

بنابر نتایج تحلیل حساسیت مدل INCA با استفاده از تقریب کریگینگ به عنوان یک مدل جانشین برای مدل واقعی INCA، دریافتیم که این جانشینی در برآورد شاخص‌های حساسیت بسیار خوب عمل کرده است و توانسته است با کاهش هزینه محاسبات، نتایجی معادل با نتایج به‌دست آمده از اجرای مدل واقعی کسب کند. لذا می‌توان مهمترین عوامل تاثیرگذار بر نیتروژن ورودی به آب رودخانه تولید در بین متغیرهای نرخ تبادلات نیتروژن را به ترتیب میزان جذب نیترات توسط گیاهان ( $X_3$ )، نرخ نیترات‌زدایی ( $X_1$ )، آلی‌سازی ( $X_6$ ) و نرخ معدنی‌سازی ( $X_5$ ) دانست.

## ۹.۴ پیشنهادات

به منظور انجام تحقیقات و پژوهش‌های آتی در زمینه مدل‌های جانشین، می‌توان از راهکارهای پیشنهادی ذیل بهره‌گرفت:

۱. استفاده از ساختارهای همسایگی، یعنی به‌جای پیش‌گویی نقطه‌آزمون با تمام داده‌های مدل‌ساز، از آن تعداد داده‌های مدل‌سازی برای پیش‌گویی استفاده کنیم که در نزدیکی و همسایگی نقطه‌آزمون قرار دارند.

۲. بررسی تاثیر توابع کوواریانس مختلف، چراکه تابع کوواریانس جزء مهمی از کریگینگ را تشکیل می‌دهد و تاثیرگذاری قابل توجهی بر تقریب حاصل از کریگینگ دارد. لذا با بررسی توابع کوواریانس مختلف می‌توانیم بهترین تابع کوواریانس را یافته و از آن در جهت بهبود مدل و بالا بردن دقت پیش‌گویی استفاده نمود.

۳. بررسی اثر قطعه‌ای روی پیش‌گو، هنگامی که دو نقطه با فاصله بسیار کم از یکدیگر دارای مقادیر بسیار متفاوت باشند، اثر قطعه‌ای را به‌وجود می‌آورند. وجود این اثر قطعه‌ای، فرض همواری مدل را دچار ایراد می‌کند، لذا در هنگام وجود اثر قطعه‌ای باید راهکار دیگری برای پیش‌گویی مدل اتخاذ شود.

۴. بررسی تاثیر روند، به عبارت دیگر می‌توان از انواع مدل‌ها که دارای روندهای متفاوتی هستند، استفاده نمود.

۵. بررسی دیگر مدل‌های آماری مانند MARS و اسپلاین در جانشینی مدل‌های کامپیوتری نیز در خور توجه می‌باشد.



# پیوست آ

## نمادها و تعریفها

### ۱. آ فرآیندهای تصادفی گاوسی

یک فرآیند تصادفی مجموعه‌ای از متغیرهای تصادفی  $\{X(t) : t \in T\}$  است که در آن  $X(t)$  به ازای هر  $t \in T$  ( $T$  مجموعه اندیس فرآیند) یک متغیر تصادفی است. فرآیند تصادفی  $X(t)$  را فرآیند گاوسی گوئیم اگر برای هر  $n$ ، و هر انتخاب  $t_1, t_2, \dots, t_n$ ، توزیع توام متغیرهای تصادفی  $X(t_1), \dots, X(t_n)$  تواماً گاوسی باشد. یعنی هر یک از متغیرهای تصادفی به تنهایی و همچنین به طور مشترک دارای چگالی احتمال نرمال باشند. خیلی از پدیده‌هایی که از عوامل طبیعی ناشی می‌شوند می‌توانند به خوبی توسط فرآیندهای گاوسی و با استفاده از قضیه حد مرکزی تقریب شوند. علاوه بر این، نتیجه تبدیل یک فرآیند تصادفی گاوسی تحت نگاشت خطی نیز فرآیند تصادفی گاوسی خواهد بود. این خاصیت در کاربردهای نظری و عملی نظریه فرآیندهای تصادفی بسیار حایز اهمیت است. همچنین دیگر ویژگی خوب فرآیند گاوسی این است که می‌توان آن را به طور کامل توسط میانگین و تابع کوواریانس‌اش توصیف نمود.

### ۲. آ درون‌یابی

در محاسبات عددی، درون‌یابی روشی است برای یافتن مقدار تابع درون یک بازه، هنگامی که مقدار تابع در تعدادی از نقاط گسسته معلوم است. یافتن مقدار تابع در خارج از این بازه را برون‌یابی<sup>۱</sup> گویند

<sup>۱</sup>Extrapolation



که عموماً از روش‌های مشابهی برای هر دو استفاده می‌شود.

در بسیاری از کاربردها در مهندسی و علوم پایه تعدادی نقاط معلوم در دسترس است، مانند داده‌های به دست آمده از آزمایش و یا نمونه‌برداری. در چنین مواردی سعی می‌شود تابعی یافت که حتی‌المقدور به داده‌ها نزدیک‌تر باشد. یکی از روش‌های یافتن چنین تابعی درون‌یابی می‌باشد که وجه مشخصه این روش آن است که تابع یافت شده از این روش از تمامی نقاط داده شده می‌گذرد.

### ۳.آ ناهم‌واریانسی

در آمار دنباله‌ای، متغیرهای تصادفی که دارای واریانس‌های متفاوتی باشند ناهم‌واریانس نامیده می‌شود. در مقابل به یک دنباله از متغیرهای تصادفی واریانس همسان می‌گویند اگر دارای واریانس ثابتی باشند. فرض کنید یک دنباله از متغیرهای تصادفی در اختیار داریم:  $\{Y_t\}_{t=1}^n$  و یک دنباله از بردارهای متغیرهای تصادفی:  $\{X_t\}_{t=1}^n$ . با استفاده از امید ریاضی  $Y_t$  به شرط  $X_t$ ، دنباله‌ی  $\{Y_t\}_{t=1}^n$  واریانس ناهمسان نامیده می‌شود اگر واریانس  $Y_t$  به شرط  $X_t$ ، با تغییر  $t$  تغییر نماید.

برخی از منابع از این تعریف تحت عنوان ”واریانس ناهمسانی مشروط” یاد می‌کنند، که این موضوع جهت تأکید بر این واقعیت است که ”دنباله‌ای از واریانس‌های شرطی” می‌باشند که متفاوت بوده و تغییر می‌یابند و نه واریانس‌های غیر شرطی. در واقع ممکن است واریانس ناهمسانی مشروط داشته باشیم در حالی که متغیرهای تصادفی به شکل غیرشرطی واریانس همسان‌اند.

### ۴.آ توزیع گامای معکوس

در نظریه احتمالات و آمار، توزیع گامای وارونه توزیعی پیوسته و دو پارامتری است که تنها روی متغیرهای با مقدار مثبت تعریف می‌شود. بسیاری از کاربردهای توزیع گاما در مدل‌بندی بیزی است. می‌گوییم متغیر تصادفی  $X$  دارای توزیع گامای معکوس با پارامترهای  $\alpha$  و  $\beta$  است، اگر دارای تابع چگالی به صورت زیر باشد

$$f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp\left(-\frac{\beta}{x}\right), \quad x \in (0, \infty), \alpha > 0, \beta > 0.$$

## ۵.آ. توزیع نرمال-گاما معکوس

در نظریه احتمالات و آمار، توزیع گامای نرمال وارونه یک خانواده چهار پارامتری از توزیع‌های احتمالی پیوسته چند متغیره است. این توزیع یک پیشین توأم از توزیع نرمال با میانگین و واریانس نامعلوم است. فرض کنید

$$x \mid \sigma^2, \mu, \lambda \sim N(\mu, \sigma^2/\lambda),$$

یک توزیع نرمال با میانگین  $\mu$  و واریانس  $\sigma^2/\lambda$  است، که

$$\sigma^2 \mid \alpha, \beta \sim \Gamma^{-1}(\alpha, \beta)$$

یک توزیع گاما وارونه است. آنگاه  $(x, \sigma^2)$  دارای توزیع گاما نرمال وارونه است که به صورت زیر نشان داده شده می‌شود

$$(x, \sigma^2) \sim N - \Gamma^{-1}(\mu, \lambda, \alpha, \beta)$$

بجای نماد  $N - \Gamma^{-1}$  از  $NIG$  نیز استفاده می‌شود. تابع چگالی  $(x, \sigma^2)$  به صورت زیر می‌باشد

$$f(x, \sigma^2 \mid \mu, V^{-1}, \alpha, \beta) = \frac{\sqrt{\lambda}}{\sigma \sqrt{2\pi}} \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{\sigma^2}\right)^{\alpha+1} \exp\left(-\frac{2\beta + \lambda(x - \mu)^2}{2\sigma^2}\right)$$

در این توزیع  $\lambda, \alpha, \beta > 0$  و  $\mu \in R, \sigma^2 \in [0, \infty), x \in R$  در حالت چند متغیره،

$$\mathbf{x} \mid \sigma^2, \mu, V^{-1} \sim N(\mu, \sigma^2 V)$$

دارای توزیع گاما نرمال وارونه است، که در آن  $\mathbf{x}$  یک بردار تصادفی  $1 \times k$  با میانگین  $\mu$  و کوواریانس

$\sigma^2 V$  است. تابع چگالی گاما نرمال وارونه در حالت چند متغیره به صورت زیر است:

$$f(\mathbf{x}, \sigma^2 \mid \mu, V^{-1}, \alpha, \beta) = |V|^{-1/2} (2\pi)^{-k/2} \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{\sigma^2}\right)^{\frac{k}{2} + \alpha + 1} \exp\left(-\frac{2\beta + (\mathbf{x} - \mu)^T V^{-1} (\mathbf{x} - \mu)}{2\sigma^2}\right)$$

که  $|V|$  دترمینان ماتریس  $V_{k \times k}$  است. در این حالت اگر  $k = 1$ ، به طوری که  $\mathbf{x}$  و  $V$  اسکالر باشند، حالت تک متغیره به دست می‌آید.

## ۶.آ تابع چگالی نرمال چند متغیره

می‌گوییم بردار تصادفی  $\mathbf{y} \in \mathbb{R}^p$  دارای توزیع نرمال چندمتغیره با بردار میانگین  $\boldsymbol{\mu} \in \mathbb{R}^p$  و ماتریس واریانس-کوواریانس  $\boldsymbol{\Sigma} > 0$  است، اگر تابع چگالی آن به صورت زیر باشد

$$f(\mathbf{y}) = \frac{1}{(\sqrt{2\pi})^p |\boldsymbol{\Sigma}|^{1/2}} e^{-(\mathbf{y}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y}-\boldsymbol{\mu})/2} \quad (1.آ)$$

چگالی نرمال چند متغیره با میانگین  $\boldsymbol{\mu}$  و ماتریس کوواریانس  $\boldsymbol{\Sigma}$  است. وقتی  $\mathbf{y}$  دارای چگالی (۱.آ) است، می‌گوییم  $\mathbf{y}$  توزیعی به صورت  $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  دارد یا به اختصار  $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  است. اندیس  $p$  بعد نرمال متغیره  $p$  است و تعداد متغیرها را نشان می‌دهد، یعنی،  $\mathbf{y}$  بردار  $1 \times p$  و  $\boldsymbol{\mu}$  بردار  $1 \times p$  و  $\boldsymbol{\Sigma}$  ماتریس  $p \times p$  است.

**تعریف ۱.۶.آ.** رتبه هر ماتریس مربع یا مستطیل  $\mathbf{A}$  به صورت زیر تعریف می‌شود:

$$\text{رتبه}(\mathbf{A}) = \text{تعداد ستون‌های مستقل } \mathbf{A}$$

$$= \text{تعداد سطرهای مستقل } \mathbf{A}$$

**تعریف ۲.۶.آ.** اگر ماتریس متقارن  $\mathbf{A}$  دارای این خاصیت باشد که به ازای تمام مقادیر ممکن بردار  $\mathbf{y}$ ، به غیر از  $\mathbf{y} = 0$ ،  $\mathbf{y}^T \mathbf{A} \mathbf{y} > 0$  باشد، آنگاه  $\mathbf{A}$  را ماتریس معین مثبت گوئیم و همین‌طور اگر برای تمام  $\mathbf{y}$ ها غیر از  $\mathbf{y} = 0$ ،  $\mathbf{y}^T \mathbf{A} \mathbf{y} \geq 0$  باشد، آنگاه  $\mathbf{A}$  را نیمه معین مثبت گوئیم.

**قضیه ۳.۶.آ.** (گوس-مارکوف) اگر  $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$  و  $Cov(\mathbf{y}) = \sigma^2 \mathbf{I}$  آنگاه برآوردگرهای کمترین مربعات  $\hat{\beta}_j$ ،  $j = 0, 1, \dots, k$  در میان تمام برآوردگرهای ناریب خطی دارای کمترین واریانس است.

قضیه گوس-مارکوف گاهی به صورت زیر بیان می‌شود: اگر  $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$  و  $Cov(\mathbf{y}) = \sigma^2 \mathbf{I}$  باشد، آنگاه برآوردگرهای کمترین مربعات  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$  بهترین برآوردگرهای ناریب خطی (BLUE) هستند. در این عبارت، بهترین یعنی دارای کمترین واریانس و خطی نشان می‌دهد که برآوردگرها تابعی خطی از  $\mathbf{y}$  هستند.

**نتیجه ۴.۶.آ.** اگر  $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$  و  $Cov(\mathbf{y}) = \sigma^2 \mathbf{I}$ ، بهترین برآوردگر خطی ناریب  $\mathbf{a}^T \boldsymbol{\beta}$  عبارتست

$$\text{از: } \hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \text{ که در آن } \hat{\beta} \text{ برآوردگر کمترین مربعات } \hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \text{ است.}$$

## ۷.آ معکوس ماتریس بلوکی

دو ماتریس بلوکی  $A$  و  $B$  با اندازه  $n \times n$  را به صورت زیر در نظر بگیرید:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

که  $A_{11}$  و  $B_{11}$  ماتریس‌هایی با اندازه  $k \times k$  هستند. فرض کنید  $A$  غیرمنفرد (تکین) و  $B = A^{-1}$ .

می‌خواهیم  $B_{ij}$  را بر حسب  $A_{ij}$ ‌ها به گونه‌ای به دست آوریم که

$$\det(A_{11}) \neq 0, \quad \det(A_{22}) \neq 0 \quad (۲.آ)$$

شرایط (۲.آ) برای تکین بودن  $A$  کافی است. اگرچه این شرایط در حالت کلی ضروری نیستند برای مثال

هنگامی که

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

از طرف دیگر اگر  $A$  معین مثبت باشد آنگاه این شرایط ضروری هستند. اگر  $B = A^{-1}$  آنگاه

$$\begin{aligned} AB &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \\ &= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix} \\ &= \begin{bmatrix} I_k & \mathbf{O}_{k,n-k} \\ \mathbf{O}_{n-k,k} & I_{n-k} \end{bmatrix} \end{aligned} \quad (۳.آ)$$

که  $I$  ماتریس همانی و  $\mathbf{O}$  ماتریس صفر می‌باشد که اندازه‌های آن‌ها توسط زیروندهایشان نشان داده

شده‌اند. برای حل (۳.آ)، نیاز به حل چهار معادله ماتریسی زیر داریم:

$$A_{11}B_{11} + A_{12}B_{21} = I_k, \quad (۴.آ)$$

$$A_{11}B_{12} + A_{12}B_{22} = \mathbf{O}_{k,n-k}, \quad (۵.آ)$$

$$A_{21}B_{11} + A_{22}B_{21} = \mathbf{O}_{n-k,k}, \quad (۶.آ)$$

$$A_{21}B_{12} + A_{22}B_{22} = I_{n-k}. \quad (۷.آ)$$

نتایج زیر از (۵.آ) و (۶.آ) به دست می‌آیند،

$$B_{12} = -A_{11}^{-1}A_{12}B_{22} \quad (۸.آ)$$

$$B_{21} = -A_{22}^{-1}A_{21}B_{11}. \quad (9.آ)$$

(۴.آ) و (۷.آ) را می‌توانیم به صورت زیر بازنویسی کنیم،

$$(A_{11} - A_{12}A_{22}^{-1}A_{21})B_{11} = I_k,$$

$$(A_{22} - A_{21}A_{11}^{-1}A_{12})B_{22} = I_{n-k},$$

پس داریم:

$$B_{11} = (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1},$$

$$B_{22} = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}.$$

با جایگذاری این جوابها در (۸.آ) و (۹.آ)، نتایج زیر را خواهیم داشت:

$$B_{12} = -A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1},$$

$$B_{21} = -A_{22}^{-1}A_{21}(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}.$$

پس  $A^{-1}$  به صورت زیر حاصل می‌شود:

$$A^{-1} = \begin{bmatrix} (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} & -A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \\ -A_{22}^{-1}A_{21}(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix}$$

## برآوردگر $\sigma^2$

با فرض این که  $\sigma^2$  برای هر  $y_i, i = 1, 2, \dots, n$  ثابت است، می‌خواهیم یک برآوردگر نااریب برای  $\sigma^2$  بر مبنای برآوردگر کمترین مربعات  $\hat{\beta}$  بسازیم.  $\sigma^2$  به صورت  $\sigma^2 = E[y_i - E(y_i)]^2$  تعریف می‌شود و فرض کردیم که داریم:

$$E(y_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} = \mathbf{x}_i^T \boldsymbol{\beta},$$

که در آن سطر  $i$ ام ماتریس  $\mathbf{X}$  است. پس  $\sigma^2$  به صورت زیر نوشته می‌شود:

$$\sigma^2 = E[y_i - \mathbf{x}_i^T \boldsymbol{\beta}]^2$$

مقدار  $\sigma^2$  را به صورت یک مقدار متوسط از نمونه برآورد می‌کنیم،

$$s^2 = \frac{1}{n - k - 1} \sum_{i=1}^n (y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2 \quad (10.آ)$$

که در آن  $n$  حجم نمونه و  $d$  تعداد  $x$  ها است. بنا به نتیجه (۴.۶.آ)، برای  $\mathbf{x}_i^T \hat{\beta}$  یک برآوردگر BLUE است.

رابطه‌ی (۱۰.آ) را می‌توان به صورت زیر نوشت:

$$s^2 = \frac{1}{n-k-1} (\mathbf{y} - \mathbf{X}\hat{\beta})^T (\mathbf{y} - \mathbf{X}\hat{\beta}) = \frac{\mathbf{y}^T \mathbf{y} - \hat{\beta}^T \mathbf{X}^T \mathbf{y}}{n-k-1} \quad (11.آ)$$

صورت رابطه‌ی (۱۱.آ) که بیانگر مجموع مربعات خطاست را می‌توان با SSE نیز نشان داد.  $s^2$  یک برآوردگر ناریب برای  $\sigma^2$  است.

**قضیه ۱.۷.آ.** اگر  $\mathbf{y}$  دارای توزیع  $N_n(\mathbf{X}\beta, \sigma^2 \mathbf{I})$  باشد که در آن  $\mathbf{X}$  ماتریس  $n \times (k+1)$  با رتبه  $k+1 < n$  است، آنگاه برآوردگرهای درست‌نمایی ماکزیم  $\beta$  و  $\sigma^2$  عبارتند از:

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}, \quad (12.آ)$$

$$\hat{\sigma}^2 = \frac{1}{n} (\mathbf{y} - \mathbf{X}\hat{\beta})^T (\mathbf{y} - \mathbf{X}\hat{\beta}). \quad (13.آ)$$

**قضیه ۲.۷.آ.** فرض کنید  $\mathbf{y}$  دارای توزیع  $N_n(\mathbf{X}\beta, \sigma^2 \mathbf{I})$  باشد، که در آن  $\mathbf{X}$  ماتریس  $n \times (k+1)$  با رتبه  $k+1 < n$  است و  $\beta = (\beta_0, \beta_1, \dots, \beta_k)^T$ . در این صورت برآوردگرهای درست‌نمایی ماکزیم  $\hat{\beta}$  و  $\hat{\sigma}^2$  در قضیه (۱.۷.آ) دارای خواص زیرند:

(۱) توزیع  $\hat{\beta}$  عبارتست از:  $N_{k+1}[\beta, \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}]$ .

(۲) توزیع  $n\hat{\sigma}^2/\sigma^2$  عبارتست از:  $\chi^2(n-k-1)$ ، یا به عبارت معادل  $s^2/\sigma^2$  دارای توزیع  $\chi^2(n-k-1)$  است.

(۳) برآوردگرهای  $\hat{\beta}$  و  $\hat{\sigma}^2$  (یا  $s^2$ ) مستقلند.

## ۸.آ تغییرنگار

محاسبه یک تابع کوواریانس نیاز به داشتن اطلاعات  $E(Y(\mathbf{x}))$  و  $E(Y(\mathbf{u}))$  دارد. چون این اطلاعات در دسترس نیستند، زمین‌شناسان اندازه‌ی دیگری از تغییرات مجزا ایجاد کرده‌اند که می‌تواند بدون میانگین‌ها

محاسبه شود. این اندازه را تغییرنگار<sup>۱</sup> می‌نامند. تغییرنگار، واریانس اختلاف بین  $Y(\mathbf{x})$  و  $Y(\mathbf{u})$  است. فرم ریاضی آن به صورت

$$\gamma(r) = \text{Var}(Y(\mathbf{x}) - Y(\mathbf{u}))$$

است. که  $r$  فاصله اقلیدسی بین دو نقطه  $\mathbf{x}$  و  $\mathbf{u}$  است.  $\gamma(r)$  نیم‌تغییرنگار نامیده می‌شود. رابطه بین کوواریانس و نیم‌تغییرنگار به صورت زیر است.

$$\text{Cov}(Y(\mathbf{x}), Y(\mathbf{u})) = \text{Var}(Y(\mathbf{x})) - \gamma(r)$$

برای پیش‌گویی مدل از مجموعه طرح  $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_n]$  به عنوان نقاط آزمایش استفاده می‌شود، و بردار پاسخها در نقاط طرح به صورت  $\mathbf{y}_S = [y(\mathbf{s}_1), \dots, y(\mathbf{s}_n)]^T$  تعریف می‌شود.

## ۹.۱ اثر قطعه‌ای

مقدار تغییرنگار در مبدا مختصات ( $r = 0$ )، اثر قطعه‌ای<sup>۲</sup> نامیده شده و به صورت

$$c_0 = \lim_{r \rightarrow 0} \gamma(r)$$

تعریف می‌شود. به لحاظ نظری اثر قطعه‌ای باید صفر باشد. زیرا دو نمونه که از موقعیت واحدی گرفته شده‌اند، باید کمیت یکسانی داشته باشند. ولی در عمل تغییرنگارهای تجربی از این وضعیت پیروی نمی‌کنند. دلایل این موضوع را می‌توان در خطای نمونه‌گیری، خطای اندازه‌گیری یا تغییرات شدید کمیت مورد بررسی در موقعیت‌های نزدیک به هم دانست.

<sup>۱</sup>Variogram

<sup>۲</sup>Nugget effect

# پیوست ب

## کد برنامه $R$ برای اجرای روش کریگینگ و برآوردهای مونت کارلویی در فصل چهارم

۱- بررسی مناسبت مدل و محاسبه برآوردهای شاخص‌های حساسیت اصلی و کل برای تابع سبل:

```
library(lhs)

library(DiceKriging)

library(sensitivity)

n <- 250

d <- 8

set.seed(0)

X <- maximinLHS(n, 2*d)

A <- X[, -(d+1):(2*d)]

B <- X[, -1:-d]

X1 <- data.frame(A)

X2 <- data.frame(B)

p <- ncol(X1)

X <- rbind(X1, X2)

for (i in 1:p) {
```



```
Xb <- X2

Xb[, i] <- X1[, i]

X <- rbind(X, Xb)

}

# plot(X)

y <- sobol.fun(X)

m <- km( design = X, response = y, covtype = "gauss")

plot(m)

n.test<-1000

set.seed(0)

X.test1 <- randomLHS(n.test, d)

colnames(X.test1) <- names(X)

y.test.1 <- sobol.fun(X.test1)

y.pred.1 <- predict(m, newdata = X.test1, type = "UK",
covtype = "gauss")$mean

X.test2 <- randomLHS(n.test, d)

colnames(X.test2) <- names(X)

y.test.2 <- sobol.fun(X.test2)

y.pred.2 <- predict(m, newdata = X.test2, type = "UK",
covtype = "gauss")$mean

y.test.C <-matrix(0,n.test,8)

y.pred.C <- matrix(0,n.test,8)

for (i in 1:8){

X.test.C <- X.test2
```

```
X.test.C[,i] <- X.test1[,i]

colnames(X.test.C) <- names(X)

y.test.C[,i] <- sobol.fun(X.test.C)

y.pred.C[,i] <- predict(m, newdata = X.test.C, type = "UK",
covtype = "gauss")$mean
}

###To make predicted values

yp <- cbind(y.pred.1,y.pred.2, y.pred.C[,1],y.pred.C[,2],y.pred.C[,3],
y.pred.C[,4],y.pred.C[,5],y.pred.C[,6],y.pred.C[,7],y.pred.C[,8])

yp <- data.frame(yp)

#### to use real values

yt <- cbind(y.test.1,y.test.2, y.test.C[,1],y.test.C[,2],y.test.C[,3],
y.test.C[,4],y.test.C[,5],y.test.C[,6],y.test.C[,7],y.test.C[,8])

yt <- data.frame(yt)

#### To compute Ujs

u=matrix(0,1,nrow=d)

for (i in 1:d){

sum = 0

for(j in 1:n.test){

sum = sum + yp[j,1]* yp[j,i+2]

}

u[i]=sum/n.test

}

###To compute mean for predicted values
```

```

sum = 0

for(i in 1:n.test){

  sum = sum +yp[i,1]

}

Ey=sum/n.test

####To compute variance for predicted values

sum=0

for(i in 1:n.test){

  sum=sum +yp[i,1]^2

}

Ey2= sum/n.test

###To compute first order SIs with predicted values

S.test.pred = matrix(0,1,nrow=d)

for (i in 1:d){

  S.test.pred[i]=(u[i]-(Ey^2))/(Ey2-(Ey^2))

}

#####

u=matrix(0,1,nrow=d)

for (i in 1:d){

  sum = 0

  for(j in 1:n.test){

    sum = sum + yt[j,1]* yt[j,i+2]

  }

  u[i]=sum/n.test

```

```

    }

###To compute mean for predicted values

sum = 0

for(i in 1:n.test){

    sum = sum +yt[i,1]

}

Ey=sum/n.test

####To compute variance for predicted values

sum=0

for(i in 1:n.test){

    sum=sum +yt[i,1]^2

}

Ey2= sum/n.test

###To compute first order SIs with real values

S.test = matrix(0,1,nrow=d)

for (i in 1:d){

    S.test[i]=(u[i]-(Ey^2))/(Ey2-(Ey^2))

}

##the real values of sensitivity indices in g.function

d <- 8

a <- c(0,1,4.5,9,99,99,99,99)

V <- c()

for (i in 1:d){

V[i] <- 1/(3*((1+a[i])^2))

```

```

}

Var<- (prod(V+1)-1)

realS <- V/Var

Ubj<- c()

for(i in 1:d){

Ubj[i] <- (sum(V)-V[i])+(prod(V)/V[i])}

realSt<- 1-(Ubj/Var)

#####

firstSI= matrix(0, ncol=3, nrow=8)

firstSI[,1] <- S.test.pred

firstSI[,2] <- S.test

firstSI[,3] <- realS

colnames(firstSI)<- c("S.test.pred","S.test","realS")

#####

y.test.D <- matrix(0,n.test,8)

y.pred.D <- matrix(0,n.test,8)

for (i in 1:8){

X.test.D <- X.test2

X.test.D[,i] <- X.test1[,i]

colnames(X.test.D) <- names(X)

y.test.D[,i] <- sobol.fun(X.test.D)

y.pred.D[,i] <- predict(m, newdata = X.test.D, type = "UK",

covtype = "gauss")$mean

}

```

Y\

---

```
####To make matrix for all Djs

yp <- cbind(y.pred.1,y.pred.2, y.pred.D[,1],y.pred.D[,2],y.pred.D[,3],
y.pred.D[,4],y.pred.D[,5],y.pred.D[,6],y.pred.D[,7],y.pred.D[,8])

yp <- data.frame(yp)

#### to use real values: note that to should replace yt with yp

yt <- cbind(y.test.1,y.test.2, y.test.D[,1],y.test.D[,2],y.test.D[,3],
y.test.D[,4],y.test.D[,5],y.test.D[,6],y.test.D[,7],y.test.D[,8])

yt <- data.frame(yt)

#### To compute Ujs

u=matrix(0,1,nrow=d)

for (i in 1:d){

  sum = 0

  for(j in 1:n.test){

    sum = sum + yp[j,1]* yp[j,i+2]

  }

u[i]=sum/n.test

}

####To compute mean for predicted values

sum = 0

for(i in 1:n.test){

  sum = sum +yp[i,1]

}

Ey=sum/n.test

####To compute variance for predicted values
```

```

sum=0

for(i in 1:n.test){

  sum=sum +yp[i,1]^2

  }

Ey2= sum/n.test

###To compute first order SIs

ST.test.pred = matrix(0,1,nrow=d)

for (i in 1:d){

  ST.test.pred[i]=1-((u[i]-(Ey^2))/(Ey2-(Ey^2)))

  }

#####

u=matrix(0,1,nrow=d)

for (i in 1:d){

  sum = 0

  for(j in 1:n.test){

    sum = sum + yt[j,1]* yt[j,i+2]

  }

  u[i]=sum/n.test

  }

###To compute mean for predicted values

sum = 0

for(i in 1:n.test){

  sum = sum +yt[i,1]

  }

```

```
Ey=sum/n.test

####To compute variance for predicted values

sum=0

for(i in 1:n.test){

  sum=sum +yt[i,1]^2

}

Ey2= sum/n.test

###To compute first order SIs

ST.test = matrix(0,1,nrow=d)

for (i in 1:d){

  ST.test[i]=1-(((u[i]-(Ey^2))/(Ey2-(Ey^2))))

}

#####

totalSI= matrix(0, ncol=3, nrow=8)

totalSI[,1] <- ST.test.pred

totalSI[,2] <- ST.test

totalSI[,3] <- realSt

colnames(totalSI)<- c("ST.test.pred","ST.test","realSt")

۲- بررسی مناسبت مدل و محاسبه شاخص حساسیت اصلی برای مدل INCA

n<- 250

d <- 7

T.inca<- maximinLHS(n,d)

write(T.inca, file="G:/resultofinca/T.inca250.txt",

ncolumns=n, sep=" ")
```



```
y.i <- read.table("G:/resultofinca/Y.T.inca250.csv",
header = FALSE,sep=",")
ytrain <- y.i[,1]
design <- read.table("G:/resultofinca/Design-250-inca.csv",
header = FALSE,sep=",")
design <- data.frame(design)
m <- km( design = design, response = ytrain,covtype = "gauss")
plot(m)
n.test <- 1000
##X.test <- data.frame(matrix(as.character(X.test),nrow=n.test,ncol=7))
set.seed(0)
X.test <- maximinLHS(n.test, d)
X.test[,1]<-X.test[,1]*0.01
X.test[,2]<-X.test[,2]*0.01
X.test[,3]<-X.test[,3]*0.05
X.test[,4]<-60*X.test[,4]+80
X.test[,6]<-X.test[,6]*0.1
X.test[,7]<-X.test[,7]*0.05
colnames(X.test) <- names(design)
write(X.test, file="G:/resultofinca/X.test20.txt",
ncolumns=n.test, sep=" ")
ytest <- read.table("G:/resultofinca/Y.test20.csv",
header = FALSE,sep=",")
ytest <- ytest[,1]
```

```
par(mfrow=c(1,2))

  y.pred.test <- predict(m, newdata = X.test,type = "UK",
  covtype = "gauss" )$mean

  plot(ytest,y.pred.test)

y.pred.train <- predict(m, newdata = design,type = "UK",
covtype = "gauss" )$mean

  plot(ytrain,y.pred.train)

#####

n.test <- 1000

set.seed(0)

X <- maximinLHS(n,2*d)

A <- X[,-(d+1):(2*d)]

B <- X[,-1:-d]

X1 <- data.frame(A)

X2 <- data.frame(B)

p <- ncol(X1)

X <- rbind(X1, X2)

  for (i in 1:p) {

    Xb <- X2

    Xb[, i] <- X1[, i]

    X <- rbind(X, Xb)

  }

X <- data.frame(X)

write(X, file="G:/resultofinca/XdesignSA.txt",
```

```
ncolumns=n, sep=" ")

y.i <- read.table("G:/resultofinca/YdesignSA.csv",
header = FALSE,sep=",")

ytrain <- y.i[,1]

design <- read.table("G:/resultofinca/DesignSA.csv",
header = FALSE,sep=",")

###computation Cjs

n.test <- 100000

X.test1 <- randomLHS(n.test, d)

  colnames(X.test1) <- names(design)

X.test1[,1]<-X.test1[,1]*0.01

X.test1[,2]<-X.test1[,2]*0.01

X.test1[,3]<-X.test1[,3]*0.05

X.test1[,4]<-60*X.test1[,4]+80

X.test1[,6]<-X.test1[,6]*0.1

X.test1[,7]<-X.test1[,7]*0.05

  y.pred.1 <- predict(m, newdata = X.test1, type = "UK",
  covtype = "gauss")$mean

X.test2 <- randomLHS(n.test, d)

  colnames(X.test2) <- names(design)

X.test2[,1]<-X.test2[,1]*0.01

X.test2[,2]<-X.test2[,2]*0.01

X.test2[,3]<-X.test2[,3]*0.05

X.test2[,4]<-60*X.test2[,4]+80
```

```
X.test2[,6]<-X.test2[,6]*0.1
X.test2[,7]<-X.test2[,7]*0.05

y.pred.2 <- predict(m, newdata = X.test2, type = "UK",
covtype = "gauss")$mean
y.pred.C <- matrix(0, n.test, 7)
for(i in 1:7){
X.test.C <- X.test2
X.test.C[,i] <- X.test1[,i]
colnames(X.test.C) <- names(design)
y.pred.C[,i] <- predict(m, newdata = X.test.C, type = "UK",
covtype = "gauss")$mean
}

###To make matrix for all Cjs
yp <- cbind(y.pred.1,y.pred.2, y.pred.C[,1],y.pred.C[,2],y.pred.C[,3],
y.pred.C[,4],y.pred.C[,5],y.pred.C[,6],y.pred.C[,7])

yp <- data.frame(yp)

#### To compute Ujs
u=matrix(0,1,nrow=d)
for (i in 1:d){
sum = 0
for(j in 1:n.test){
sum = sum + yp[j,1]* yp[j,i+2]
}
u[i]=sum/n.test
```

```
    }

###To compute mean for predicted values

sum = 0

for(i in 1:n.test){

  sum = sum +yp[i,1]

}

Ey=sum/n.test

####To compute variance for predicted values

sum=0

for(i in 1:n.test){

  sum=sum +yp[i,1]^2

}

Ey2= sum/n.test

###To compute first order SIs

S = matrix(0,1,nrow=d)

for (i in 1:d){

  S[i]=(u[i]-(Ey^2))/(Ey2-(Ey^2))

}
```

# مراجع

- [۱] جانفدا م، (۱۳۹۰)، پایان‌نامه کارشناسی ارشد: ”بررسی روش‌های واریانس-مبنا در تحلیل حساسیت مدل‌های تعیینی“، دانشکده ریاضی، دانشگاه صنعتی شاهرود.
- [۲] رنچرا، (۱۳۸۲)، ”مدل‌های خطی برای آمار“، چاپ اول، آذرنوش ح، بزرگ‌نیا ا، انتشارات دانشگاه فردوسی مشهد، مشهد.
- [۳] محمدزاده م، (۱۳۹۱)، ”آمار فضایی و کاربردهای آن“، چاپ اول، انتشارات دانشگاه تربیت مدرس، تهران.
- [4] Bates R. A., Kennett R. S., Steinberg D. M. and Wynn H. P. (2006), ”Achieving robust design from computer simulations”, **Quality Technology and Quantitative Management**, 3, 161-177.
- [5] Bayarri M. J., Berger J. O. and Steinberg D. M. (2009), ”Special issue on computer modeling”, **Technometrics**, 51, 353-361.
- [6] Bayarri M. J., Berger J. O., Kennedy M., Kottas A., Paulo R., Sacks J., Cafeo J. A., Lin C. H., and Tu J. (2009), ”Predicting vehicle crashworthiness: calibration of computer models for functional and hierarchical data”, **Journal of the American Statistical Association**, 104, 929-943.
- [7] Blanning R. (1975), ”The construction and implementation of metamodels”, **Simulation**, 24, 177-184.

- 
- [8] Busby D., Farmer C. L. and Iske A. (2007), "Hierarchical nonlinear approximation for experimental design and statistical data fitting", **Society for Industrial and Applied Mathematics**, 29, 49-69.
- [9] Carnevale C., Finzi G., Guariso G., Pisoni E. and Volta M. (2012), "Surrogate models to compute optimal air quality planning policies at a regional scale", **Environmental Modelling and Software**, 34, 44-50.
- [10] Craig P. S., Goldstein M., Rougier J. C., and Seheult A. H. (2001), "Bayesian forecasting for complex systems using computer simulators", **Journal of the American Statistical Association**, 96, 717-729.
- [11] Cressie N. (1993), "Statistics for Spatial Data", **John Wiley and Sons**, New York.
- [12] Cressie N. (1990), "The Origins of Kriging", **Mathematical Geology**, 22, 239-252.
- [13] Diaconis P. (1988), "Baysian numerical analysis in Gupta S. S. and Berger J. O., editors, Statistical Decision Theory and Related Topics IV: Vol. 1, pages 163-176", **Springer-Verlag**, New York.
- [14] Dubrule O., (1983), "Two methods with different objectives: splines and kriging", **Mathematical Geology**, 15, 245-257.
- [15] Fang K. T., Li R. and Sudjanto A., (2006), "Design and Modeling for Computer Experiments", **Chapman and Hall**.
- [16] Friedman J. H. (1991), "Multivariate adaptive regression splines", **The Annals of Statistics**, 19, 1-141.

- 
- [17] Goldstein M. and Rougier J. C. (2006), "Bayes linear calibrated prediction for complex systems", **Journal of the American Statistical Association**, 101, 1132–1143.
- [18] Gramacy R. B. and Lian H. (2012), "Gaussian process single-index models as emulators for computer experiments", **Technometrics**, 54, 30–41.
- [19] Harari O. and Steinberg D. M. (2013), "Convex combination of gaussian processes for bayesian analysis of deterministic computer experiments", **Technometrics**, DOI: 10.1080/00401706.2013.361629.
- [20] Haylock R. G. E. (1997), PhD. thesis, "Bayesian inference about outputs of computationally expensive algorithms with uncertainty on the inputs ", MSc. depart., Nottingham University.
- [21] Hemyari P. and Nofziger D. L. (1987), "Analytical solution for punctual kriging in one dimension", **Soil Science Society of America journal**, 51, 268-269.
- [22] Homma T. and Saltelli A. (1996), "Importance measures in global sensitivity analysis of model output", **Reliability Engineering and System Safety**, 52, 1-17.
- [23] Jarvie H. P., Wade A. J., Butterfield D., Whitehead, P. G., Tindall C. I., Virtue W. A., Dryburgh W. and McCraw A. (2002), "Modelling nitrogen dynamics and distributions in the river Tweed, Scotland: an application of the INCA-N model", **J. Hydrol. Earth Sys. Sci.**, 6, 443-453.



- [24] Jin R., Chen W. and Sudjianto A. (2004), "Analytical metamodel-based global sensitivity analysis and uncertainty propagation for robust design", **Springer**, 60, 31-38.
- [25] Kennedy, M. C. and O'Hagan, A. (2001), "Bayesian calibration of computer models (with discussion)", **Journal of the Royal Statistical Society B**, 63, 425-464.
- [26] Kleijnen J. P. C. (1975), "A comment on Blanning's "metamodel for sensitivity analysis: the regression metamodel in simulation", **Interfaces**, 5, 21-23.
- [27] Kleijnen J. P. C. (1998), "Experimental Design for Sensitivity Analysis, Optimization, and Validation of Simulation Models", **John Wiley and Sons**, New York, 173-223.
- [28] Kleijnen J. P. C. (2010), "Design and Analysis of Simulation Experiments", **Springer**.
- [29] Kleijnen J. P. C. and van Beers W. C. M. (2005), "Robustness of kriging when interpolating in random simulation with heterogeneous variances: some experiments", **European Journal of Operational Research**, 165, 826-834.
- [30] Kolmogorov A. N. (1939), "Sur l'interpolation et Extrapolation des Suites Stationnaires", **Comptes Rendus Academie des Sciences**, Paris, 208, 2043-2045.

- [31] Kolmogorov A. N. (1941), "The distribution of energy in local isotropic turbulence", **Dok-lady Akademii Nauk SSSR, Moscow**, 32, 19-31.
- [32] Krige D. G. (1951), "A statistical approach to some basic mine valuation problems on the Witwatersrand", **Journal of the Chemical, Metallurgical and Mining Society of South Africa**, 52, 119-139.
- [33] Kuhnt S. and Steinberg D. M. (2010), "Design and analysis of computer experiments: editorial", **Advances in Statistical Analysis**, 94, 307-309.
- [34] Levy S. and Steinberg D. M. (2010), "Computer experiments: a review", **Advances in Statistical Analysis**, 94, 311-324.
- [35] Liu H. (2009), "Taylor Kriging Metamodeling For Simulation Interpolation, Sensitivity Analysis And Optimization", **A Dissertation Submitted to the Graduate Faculty of Auburn University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy**.
- [36] Matheron G. (1963), "Principles of Geostatistics", **Economic Geology**, 58, 1246-1266.
- [37] McKay M. D., Beckman R. J. and Conover W. J. (1979), "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code", **Technometrics**, 21, 239-245.
- [38] McKay M. D., Morrison J. D. and Upton S. C. (1999), "Evaluating prediction uncertainty in simulation models", **Computer Physics Communications**, 117, 44-51.

- [39] Meckesheimer M., Barton R. R., Limayem F. and Yannou B. (2001), "Meta-modeling of combined discrete or continuous responses", **AIAA Journal**, 39, 1950-1959.
- [40] Mitchell T. J. and Morris M. D. (1992), "The Spatial Correlation Function Approach to Response Surface Estimation", **Proceedings of the 1992 Winter Simulation Conference**, 565-571.
- [41] Oakley J. and O' Hagan A. (2004), "Probabilistic sensitivity analysis of complex models: a Bayesian approach", **Journal of the Royal Statistical Society: Series B**, 66, 751-769.
- [42] O'Hagan A. (1991), "Bayes-Hermite quadrature", **Journal of Statistical Planning and Inference**, 29, 245-260.
- [43] Ord J. K. (1983), "Kriging, Entry", **Encyclopedia of Statistical Science**, New York, 4,411-413.
- [44] Owen A. B. (1992), "Orthogonal arrays for computer experiments integeation and visualisation Statist", **Sinica**, 2, 439-452.
- [45] Ratto M., Castelletti A. and Pagano A. (2012), "Emulation techniques for the reduction and sensitivity analysis of complex environmental models", **Environmental Modelling software**, 34, 1-4.
- [46] Ratto M. and Pagano A. (2010), "Recursive algorithms for efficient identification of smoothing spline ANOVA models", **Advances in Statistical Analysis**, 94, 367-388.

- [47] Razavi S., Tolson B. A. and Burn D. H. (2012), "Numerical assessment of metamodelling strategies in computationally intensive optimization", **Environmental Modelling and Software**, 34, 67-86.
- [48] Razavi S., Tolson B. A. and Burn D. H. (2012), "Review of surrogate modeling in water resources", **Water Resource Research**, 48, W07401.
- [49] Roustant O., Ginsbourger D. and Deville Y. (2012), "DiceKriging, DiceOptim: Two R Packages for the Analysis of Computer Experiments by Kriging-Based Metamodeling and Optimization", **Journal of Statistical Software**, 51,104-124.
- [50] Rui T., Jeff Wu C. F. and Yu D. (2013), "Surrogate Modeling of Computer Experiments with Different Mesh Densities", **Technometrics**, DOI: 10.1080/00401706.2013.842935.
- [51] Sacks J. (1989), "Design analysis of computer experiments", **statistical Science**, 4,409-423.
- [52] Sacks J., Welch W. J., Mitchell T. J. and Wynn H. P. (1989), "Design and Analysis of Computer Experiments", **Statistical Science**, 4, 409-435.
- [53] Saltelli A., Annoni P., Azzini I., Campolongo F., Ratto M. and Tarantola S. (2010), "Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index", **J. Compute Physics Communications**, 181, 259-270.
- [54] Saltelli A. (2002), "Making best use of model valuations to compute sensitivity indices", **Computer Physics Communications**, 145, 280-297.

- [55] Saltelli A., Ratto M., Andres T., Campolongo F., Cariboni J., Gatelli D., Saisana M. and Tarantola S. (2008), "Global Sensitivity Analysis. The Primer".
- [56] Saltelli A., Tarantola S. and Chan K. (1999), "A quantitative model independent method for global sensitivity analysis of model output", **Technometrics**, 41, 39-56.
- [57] Saltelli A., Tarantola S., Ratto M. and Campolongo F. (1953), "Sensitivity analysis in practice :a guide to assessing scientific models", **Jone Wiley and Sons**.
- [58] Santner T., Williams B. and Notz W. (2003), "The Design and Analysis of Computer Experiments", **Springer**, New York, NY, USA.
- [59] Shahsavani D. and Grimvall A. (2009), "An Adaptive design and interpolation technique for extracting highly nonlinear response surface from deterministic models", **reliability Engineering and system safety**, 94(7), 1173-1182.
- [60] Shahsavani D. and Grimvall A. (2011), "Variance-based sensitivity analysis of model outputs using surrogate models", **Environmental Modelling and Software**, 26, 723-730.
- [61] Shahsavani D., Tarantola S. and Ratto M. (2010), "Evaluation of MARS modeling technique for sensitivity analysis of model output", **Procedia Social and Behavioral Sciences**, 2, 7737-7738.

- [62] Sobol I. M. (1993), "Sensitivity Analysis for Nonlinear Mathematical Models", **Mathematical Modeling and Computational Experiment**, 1, 407-414.
- [63] Storlie C. B. and Helton J. C. (2007), "Multiple predictor smoothing methods for sensitivity analysis: description of techniques", **Reliability Engineering and System Safety**, 93, 28-54.
- [64] Storlie C. B., Swiler L. P., Helton J. C. and Sallaberry C. J. (2009). "Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models", **Reliability Engineering and System Safety** , 94(11), 1735-1763.
- [65] Trochu F., Sacepe N., Volkov O. and Turenne S. (1999), "Characterization of NiTi Shape Memory Alloys Using Dual Kriging Interpolation", **Materials Science and Engineering**, 273-275, 395-399.
- [66] van Gigch J. (1991), "System Design Modeling and Metamodeling", **Springer**, New York, USA.
- [67] Villa-Vialaneix N., Follador M., Ratto M. and Leip A. (2012), "A comparison of eight metamodeling techniques for the simulation of N<sub>2</sub>O fluxes and N leaching from corn crops", **Environmental Modelling and Software**, 34, 51-66.
- [68] Wade A. J., Durand P., Beaujouan V., Wessel W. W., Raat K. J., Whitehead P. G., Butterfield D., Rankinen K. and Lepisto A., (2002), "A nitrogen model for European catchments: INCA, new model structure and equations", **J. Hydrology and Earth System Sciences**, 6, 559-582.

- 
- [69] Welch W., Buck R., Sacks J., Wynn H., Mitchell T. and Morris M. (1992), "Screening predicting and computer experiments", **Technometrics**, 34, 15-25.
- [70] Whitehead P. G., Wilson E. J. and Butterfield D., (1998), "A semi-distributed nitrogen model for multiple source assessments in catchments (INCA): Model structure and process equations", **J. Science of the Total Environment**, 210, 547-558.
- [71] Wiener N. (1949), "The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications", **Published jointly by the Technology Press of the Massachusetts Institute of Technology, Cambridge, Massachusetts, John Wiley and Sons**, 163-177.
- [72] Wold H. (1938), "A Study in the Analysis of Stationary Time Series", **Almqvist and Wiksell**, Stockholm, Sweden, 211-222.

## **Abstract**

Computer models contain many inputs, the real value of their are uncertain. In current methods of sensitivity analysis for assessing the effects of the input uncertainties on the output of the computer models are used from Monte Carlo methods. For computer model with complex and expensive computational, it is often not possible to get a large enough sample for a meaningful uncertainty analysis. In this thesis, we are doing sensitivity analysis with presentation kriging model as a surrogate model, that is estimator of computer model, based on the use of gaussian stochastic process models, in a Bayesian context. All of the results and measures provided by the Monte Carlo sensitivity analysis of the computer model are obtained with surrogate model, but using a far smaller sample. The accuracy of the kriging model is shown with an analytical function and a case study.

keywords: Computer models, Sensitivity analysis, Surrogate model, Gaussian stochastic, Kriging, Bayesian approach





Shahrood University Of Technology

Shahrood University  
Faculty Of Mathematical Sciences

Dissertation Submitted in Partial  
Fulfillment of The Requirements For The  
Degree of Master of Science in  
Statistics

**Sensitivity analysis of computer models  
using Bayesian surrogate model**

Supervisor

**Dr. Davood Shahsavani**

by

**Fatemeh Noormohammadzadeh**

February 2014