



دانشگاه صنعتی شریف

دانشکده فیزیک

تمرین سری (۱) درس موضوع‌های منتخب در فیزیک آماري

نام و نام خانوادگی: اشکان دماوندی

شماره دانشجویی: ۴۰۰۱۰۲۵۷۷

استاد درس: دکتر محمدرضا رحیمی تبار

نیم سال ۱۴۰۳-۲

۱ توصیف مسئله

یک دستگاه معادلات خطی به صورت $Ax = B$ داریم و می‌خواهیم جواب آن را با روش optimization به دست بیاوریم. باید norm دستگاه معادلات خطی را حساب کرده و این norm را minimize کنیم و برای این کار، می‌خواهیم از Gradient Decent Algorithm استفاده کنیم. در این کار، اگر ماتریس A یک ماتریس full-ranked نباشد، در دترمینان‌گیری از آن به مشکل می‌خوریم و الگوریتم حل از کار می‌افتد. برای رفع این مشکل پنج راه وجود دارد که عبارتند از: Ridge Regression، Pseudoinverse، QR Decomposition، Singular Value Decomposition (SVD) و روش‌های Iterative. هر کدام از این راه‌ها را توضیح می‌دهیم و مسئله را با راه Ridge Regression حل می‌کنیم.

۲ روش Ridge Regression

روش Ridge Regression به صورت کلی روشی برای تقریب زدن ضرایب مدل‌های رگرسیون چندگانه است که در زمان‌هایی که متغیرهای مستقل مسئله correlation زیادی دارند استفاده می‌شود. در این روش یک ترم رگرسیونی کوچک و مثبت به ماتریس مورد نظرمان (A) اضافه می‌شود که در این صورت داریم:

$$(A^T A + \lambda I)x = A^T B$$

در عبارت بالا، ترم $A^T A$ در زمانی که دترمینان ماتریس مورد نظر نزدیک صفر باشد باعث پایداری و رفع مشکلات معکوس‌ناپذیری ماتریس می‌شود.

۳ روش Pseudoinverse

روش Pseudoinverse روشی قوی برای حل سیستم‌هایی از معادلات است که جواب قطعی ندارند و برای یافتن نُرم اقلیدسی سیستم‌هایی معادلاتی که جواب‌های مختلف دارند، کاربردی است؛ حتی برای مواقعی که ماتریس‌های موجود full rank نباشند. این روش برای همهٔ ماتریس‌های مربعی که ورودی‌هایشان اعداد حقیقی و یا مختلط هستند تعریف می‌شود. هر ماتریس مربعی، یک pseudoinverse منحصر به خودش را دارد. pseudoinverse یک ماتریس را می‌توان با روش SVD نیز به دست آورد که این روش را در ادامه توضیح می‌دهیم.

۴ روش QR Decomposition

در روش QR Decomposition، ماتریس A به ضرب یک ماتریس متعامد Q در ماتریس بالامثلثی R تبدیل می‌شود ($A = QR$). این روش از لحاظ عددی پایدار است و برای ماتریس‌هایی که full rank نیستند هم قابل استفاده است.

۵ روش Singular Value Decomposition (SVD)

روش Singular Value Decomposition (SVD) روشی است که در آن یک ماتریس به سه ماتریس ساده‌تر تجزیه می‌شود که با چشم‌پوشی از singular value های کوچکی که ناپایداری ایجاد می‌کنند، باعث آسان‌تر شدن تحلیل و ایجاد تغییرات در آن می‌شود. SVD به ما کمک می‌کند تا مرتبه ماتریس را تحلیل کنیم و norm مناسبی استفاده کنیم که خطا را به حداقل مقدارش برساند.

۶ روش‌های Iterative

روش‌های Iterative روش‌هایی هستند که از یک مقدار اولیه برای تولید دنباله‌ای از جواب‌های تقریبی برای دسته‌ای از مسائل استفاده می‌کنند که در این دنباله تولید شده، هر کدام از جواب‌های تقریبی از قبلی‌ها درست شده‌اند. با استفاده از روش‌های Conjugate Gradient و یا GMRES که از جمله روش‌های Iterative هستند، دیگر نیازی به محاسبه مستقیم معکوس ماتریس‌ها، حتی در دستگاه‌های بزرگ و یا زمانی که ماتریس مورد نظرمان معکوس ناپذیر باشد، نداریم.

۷ حل مسئله با استفاده از Gradient Descent Algorithm و استفاده از

روش Ridge Regression

برای حل مسئله با استفاده از Gradient Descent Algorithm، ابتدا باید گرادیان تابع هزینه را بر حسب x محاسبه کنیم. تابع هزینه برابر است با:

$$L(x) = \|Ax - B\|^2$$

حال برای گرادیان تابع هزینه بر حسب x داریم:

$$\nabla_x L(x) = 2A^T(Ax - B)$$

با استفاده از رابطه گرادیان بالا، x را در جهتی خلاف جهت گرادیان تابع هزینه و با learning rate مشخصی مثل α عوض می‌کنیم (learning rate همان step size برای الگوریتم مسئله است). پس می‌توان گفت که برای x_{new} داریم:

$$x_{\text{new}} = x_{\text{old}} - \alpha \nabla_x L(x)$$

با توجه به اینکه ممکن است ماتریس A یک ماتریس full rank نباشد که باعث از کار افتادن الگوریتم حل مسئله می‌شود، باید از یکی از پنج راهی که گفتیم به دلخواه استفاده کنیم تا جواب پایدارتری به دست بیاوریم. از روش Ridge Regression استفاده می‌کنیم؛ در این روش یک تابع هزینه جدید تعریف می‌کنیم (به ماتریس موجود در تابع هزینه قبلی یک ترم رگرسیونی کوچک و مثبت اضافه می‌کنیم) که داریم:

$$L_R(x) = \|Ax - B\|^2 + \lambda \|x\|^2$$

این ترم λ اضافه شده، باعث می‌شود تا از ناپایدار شدن الگوریتم حل مسئله جلوگیری شود.

۸ کد استفاده شده برای حل مسئله

این به همراه توضیحاتش در پیوست ایمیل قرار گرفته است.