

بسم الله الرحمن الرحيم



دانشگاه حکیم بسزوری

دانشکده ریاضی و علوم کامپیوتر

پایان نامه برای دریافت درجه کارشناسی ارشد در رشته ریاضی کاربردی  
گرایش تحقیق در عملیات

## برش هوشمند تصویر

استاد راهنما

دکتر محمود امین طوسی

استاد مشاور

دکتر علیرضا قدسی

پژوهشگر:

فرزانه رشیدآبادی

۱۳۹۴/۱۰/۱۳



دانشگاه حکیم سبزواری

بسمه تعالی

فرم ۱۱۳ - ت

فرم ارزشیابی و صورتجلسه دفاع از پایان نامه کارشناسی ارشد\*

جلسه دفاع از پایان نامه خانم فرزانه رشیدآبادی دانشجوی رشته ریاضی کاربردی گرایش تحقیق در عملیات به شماره دانشجویی ۹۲۱۳۱۳۳۰۴۲ با عنوان برش هوشمند تصویر در مورخه ۱۳/۱۰/۱۳۹۴ در دانشکده ریاضی و علوم کامپیوتر تشکیل و توسط هیات داوران مورد ارزشیابی قرار گرفت و نمره...۱۹.۲۵... برابر درجه... عالی... برای آن تعیین گردید.

به این ترتیب از این تاریخ خانم فرزانه رشیدآبادی به عنوان کارشناس ارشد در رشته مذکور شناخته می شود.

مورد ارزشیابی	موارد	حداکثر نمره	نمره کسب شده
۱- کیفیت نگارش	رعایت اصول نگارش انسجام در تنظیم بخشهای مختلف، کیفیت تصاویر، جداول و اشکال، تنظیم فهرست ها، منابع و ماخذ	۴	۳,۷۵
۲- کیفیت علمی	بررسی تاریخچه و سابقه تجربی و نظری موضوع انسجام منطقی در بخش های مختلف پایان نامه، ابتکار و نوآوری، اهمیت و ارزش علمی پایان نامه، استفاده از منابع معتبر و جدید، کیفیت تجزیه و تحلیل یافته ها و نتیجه گیری، روشن بودن روش کار، هدف ها و فرضیه های تحقیق، جدید بودن روش تحقیق	۱۰	۱۰
۳- کیفیت ارائه در جلسه دفاع	تسلط بر موضوع و بیان واضح و تفهیم آن، توانایی در پاسخگویی به سوالات مطرح شده در جلسه، رعایت زمان ارائه، روش ارائه	۴	۴
۴- ارزشیابی گزارشات	گزارش های دوره ای بیشتر (کار حداقل ۴ مورد)	۱	۱
۵- خروجی پایان نامه	مقاله مستخرج از پایان نامه: این نمره به صورت زیر اختصاص می یابد (۱) چکیده کنفرانسی هر مورد ۲۵/نمره تا سقف ۵۰- نمره (۲) مقاله کامل در مجموع مقالات همایشهای معتبر یا مقاله در مجلات علمی-ترویجی معتبر پذیرفته شده یا چاپ شده هر مورد ۵۰/ نمره تا سقف ۱۰۰ نمره (۳) مقاله پذیرفته شده یا چاپ شده در مجلات علمی پژوهشی معتبر ۱ نمره (۴) مقاله ارسال شده مجلات علمی پژوهشی معتبر هر مورد ۲۵/ نمره تا سقف ۵۰- نمره (۵) دستگاه ساخته شده دارای گواهی ثبت اختراع یا به سفارش سازمان ها تا سقف ۱ نمره (۶) دستگاه ساخته شده کاربردی که به تایید رئیس دانشکده رسیده باشد تا سقف ۵۰/ نمره	۱	۶۵
جمع			۱۹,۲۵

درجه معادل کسب شده: (از ۱۹ تا ۲۰ عالی)  از ۱۸ تا ۱۹ بسیار خوب  از ۱۶ تا ۱۷ خوب  از ۱۴ تا ۱۵ قابل قبول  کمتر از ۱۴ غیر قابل قبول

مشخصات هیات داوران

ردیف	نام و نام خانوادگی	سمت	مرتبۀ علمی	محل کار	امضاء
۱	دکتر محمود امین طوسی	استاد راهنما	استادیار	دانشگاه حکیم سبزواری	
۲	دکتر علیرضا قدسی	استاد مشاور	استادیار	دانشگاه حکیم سبزواری	
۳	دکتر محمدعلی پرتانین	استاد داور	استادیار	دانشگاه حکیم سبزواری	
۴	دکتر قدیر صادقی	نماینده تحصیلات تکمیلی	دانشیار	دانشگاه حکیم سبزواری	

امضاء  
رئیس دانشکده: دکتر غلامرضا مقدسی

امضاء  
مدیر گروه: دکتر عبدالله قلی زاده

\* این فرم الزاما باید به صورت تایپ شده تهیه، ارسال و در پایان نامه درج شود



## سوگند نامه دانش آموختگان دانشگاه حکیم سبزواری

به نام خداوند جان و خرد      کزین برتر اندیشه بر نگذرد

اینک که به خواست آفریدگار پاک، کوشش خویش و بهره گیری از دانش استادان و سرمایه های مادی و معنوی این مرز و بوم، توشه ای از دانش و خرد گردآورده ام، در پیشگاه خداوند بزرگ سوگند یاد می کنم که در به کارگیری دانش خویش، همواره بر راه راست و درست گام بردارم. خداوند بزرگ، شما شاهدان، دانشجویان و دیگر حاضران را به عنوان داورانی امین گواه می گیرم که از همه دانش و توان خود برای گسترش مرزهای دانش بهره گیرم و از هیچ کوششی برای تبدیل جهان به جایی بهتر برای زیستن، دریغ نورزم. پیمان می بندم که همواره کرامت انسانی را در نظر داشته باشم و ممنوعان خود را در هر زمان و مکان تا سر حد امکان یاری دهم. سوگند می خورم که در به کارگیری دانش خویش به کاری که باراه و رسم انسانی، آیین پرهیزگاری، شرافت و اصول اخلاقی برخاسته از ادیان بزرگ الهی، به ویژه دین مبین اسلام، مبادنت دارد دست نیازم. همچنین در سایه اصول جهان شمول انسانی و اسلامی، پیمان می بندم از هیچ کوششی برای آبادانی و سرافرازی میهن و هم میهنانم فروگذاری نکنم و خداوند بزرگ را به یاری طلبم تا همواره در پیشگاه او و در برابر وجدان بیدار خویش و ملت سرافراز، بر این پیمان تا ابد استوار بمانم.

نام و نام خانوادگی: فرزانه رشیدآبادی

تاریخ و امضا:

## تأییدی صحت و اصالت نتایج

باسمه تعالی

اینجانب فرزانه رشیدآبادی به شماره دانشجویی ۹۲۱۳۱۳۳۰۴۲ دانشجوی رشته ریاضی کاربردی مقطع تحصیلی کارشناسی ارشد تأیید می‌نمایم که کلیه نتایج این پایان‌نامه حاصل کار اینجانب و بدون هرگونه دخل و تصرف است و موارد نسخه برداری شده از آثار دیگران را با ذکر کامل مشخصات منبع ذکر کرده‌ام. در صورت اثبات خلاف مندرجات فوق، به تشخیص دانشگاه مطابق با ضوابط و مقررات حاکم (قانون حمایت از حقوق مؤلفان و مصنفان و قانون ترجمه و تکثیر کتب و نشریات و آثار صوتی، ضوابط و مقررات آموزشی، پژوهشی و انضباطی ... ) با اینجانب رفتار خواهد شد و حق هرگونه اعتراض در خصوص احقاق حقوق مکتسب و تشخیص و تعیین تخلف و مجازات را از خویش سلب می‌نمایم. در ضمن، مسئولیت هرگونه پاسخگویی به اشخاص اعم از حقیقی و حقوقی و مراجع ذی صلاح (اعم از اداری و قضایی) به عهده ی اینجانب خواهد بود و دانشگاه هیچ گونه مسئولیتی در این خصوص نخواهد داشت.

نام و نام خانوادگی: فرزانه رشیدآبادی

تاریخ و امضا:

## مجوز بهره برداری از پایان نامه

بهره برداری از این پایان نامه در چهارچوب مقررات کتابخانه و با توجه به محدودیتی که توسط استاد راهنما

به شرح زیر تعیین می شود، بلامانع است:

بهره برداری از این پایان نامه برای همگان بلامانع است.

بهره برداری از این پایان نامه با اخذ مجوز از استاد راهنما، بلامانع است.

بهره برداری از این پایان نامه تا تاریخ ..... ممنوع است.

استاد راهنما: دکتر محمود امین طوسی

تاریخ و امضا:

تقدیم به:



## خدایی که آفرید

جهان را، انسان را، عقل را، علم را، عشق را  
و به کسانی که عشقشان را در وجودم دمید.



سپاس خدای را که سخنوران در ستودن او بمانند و شمارندگان شمردن نعمت‌های او ندانند.

بدون شک جایگاه و منزلت معلم، اجل از آن است که در مقام قدردانی از زحمات بی شائبه‌ی 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 می‌کند و سلامت امانت‌هایی را که به دستش سپرده‌اند، تضمین بر حسب وظیفه و از باب

”من لم یشکر المنعم من المخلوقین لم یشکر الله عز و جل“

در آغاز وظیفه خود می‌دانم از زحمات بی دریغ استاد فرهیخته جناب آقای دکتر محمود امین‌طوسی، صمیمانه قدردانی کنم که در کمال سعه‌صدر، با حسن خلق از هیچ کمکی در این عرصه بر من دریغ ننمودند و زحمت راهنمایی این رساله را بر عهده گرفتند؛ که قطعاً بدون راهنمایی‌های ارزنده ایشان، این مجموعه به انجام نمی‌رسید. از جناب آقای دکتر علیرضا قدسی که زحمت مشاوره این رساله را تقبل فرمودند و در آماده‌سازی این رساله، به نحو احسن اینجانب را مورد راهنمایی قرار دادند، کمال امتنان را دارم. در پایان بوسه می‌زنم بر دستان خداوندگاران مهر و مهربانی، پدر و مادر عزیزم و بعد از خدا ستایش می‌کنم وجود مقدس شان را. امروز هستی‌ام به امید شماست و فردا کلید باغ بهشت رضای شما. ره آوردی گران سنگ‌تر از این ارزان نداشتم تا به خاک پایتان نثار کنم، باشد که حاصل تلاشم نسیم گونه غبار خستگی‌تان را بزدايد. هم‌چنین تشکر می‌کنم از همسر عزیزم که مسیح‌وار با صبرش در تمامی لحظات رفیق راه بود. باشد که این خردترین بخشی از زحمات آنان را سپاس گوید.

فرزانه رشیدآبادی

۱۳۹۴/۱۰/۱۳

# فهرست مطالب

ج	فهرست علائم اختصاری
د	فهرست جداول
ه	فهرست تصاویر
و	فهرست الگوریتم‌ها
۱	چکیده
۲	پیش‌گفتار
۳	فصل ۱: مقدمه و پیش‌نیازها
۴	۱-۱ تعاریف اولیه
۹	۲-۱ رگرسیون خطی
۱۴	۱-۲-۱ رگرسیون ریبج
۱۵	۱-۱-۲-۱ نمایش اصلی رگرسیون ریبج
۱۶	۲-۱-۲-۱ نمایش دوگان رگرسیون ریبج
۱۷	۳-۱ نگاشت غیرخطی
۲۰	۴-۱ روش لاسو
۲۱	۱-۴-۱ مقدمات لاسو
۲۵	۵-۱ الگوریتم شوتینگ برای لاسو
۲۵	۱-۵-۱ نمونه یک متغیره: $n = 1$
۳۲	۲-۵-۱ حالت چند متغیره: $n > 1$
۳۷	۶-۱ چگونگی انتخاب $\lambda$

فصل ۲: برش هوشمند تصویر براساس روش یادگیری ۳۹

- ۱-۲ برآورد آلفا با یادگیری ۴۱ . . . . .
- ۲-۲ برآورد آلفا با یادگیری محلی ۴۲ . . . . .
- ۱-۲-۲ برآورد آلفا با به کارگیری رگرسیون ساده ۴۳ . . . . .
- ۲-۲-۲ برآورد آلفا با به کارگیری رگرسیون ریج ۴۶ . . . . .
- ۳-۲-۲ برآورد آلفا با به کارگیری تکنیک کرنل ۴۷ . . . . .
- ۳-۲ برآورد آلفا با یادگیری سراسری ۴۸ . . . . .

فصل ۳: استفاده از فرم بسته برای برش تصاویر ۵۱

- ۱-۳ تصاویر خاکستری ۵۱ . . . . .
- ۲-۳ تصاویر رنگی ۵۶ . . . . .
- ۳-۳ راهنمایی‌ها و محدودیت‌های کاربر ۶۰ . . . . .
- ۴-۳ برش تصویر با به کارگیری لاسو ۶۳ . . . . .

فصل ۴: روش پیشنهادی ۶۶

- ۱-۴ آشکارسازی بهتر شریان‌های کرونری قلب با یادگیری نیمه نظارتی خودکار ۶۶ . . . . .
- ۲-۴ ایجاد نقشه‌ی سه‌گانه خودکار ۶۶ . . . . .
- ۳-۴ پیاده‌سازی روش ۶۸ . . . . .

فهرست منابع ۷۰

پیوست آ: ۷۳

واژه نامه فارسی به انگلیسی ۸۲

واژه نامه انگلیسی به فارسی ۸۶



# فهرست علائم اختصاری

$()^T$ .....	ترانهاده
$\langle, \rangle$ .....	ضرب داخلی
$ \cdot $ .....	قدر مطلق
$\nabla$ .....	گرادیان
$\frac{\partial L(\beta, s)}{\partial \beta}$ .....	مشتق جزئی
$()^{-1}$ .....	معکوس
$\ \cdot\ _1$ .....	نرم ۱
$\ \cdot\ _2$ .....	نرم ۲

## فهرست جداول

۱-۳ گراف شش راسی و ماتریس لاپلاسیان آن ..... ۵۹

# فهرست تصاویر

- ۱ نمونه‌ای از کاربرد برش هوشمند تصویر در تولید فیلم . . . . . ۲
- ۱-۱ نمودار پراکندگی . . . . . ۱۰
- ۲-۱ نمایش رگرسیون ریج و لاسو . . . . . ۲۱
- ۳-۱ انتخاب ۸ با روش اعتبارسنجی متقابل . . . . . ۳۸
- ۱-۲ خروجی حاصل از نرم‌افزار Picture Cut Out و روش مبتنی بر یادگیری بر روی تصویر عروسک ۳۹
- ۲-۲ اجرای فرآیند نیمه‌نظارتی برش هوشمند بر روی تصویر عروسک . . . . . ۴۰
- ۳-۲ نقشه سه‌گانه . . . . . ۴۲
- ۴-۲ خروجی حاصل از برش هوشمند تصویر بر اساس روش یادگیری با به کارگیری رگرسیون ساده و رگرسیون ریج بر روی تصویر عروسک . . . . . ۴۷
- ۵-۲ خروجی حاصل از برش هوشمند تصویر بر اساس روش یادگیری با به کارگیری رگرسیون ساده و رگرسیون ریج بر روی تصویر گیاه . . . . . ۴۷
- ۶-۲ نمونه‌ای از برش هوشمند بر اساس روش یادگیری سراسری . . . . . ۴۹
- ۱-۳ خروجی حاصل از پیاده‌سازی روش فرم بسته . . . . . ۶۲
- ۲-۳ خروجی حاصل از برش هوشمند تصویر با به کارگیری لاسو بر روی تصویر عروسک . . . . . ۶۴
- ۳-۳ مقایسه خروجی حاصل از برش هوشمند تصویر با به کارگیری لاسو و روش فرم بسته . . . . . ۶۴
- ۱-۴ نمونه‌ای از تصاویر آنژیوگرافی و نمایش خروجی حاصل از روش برش هوشمند با استفاده از نقشه سه‌گانه پیشنهادی . . . . . ۶۷
- ۲-۴ مراحل تولید نقشه‌ی سه‌گانه خودکار متناظر با تصویر شریان‌های کرونری . . . . . ۶۸
- ۳-۴ ادامه مراحل تولید نقشه‌ی سه‌گانه خودکار متناظر با تصویر شریان‌های کرونری . . . . . ۶۸
- ۴-۴ خروجی روش برش هوشمند بر روی عروق کرونری . . . . . ۶۹
- ۵-۴ نمایش بهتر نتایج حاصل از روش برش هوشمند برای تصاویر آنژیوگرافی قلب . . . . . ۶۹

# فهرست الگوریتم‌ها

۱-۱ الگوریتم شوتینگ برای حل مسئله لاسو: ..... ۳۳



دانشگاه گیلان

## فرم چکیده ی پایان نامه ی دوره ی تحصیلات تکمیلی

مدیریت تحصیلات تکمیلی

نام خانوادگی دانشجو: رشیدآبادی	نام: فرزانه	ش. دانشجویی: ۹۲۱۳۱۳۳۰۴۲
استاد راهنما: دکتر محمود امین طوسی		
استاد مشاور: دکتر علیرضا قدسی		
دانشکده ریاضی و علوم کامپیوتر	رشته: ریاضی کاربردی	گرایش: تحقیق در عملیات
مقطع: کارشناسی ارشد	تاریخ دفاع: ۱۳۹۴/۱۰/۱۳	تعداد صفحات: ۹۰
عنوان پایان نامه: برش هوشمند تصویر		
کلید واژه ها: برش هوشمند تصویر، رگرسیون ریح، منظم سازی، داده های آموزشی		
<p>چکیده: شناسایی اشیا موجود در تصویر یکی از معروف ترین مسائل در حوزه پردازش تصویر و یادگیری ماشین است که کاربردهای عملی بسیاری دارد.</p> <p>شناسایی اشیا در تصویر ارتباط تنگاتنگی با قطعه بندی تصویر دارد، که یکی از روش های بسیار کارا در این زمینه Image matting است که ما از آن با نام برش هوشمند تصویر یاد می کنیم.</p> <p>برش هوشمند تصویر، فرآیند استخراج شی پیش زمینه از تصویر براساس محدودیت های ورودی کاربر است، که نقش مهمی در ویرایش فیلم و عکس ایفا می کند.</p> <p>در این پایان نامه، به بیان بعضی از روش های مشهور در این زمینه از جمله روش مبتنی بر یادگیری، روش فرم بسته می پردازیم. در پایان هم با به کارگیری این روش ها و اعمال تغییراتی در آن، روشی مناسب برای آشکارسازی بهتر شریان های کرونری قلب ارائه می دهیم.</p>		

## پیش‌گفتار

استخراج پس‌زمینه<sup>۱</sup> و پیش‌زمینه<sup>۲</sup> تصویر یکی از مسائل مهم پردازش تصویر است. برش هوشمند تصویر<sup>۳</sup> به معنای تخمین دقیق پیش‌زمینه در تصویر و ویدیو می‌باشد. این مسئله یکی از فنون کلیدی در بسیاری از کاربردهای تولید فیلم و ویرایش تصویر است که مطالعات زیادی را به خود معطوف نموده است [۱]. یکی از مرسوم‌ترین و قدیمی‌ترین روش‌های مونتاز فیلم، استفاده از زمینه ثابت یا تکنیک پرده آبی است. به عنوان مثال استخراج تصویر هنرپیشه از یک تصویر و قرار دادن آن در یک زمینه جدید از کاربردهای این مسئله است.



شکل ۱: نمونه‌ای از کاربرد برش هوشمند تصویر در تولید فیلم

یکی از کاربردهای حیاتی شناسایی اشیاء، شناسایی شریان‌های کرونری در تصاویر آنژیوگرافی قلب بیماران مبتلا به نارسایی قلبی است که این امر در تشخیص و درمان بیماری‌های قلبی اهمیت به‌سزایی دارد.

این پایان‌نامه شامل ۴ فصل است:

در فصل ۱ تعاریف و مفاهیم مورد نیاز را مطرح خواهیم نمود.

در فصل ۲ یکی از روش‌های نیمه‌نظارتی برش هوشمند تصویر تحت عنوان روش مبتنی بر یادگیری را بیان خواهیم نمود.

در فصل ۳ به بیان روش فرم بسته برای مسئله برش هوشمند تصویر می‌پردازیم.

در فصل ۴ روشی مناسب برای نمایش بهتر رگ‌ها در تصاویر آنژیوگرافی قلب بیماران مبتلا به نارسایی قلبی، ارائه می‌کنیم.

---

<sup>۱</sup>Background    <sup>۲</sup>Foreground    <sup>۳</sup>Image matting

# فصل ۱

## مقدمه و پیش نیازها

مسئله استخراج شی پیش زمینه از یک تصویر به عنوان برش هوشمند تصویر مطرح می شود. در [۲] روش های مختلف مواجهه با این مسئله در تصویر و ویدیو مورد بررسی قرار گرفته است، که مقاله مروری بسیار خوبی در این زمینه می باشد.

روش های اصلی را می توان به دسته های زیر تقسیم بندی نمود:

- روش فرم بسته <sup>۱</sup>
- روش مبتنی بر یادگیری <sup>۲</sup>
- روش مبتنی بر لاسو (نرم ۱) <sup>۳</sup>
- روش مبتنی بر معادله پواسن <sup>۴</sup>
- روش مبتنی بر نظریه بیز <sup>۵</sup>
- روش مبتنی بر قدم زدن تصادفی <sup>۶</sup>

در این پایان نامه به بیان روش مبتنی بر یادگیری ، روش فرم بسته و روش مبتنی بر لاسو خواهیم پرداخت، که برای فهم آن ها دانستن بعضی از مفاهیم ریاضی این روش ها ضروری است که در ادامه توضیح می دهیم.

---

<sup>۱</sup>The closed form solution    <sup>۲</sup>Learning based digital matting    <sup>۳</sup> $L_1$  matting    <sup>۴</sup>Poisson matting

<sup>۵</sup>Bayesian matting    <sup>۶</sup>Random walk matting

## ۱-۱ تعاریف اولیه

کوچک ترین جزء هر تصویر *pixel* است. پیکسل مخفف *picture element* به معنای المان تصویر است. یک تصویر متشکل از تعداد زیادی پیکسل است که در کنار یکدیگر قرار گرفته و هر کدام از این نقاط دارای مکان و مقدار مشخصی هستند.

برای نمایش یک تصویر  $M \times N$  از یک ماتریس که  $M$  سطر و  $N$  ستون دارد، استفاده می‌شود. مقدار هر عنصر نشان‌دهنده‌ی شدت روشنایی تصویر در آن نقطه است. هر عنصر می‌تواند مقداری بین ۰ و ۲۵۵ داشته باشد. مقدار صفر نشان‌دهنده‌ی رنگ سیاه و ۲۵۵ نشان‌دهنده‌ی رنگ سفید می‌باشد. یک تصویر RGB متشکل از سه ماتریس است که هر یک از آن‌ها مقادیر قرمز (Red) و سبز (Green) و آبی (Blue) تصویر رنگی را نگه می‌دارند. یک پیکسل زمانی مقدار خاکستری خواهد داشت که مؤلفه‌های  $B, G, R$  آن مقدار یکسانی داشته باشد.

معمولاً از لغت سطح خاکستری<sup>۱</sup> در ارتباط با شدت روشنایی<sup>۲</sup> تصاویر تک‌رنگ استفاده می‌شود. تصاویر رنگی از ترکیب تصاویر تک‌رنگ تشکیل می‌شوند. به عنوان مثال سیستم رنگی RGB، تصویر رنگی از سه تصویر (قرمز، سبز و آبی) تشکیل می‌شود.

آستانه‌گیری<sup>۳</sup>: هنگامی که قرار است از یک تصویر رنگی تصویر سیاه و سفید ساخته شود، برای این کار باید تک تک پیکسل‌ها با مقدار مشخصی مقایسه شده اگر بزرگ‌تر از آن مقدار بود به جای آن رنگ سفید و اگر کوچک‌تر بود به جای آن رنگ سیاه قرار می‌گیرد.

هیستوگرام تصویر: هیستوگرام در واقع همان نمودار میله‌ای افزونی هر رنگ می‌باشد. از هیستوگرام برای آستانه‌گیری استفاده می‌شود.

اعتبارسنجی متقابل<sup>۴</sup>: در این نوع اعتبارسنجی داده‌ها به  $K$  زیرمجموعه افزای می‌شوند. از این  $K$  زیرمجموعه، هر بار یکی برای اعتبارسنجی و  $K - 1$  تای دیگر برای مدل‌سازی به کار می‌روند. این روال  $K$  بار تکرار می‌شود و همه داده‌ها دقیقاً یک‌بار برای مدل‌سازی و یک‌بار برای اعتبارسنجی به کار می‌روند. در نهایت میانگین نتیجه این  $K$  بار اعتبارسنجی به عنوان یک تخمین نهایی برگزیده می‌شود. به طور معمول از  $K = 10$  استفاده می‌شود.

بیش‌برازش<sup>۵</sup>: به پدیده‌ی نامطلوبی در آمار گفته می‌شود که در آن درجه آزادی مدل بیش‌تر از درجه آزادی واقعی انتخاب شده و در نتیجه اگر چه مدل روی داده‌های استفاده شده برای یادگیری بسیار خوب نتیجه می‌دهد،

---

<sup>۱</sup>Gray    <sup>۲</sup>Intensity    <sup>۳</sup>Threshold    <sup>۴</sup>Cross validation    <sup>۵</sup>Overfitting



اما بر روی داده جدید دارای خطای زیاد است. منظم‌سازی<sup>۱</sup> از راه‌های مقابله با این پدیده است.

نکته: منظور از  $\operatorname{argmin}_x f(x)$  پیدا کردن مقدار  $x$  است به طوری که به ازای آن تابع مینیمم شود. به طور مثال

$$\min[x^2 + 1] = 1$$

در حالی که

$$\operatorname{argmin}[x^2 + 1] = 0$$

می‌باشد.

**تعریف ۱-۱-۱.** ترانواده: فرض کنید  $A$  یک ماتریس  $m \times n$  باشد. ترانواده<sup>۲</sup> ماتریس  $A$  با نماد  $A^T$  نمایش داده می‌شود، که یک ماتریس  $n \times m$  است که به صورت زیر تعریف می‌شود [۳].

$$A = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \iff A^T = (a_{ji})_{\substack{1 \leq j \leq n \\ 1 \leq i \leq m}}$$

**قضیه ۱-۱-۲.** فرض کنید

$$\begin{aligned} T : M_{m \times n} &\longrightarrow M_{n \times m} \\ A &\longrightarrow A^T \end{aligned}$$

باشد، آن‌گاه روابط زیر برقرار می‌باشند:

$$A^{TT} = A \bullet$$

$$(A + B)^T = A^T + B^T \bullet$$

$$(AB)^T = B^T A^T \bullet \quad \text{اگر } n = m \text{ باشد.}$$

$$(A^{-1})^T = (A^T)^{-1} \bullet \quad \text{اگر } A \text{ معکوس پذیر و } n = m \text{ باشد.}$$

□

برهان. [۳].

<sup>۱</sup>Regularization

<sup>۲</sup>Transpose

تعریف ۱-۱-۳. الف)  $\|\cdot\|_1$  نرم  $L_1$  در  $\mathbb{R}^n$  است و به صورت زیر تعریف می‌شود:

$$\forall X = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n; \quad \|X\|_1 = \sum_{k=1}^n |x_k|$$

ب)  $\|\cdot\|_2$  نرم  $L_2$  در  $\mathbb{R}^n$  است و به صورت زیر تعریف می‌شود:

$$\forall X = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n; \quad \|X\|_2^2 = X^T X = \sum_{k=1}^n x_k^2$$

تعریف ۱-۱-۴. اگر  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  تابعی مشتق‌پذیر باشد، آن‌گاه گرادیان تابع را به صورت زیر به دست می‌آوریم:

$$\nabla f = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)^T \in \mathbb{R}^{n \times 1}$$

و مشتق مرتبه دوم تابع را با نماد  $\nabla^2 f = H$  نمایش داده و از آن با نام هسیان<sup>۱</sup> تابع نام می‌بریم که به صورت زیر محاسبه می‌شود:

$$H(x) = \begin{bmatrix} f_{11}(x) & f_{12}(x) & \dots & f_{1n}(x) \\ f_{21}(x) & f_{22}(x) & \dots & f_{2n}(x) \\ \vdots & \vdots & \dots & \vdots \\ f_{n1}(x) & f_{n2}(x) & \dots & f_{nn}(x) \end{bmatrix} \in \mathbb{R}^{n \times n} \quad (1-1)$$

و اگر  $f = (f_1, f_2, \dots, f_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  نگاهی مشتق‌پذیر باشد، آنگاه

$$\nabla f = \frac{\partial (f_1, f_2, \dots, f_m)}{\partial (x_1, x_2, \dots, x_n)} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m \times n}. \quad (2-1)$$

قضیه ۱-۱-۵. فرض کنید  $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  دو تابع مشتق‌پذیر به صورت  $f = (f_1, f_2, \dots, f_m)^T$  و  $g = (g_1, g_2, \dots, g_m)^T$  باشند، و تابع  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  به صورت  $h(x) = f(x)^T g(x)$  تعریف شده باشد.

<sup>۱</sup>Hessian

آن‌گاه تابع  $h$  مشتق‌پذیر است و مشتق آن از رابطه‌ی زیر به دست می‌آید:

$$\nabla h(x) = \nabla f(x)^T g(x) + \nabla g(x)^T f(x)$$

□

برهان. [۴]

قضیه ۱-۱-۶. قضیه تیلور: فرض کنید  $f$  تابعی باشد که خود و  $n$  مشتق اولش بر بازه بسته  $[a, b]$  پیوسته هستند. علاوه بر این، فرض کنید  $f^{(n+1)}(x)$  به ازای هر  $x$  در بازه‌ی باز  $(a, b)$  وجود داشته باشد. آن‌گاه عددی مانند  $\xi$  در بازه‌ی باز  $(a, b)$  وجود دارد، به طوری که رابطه‌ی

$$(3-1) \quad f(b) = f(a) + \frac{f'(x)}{1!}(b-a) + \frac{f''(x)}{2!}(b-a)^2 + \dots + \frac{f^{(n)}(x)}{n!}(b-a)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(b-a)^{n+1}$$

برقرار است.

□

برهان. [۵]

تعریف ۱-۱-۷. مسئله بهینه‌سازی زیر مفروض است:

$$\text{Min } f(x) \quad (4-1)$$

st.

$$g(x) \geq 0$$

که  $f(x)$  و  $g(x)$  دو تابع مشتق‌پذیر هستند. تابع لاگرانژ<sup>۱</sup> به صورت زیر تعریف می‌شود:

$$L(x, \lambda) = f(x) - \lambda g(x) \quad \lambda \geq 0 \quad (5-1)$$

جواب مسئله بهینه‌سازی ۱-۴، به وسیله بهینه‌سازی تابع لاگرانژ ۱-۵ نسبت به  $x$  و  $\lambda$  حاصل می‌شود به طوری که  
اولاً

$$\nabla f(x) - \lambda \nabla g(x) = 0$$

---

<sup>۱</sup>Lagrange

و ثابثاً شرایط زیر نیز برقرار باشند:

$$g(x) \geq 0$$

$$\lambda \geq 0$$

$$\lambda g(x) = 0$$

این شرایط به عنوان شرایط کروش کان تا کر<sup>۱</sup> شناخته می‌شوند [۶].

**تعریف ۱-۱-۸.** شبه معکوس<sup>۲</sup>: فرض می‌کنیم ماتریس  $A \in \mathbb{R}^{m \times n}$  باشد. آنگاه ماتریس  $A^\dagger \in \mathbb{R}^{n \times m}$  شبه معکوس ماتریس  $A$  است، اگر

$$AA^\dagger A = A$$

و ماتریس‌های  $U \in \mathbb{R}^{n \times n}$  و  $V \in \mathbb{R}^{m \times m}$  موجود باشند به طوری که

$$A^\dagger = UA^T$$

$$A^\dagger = A^T V$$

برقرار باشند [۴].

شرایط لازم  $A^\dagger = UA^T = A^T V$  را می‌توان به صورت زیر تعبیر نمود. هر سطر از ماتریس شبه معکوس  $A^\dagger$  از  $A$  ترکیب خطی از سطرهاى  $A^T$  است و هر ستون  $A^\dagger$  ترکیب خطی از ستون‌های  $A^T$  است.

**قضیه ۱-۱-۹.** فرض می‌کنیم ماتریس  $A \in \mathbb{R}^{m \times n}$  و  $m \geq n$  و  $rank(A) = n$  باشد. آنگاه شبه معکوس  $A$  به صورت

$$A^\dagger = (A^T A)^{-1} A^T$$

حاصل می‌شود.

**برهان.** اولاً  $A(A^T A)^{-1} A^T A = A$  است. اگر  $U = (A^T A)^{-1}$  و  $V = A(A^T A)^{-1} A^T$  قرار

<sup>۱</sup>Karush Kuhn Tucker

<sup>۲</sup>Pseudo-inverse

دهیم، آن‌گاه  $A^\dagger = UA^T = A^T V$  می‌باشد. در حقیقت

$$A^\dagger A = I$$

می‌باشد [۴]. □

قضیه ۱-۱-۱۰. اگر ماتریس  $A \in \mathbb{R}^{m \times n}$  و  $m \leq n$  و  $\text{rank}(A) = m$  باشد، آن‌گاه شبه‌معکوس  $A$  به صورت

$$A^\dagger = A^T (AA^T)^{-1}$$

خواهد بود.

برهان. [۴] □

## ۲-۱ رگرسیون خطی

مجموعه‌ی آموزشی<sup>۱</sup> داده شده‌ی  $S = \{(x_1, y_1), \dots, (x_l, y_l)\}$  از نقاط  $x_i$  از  $X \subseteq \mathbb{R}^n$  با برچسب‌های نظیر به نظیر  $y_i$  در  $Y \subseteq \mathbb{R}$  رادر نظر می‌گیریم. هدف یافتن یک تابع خطی مانند  $g(\cdot)$  به صورت زیر است:

$$g(\mathbf{x}) = \langle \beta, \mathbf{x} \rangle = \beta^T \mathbf{x} = \sum_{i=1}^n \beta_i x_i.$$

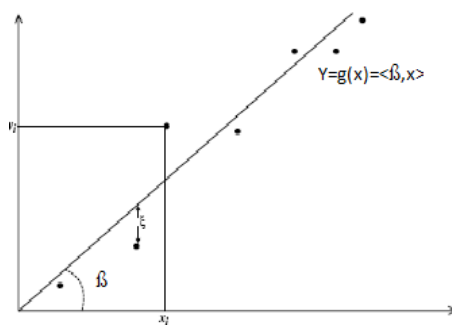
و ایجاد یک تابع الگو که باید تقریباً معادل صفر باشد:

$$f(\mathbf{x}, y) = |y - g(\mathbf{x})| = |y - \langle \beta, \mathbf{x} \rangle| \approx 0.$$

در واقع دنبال یافتن  $\beta$  بی‌هستیم که در روابط فوق صدق کند، که بهترین درونیابی مجموعه‌ی آموزشی<sup>۲</sup> داده شده‌ی  $S$  است. این به طور طبیعی یکی از ساده‌ترین روابطی است که می‌تواند در مبدا  $X \times Y$  یافت شود، یعنی یک تابع خطی  $g$  از  $X$  مقدار نظیر به نظیر  $Y$  را تنظیم می‌کند. از لحاظ هندسی متناظر ساختن به منظور برازش یک ابر صفحه از میان نقاط داده  $n$ -بعدی می‌باشد. که از نمادگذاری  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$  برای بردارهای داده  $n$ -بعدی استفاده می‌کنیم [۷]. شکل ۱-۱ یک مثال برای  $n = 1$  را نشان می‌دهد.

<sup>۱</sup>Training set

<sup>۲</sup>Training set



شکل ۱-۱: نمودار پراکندگی

در نمونه دقیق، هنگامی که داده‌ها به صورت  $(x, g(x))$  تولید شده باشند که  $g(x) = \langle \beta, x \rangle$  و دقیقاً  $l = n$  نقطه مستقل خطی وجود داشته باشد، پارامتر  $\beta$  به وسیله حل سیستم معادلات خطی زیر حاصل می‌شود:

$$X\beta = Y.$$

که از نماد  $X$  برای نمایش ماتریسی که سطرهای آن بردارهای سطری  $x_1^T, \dots, x_l^T$  و  $Y$  برای نمایش بردار  $(y_1, \dots, y_l)^T$  استفاده می‌شود. در واقع داده‌ها بردارهای ستونی اند اما در ماتریس  $X$  به عنوان بردارهای سطری ذخیره می‌شوند. اگر تعداد نقاط کمتر از بعد باشد، ممکن است چندین  $\beta$  وجود داشته باشد که دقیقاً داده‌ها را توصیف کند. یک معیار برای انتخاب بین آن‌ها لازم است، که در این حالات  $\beta$  با نرم کمینه مطلوب است [۷]. اگر تعداد نقاط بیشتر از بعد باشد و در فرآیند تولید داده‌ها، نویز وجود داشته باشد پس نمی‌توان انتظار داشت الگو دقیق باشد بنابراین معیاری لازم است. در این وضعیت الگویی با کمترین خطا را انتخاب می‌کنیم. به طور کلی اگر داده‌ها همراه با خطا باشد ترکیبی از دو استراتژی به کار گرفته می‌شود: یافتن بردار  $\beta$  ای با کمترین نرم و کمترین خطا.

در شکل ۱-۱ فاصله نشان داده شده با  $\xi$  نشان دهنده خطای تابع خطی در مثال آموزشی است، یعنی

$$\xi = (y - g(x))$$

$$f((x, y)) = |y - g(x)| = |\xi|.$$

می‌باشد. هدف یافتن تابعی است که مجموع خطاها روی داده‌های آموزشی کمینه باشد.

مجموع مربعات خطا رایج ترین معیار انتخاب اختلاف بین داده‌های آموزشی و تابع موردنظر ( $g(\cdot)$ ) است:

$$L(g, S) = L(g, \beta) = \sum_{i=1}^l (y_i - g(x_i))^2 = \sum_{i=1}^l \xi_i^2 = \sum_{i=1}^l L((x_i, y_i), g).$$

از نمادگذاری  $L((x_i, y_i), g) = \xi_i^2$  برای نمایش مربعات خطا یا اختلاف  $g$  در نمونه  $(x_i, y_i)$  و از نماد  $L(f, S)$  برای نمایش مجموع اختلاف تابع  $f$  بر مجموعه‌ی آموزشی  $S$  استفاده می‌شود. مسئله یادگیری اکنون به برآورد بردار  $\beta$  که کمینه خطا را دارد، تبدیل می‌شود. این مسئله توسط گاوس<sup>۱</sup> و با عنوان تقریب کمترین مربعات<sup>۲</sup> معرفی شد [۷].

با به کارگیری نمادگذاری بالا، بردار اختلاف خروجی را می‌توان به صورت زیر نوشت:

$$\xi = Y - \mathbf{X}\beta.$$

بنابراین تابع خطا<sup>۳</sup> به صورت زیر حاصل می‌شود:

$$L(\beta, s) = \|\xi\|_2^2 = (Y - \mathbf{X}\beta)^T (Y - \mathbf{X}\beta). \quad (6-1)$$

مقدار بهینه  $\beta$  به وسیله مشتق‌گیری از تابع خطا نسبت به پارامتر  $\beta$  و برابر صفر قرار دادن آن، تعیین می‌شود:

$$\frac{\partial L(\beta, s)}{\partial \beta} = -2\mathbf{X}^T Y + 2\mathbf{X}^T \mathbf{X}\beta = 0.$$

بنابراین معادلات زیر حاصل می‌شوند که به آن معادلات نرمال<sup>۴</sup> می‌گویند.

$$\mathbf{X}^T \mathbf{X}\beta = \mathbf{X}^T Y. \quad (7-1)$$

اگر  $\mathbf{X}^T \mathbf{X}$  معکوس‌پذیر<sup>۵</sup> باشد، جواب مسئله کمترین مربعات به صورت زیر حاصل می‌شود:

$$\beta = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T Y.$$

---

<sup>۱</sup>Gauss    <sup>۲</sup>Least squares approximation    <sup>۳</sup>Loss function    <sup>۴</sup>normal equation    <sup>۵</sup>Invertible

پیش‌بینی خروجی برای نقاط داده‌ی جدید می‌تواند با به کارگیری تابع پیش‌بینی محاسبه شود:

$$g(x) = \langle \beta, x \rangle$$

ملاحظه ۱-۲-۱ (نمایش دوگان<sup>۱</sup>). اگر معکوس  $\mathbf{X}^T \mathbf{X}$  موجود باشد می‌توان  $\beta$  را به روش زیر نیز بیان کرد:

$$\beta = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T Y = \mathbf{X}^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-2} \mathbf{X}^T Y = \mathbf{X}^T \alpha = \langle \mathbf{X}, \alpha \rangle$$

که در آن

$$\alpha = \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-2} \mathbf{X}^T Y.$$

می‌باشد  $[N]$ .

ملاحظه ۲-۲-۱ (شبه معکوس). اگر  $\mathbf{X}^T \mathbf{X}$  منفرد<sup>۲</sup> (معکوس ناپذیر) باشد، در این حالت برای یافتن جواب معادله (۷-۱) از شبه معکوس استفاده می‌کنیم.

تعریف ۳-۲-۱. دستگاه معادلات خطی زیر را در نظر می‌گیریم:

$$A_{n \times n} x_{n \times 1} = b_{n \times 1}$$

با فرض اینکه  $A$  غیر منفرد باشد، خواهیم داشت:

$$x = A^{-1} b$$

حال اگر  $b$  شامل نویز یا خطای محاسباتی ناشی از گرد کردن مانند  $\Delta b$  باشد، در این صورت این خطا به صورت زیر در پاسخ ظاهر خواهد شد:

$$x + \Delta x = A^{-1} (b + \Delta b)$$

لذا می‌توان نوشت:

$$\Delta x = A^{-1} \Delta b$$

<sup>۱</sup>Dual representation

<sup>۲</sup>Singular



با توجه به خواص نرم ماتریس‌ها نتیجه می‌شود:

$$\|\Delta x\|_p \leq \|A^{-1}\|_p \|\Delta b\|_p.$$

از عبارت اخیر می‌توان تعبیر کرد که، اگر  $\|A^{-1}\|_p$  مقدار کوچکی باشد، برای تغییرات کم در  $b$  یعنی  $\|\Delta b\|_p$  مقدار  $\|\Delta x\|_p$  کم خواهد بود. ولی برای  $\|A^{-1}\|_p$  های بزرگ، مقدار  $\|\Delta x\|_p$  بزرگ است، حتی اگر  $\|\Delta b\|_p$  مقدار کوچکی باشد.

لذا برای تشخیص بد حالت بودن یک دستگاه معادلات، پارامتری به نام عدد حالت تعریف می‌گردد:

$$\kappa = \|A\|_p \|A^{-1}\|_p \quad \kappa \geq 1$$

اگر عدد حالت کوچک باشد، بیان‌کننده آن است که ماتریس  $A$  و دستگاه معادلات حاصل خوش حالت (خوش وضع)<sup>۱</sup> است و اگر عدد حالت مقدار خیلی بزرگی باشد، بیان‌گر آن است که ماتریس نزدیک به منفرد شدن است، لذا آن ماتریس را بد حالت (بد وضع)<sup>۲</sup> نامیده و خطای محاسباتی در معکوس کردن ماتریس  $A$  زیاد است [۸].

**تعریف ۱-۲-۴.** مسائل بد طرح: در اوایل قرن پیش، هادامارد<sup>۳</sup> اولین کسی بود که بحث دستگاه‌های معادلات خوش طرح<sup>۴</sup> و بد طرح<sup>۵</sup> را مطرح نمود. ایشان دستگاه معادلاتی را که در آن‌ها سه شرط وجود، یکتایی و پایداری جواب تأمین باشند، را مسائل خوش طرح و مسائلی که در آن‌ها یکی از سه شرط نقض گردیده باشد، را مسائل بد طرح نام‌گذاری کردند.

در بسیاری از عملیات جبرخطی، لازم است یک تقریب مناسب مانند  $\hat{x}$  از بردار  $x \in \mathbf{R}^n$  به دست آورده به طوری که در معادله‌ی تقریبی  $Ax \approx b$  صدق کند با این وجود که  $A \in \mathbf{R}^{m \times n}$  منفرد بوده یا مسئله بد-وضع باشد. در این وضعیت بردار  $\tilde{x} = A^{-1}b$  (یا در نمونه‌های رتبه کامل فرامعین  $A^\dagger b$  با شبه‌معکوس  $A^\dagger = (A^T A)^{-1} A^T$ ) معمولاً یک تقریب نامعتبر برای  $x$  است. علاوه بر این حتی زمانی که بردار  $\hat{x}$  یک تقریب قابل قبول برای بردار  $x$  با  $Ax = b$  باشد، خطای برآورد به صورت زیر است [۹]:

$$\|x - \hat{x}\| \leq \|A^{-1}\| \|A\hat{x} - b\|$$

<sup>۱</sup>Well-conditioned

<sup>۲</sup>Ill-conditioned

<sup>۳</sup>Hadamard

<sup>۴</sup>Well-posed

<sup>۵</sup>Ill-posed

و در حالت فرامعین نیز خطا به صورت زیر خواهد بود [۹]:

$$\|x - \hat{x}\| \leq \|A^\dagger\| \|A\hat{x} - b\|$$

حالت‌هایی وجود دارد که برازش دقیق داده‌ها ممکن نیست یا حتی داده‌ها برای به دست آوردن ماتریس  $X^T X$  که معکوس‌پذیر باشد، کافی نیستند یا ممکن است خطایی در داده‌ها باشد که تلاش برای رسیدن به خروجی بی‌فایده باشد.

در مسائل بد-طرح بعضی پارامترها با روش کمترین مربعات به خوبی تعیین نمی‌شوند. به ویژه زمانی که تعداد پارامترها بیشتر از کمیت‌های معلوم باشد، بنابراین در این وضعیت روش کمترین مربعات استاندارد کارآمد نیست، در این حالت‌ها روشی که غالباً به کار گرفته می‌شود منظم‌سازی (پایدارسازی) است. روش منظم‌سازی برای به دست آوردن برآورد معناداری از جواب برای مسائل بد-طرح به کار گرفته می‌شود [۹].

منظم‌سازی در ریاضیات، آمار و به ویژه در شاخه‌هایی از یادگیری ماشین و مسائل معکوس برای فرآیند به کارگیری اطلاعات اضافی به منظور حل مسائل بد-طرح یا به منظور جلوگیری از بیش‌برازش، مطرح می‌شود. روش منظم‌سازی تیخونوف به افتخار آندری تیخونوف<sup>۱</sup> نام‌گذاری شده است و یک روش رایج به کارگیری منظم‌سازی برای مسائل بد-طرح است. این روش همچنین با نام‌های رگرسیون ریج<sup>۲</sup> و تیخونوف-میلر یا روش معکوس خطی مقید نیز عنوان می‌شود.

## ۱-۲-۱ رگرسیون ریج

رگرسیون ریج متناظر با حل مسئله بهینه‌سازی زیر است [۷]: (می‌توان با یک ضریب وزنی تاثیر تابع خطا و نرم  $\beta$  را متفاوت در نظر گرفت.)

$$\operatorname{argmin}_{\beta} L_{\lambda}(\beta) = \lambda \|\beta\|^2 + \sum_{i=1}^l (y_i - g(x_i))^2 \quad (۸-۱)$$

یا

$$\operatorname{argmin}_{\beta} L_{\lambda}(\beta) = \lambda \|\beta\|^2 + \|Y - X\beta\|_2^2 \quad (۹-۱)$$

<sup>۱</sup>Andrey Tikhonov

<sup>۲</sup>Ridge regression

$\lambda$  یک عدد مثبت است که میزان اهمیت<sup>۱</sup> بین نرم و تابع خطا را مشخص می‌کند و بنابراین درجه‌ی منظم‌سازی را نیز کنترل می‌کند.

لذا مسئله یادگیری به حل یک مسئله بهینه‌سازی در  $\mathbb{R}^n$  تبدیل می‌شود.

برای حل مسئله بهینه‌سازی ۱-۹ دو روش وجود دارد: روش اول برآوردی از  $\beta$  را براساس نمایش اصلی رگرسیون ریبج و روش دوم برآوردی از  $\beta$  را مبتنی بر نمایش دوگان ارائه می‌دهد، که در قسمت بعد این دو روش را توضیح خواهیم داد.

### ۱-۱-۲-۱ نمایش اصلی رگرسیون ریبج

به صورتی مشابه با روش بدست آوردن  $\beta$  در رگرسیون خطی با مشتق‌گیری از تابع هدف (۱-۹) نسبت به پارامتر  $\beta$  و برابر صفر قرار دادن آن داریم [۷]:

$$\frac{\partial L_\lambda(\beta)}{\partial \beta} = \lambda \beta - \mathbf{X}^T \mathbf{Y} + \mathbf{X}^T \mathbf{X} \beta = 0.$$

$$\mathbf{X}^T \mathbf{X} \beta + \lambda \beta = (\mathbf{X}^T \mathbf{X} + \lambda I_n) \beta = \mathbf{X}^T \mathbf{Y} \quad (10-1)$$

که  $I_n$  ماتریس همانی  $n \times n$  است. در این حالت ماتریس  $(\mathbf{X}^T \mathbf{X} + \lambda I_n)$  معکوس پذیر است اگر  $\lambda > 0$  باشد، بنابراین جواب به صورت زیر حاصل می‌شود:

$$\beta = (\mathbf{X}^T \mathbf{X} + \lambda I_n)^{-1} \mathbf{X}^T \mathbf{Y} \quad (11-1)$$

در نتیجه تابع پیش‌بینی<sup>۲</sup> به صورت زیر است:

$$g(\mathbf{x}) = \langle \beta, \mathbf{x} \rangle = \beta^T \mathbf{x} = \mathbf{Y}^T \mathbf{X} (\mathbf{X}^T \mathbf{X} + \lambda I_n)^{-1} \mathbf{x}$$

<sup>۱</sup>trade off

<sup>۲</sup>Prediction function

## ۲-۱-۲-۱ نمایش دوگان رگرسیون ریج

مجددا رابطه (۱۰-۱) را در نظر می‌گیریم:

$$\mathbf{X}^T \mathbf{X} \beta + \lambda \beta = (\mathbf{X}^T \mathbf{X} + \lambda I_n) \beta = \mathbf{X}^T Y$$

با بازنویسی آن بر حسب  $\beta$  داریم:

$$\beta = \lambda^{-1} \mathbf{X}^T (Y - \mathbf{X} \beta)$$

که اگر  $\alpha$  را به صورت زیر در نظر بگیریم:

$$\alpha = \lambda^{-1} (Y - \mathbf{X} \beta)$$

خواهیم داشت:

$$\beta = \mathbf{X}^T \alpha$$

در ادامه خواهیم دید که می‌توان  $g(x)$  را بر حسب ضرب داخلی  $x$  و دادگان آموزشی به دست آورد:

$$\alpha = \lambda^{-1} (Y - \mathbf{X} \beta)$$

$$\Rightarrow \lambda \alpha = (Y - \mathbf{X} \mathbf{X}^T \alpha)$$

$$\Rightarrow (\mathbf{X} \mathbf{X}^T + \lambda I_l) \alpha = Y$$

$$\Rightarrow \alpha = (G + \lambda I_l)^{-1} Y \quad (۱۲-۱)$$

که  $G = \mathbf{X} \mathbf{X}^T$  یا به صورت مؤلفه‌ای  $G_{ij} = \langle x_i, x_j \rangle$  است.

بنابراین در نمایش دوگان

$$\beta = \mathbf{X}^T (G + \lambda I_l)^{-1} Y \quad (۱۳-۱)$$

است، با فرض این که  $k_i = \langle x_i, x \rangle$  باشد در نتیجه تابع پیش‌بینی به صورت زیر حاصل می‌شود:

$$\begin{aligned} g(x) &= \langle \beta, x \rangle = \langle \sum_{i=1}^l \alpha_i x_i, x \rangle = \sum_{i=1}^l \alpha_i \langle x_i, x \rangle \\ &= \sum_{i=1}^l \alpha_i k_i = \langle \alpha, K \rangle = \alpha^T K = Y^T (G + \lambda I_l)^{-1} K \end{aligned}$$

بنابراین دو روش متمایز برای حل مسئله بهینه‌سازی رگرسیون ریبج از معادله (۹-۱) ارائه شد. در روش اول معادله (۱۱-۱) بردار وزن را محاسبه می‌کند و به عنوان جواب اصلی شناخته می‌شود. در حالی که معادله (۱۳-۱) جواب را به صورت ترکیب خطی از داده‌های آموزشی ارائه می‌دهد و به عنوان جواب دوگان شناخته می‌شود. پارامتر  $\alpha$ ، به عنوان متغیر دوگان معرفی شد.

نتیجه‌ی واضح از راه‌حل دوگان رابطه‌ی (۱۳-۱) آن است که مثال‌های آموزشی توسط حاصل ضرب داخلی بین جفت نقاط آموزشی در ماتریس  $G = \mathbf{X}\mathbf{X}^T$  مورد استفاده قرار می‌گیرد. ماتریس  $G$  به عنوان ماتریس گرام<sup>۱</sup> معرفی می‌شود. ماتریس گرام و ماتریس  $(G + \lambda I_l)$  با بعد  $l \times l$  می‌باشند. اگر بعد فضای ویژگی ( $n$ ) بزرگ‌تر از تعداد نمونه‌های آموزشی ( $l$ ) باشد، برای حل معادله (۱۳-۱) از معادله اصلی (۱۱-۱) که شامل ماتریس  $(\mathbf{X}^T \mathbf{X} + \lambda I_n)$  از بعد  $n \times n$  است، کارآمدتر است.

### ۳-۱ نگاشت غیرخطی

روش رگرسیون ریبج که در قسمت قبل مطرح شد، مسئله‌ی مشخص کردن یک رابطه‌ی خطی بین یک متغیر انتخابی و سایر ویژگی‌ها است، که فرض می‌شود چنین رابطه‌ای عملی باشد. با این حال در بسیاری از موارد، روابط مشاهده شده غیرخطی هستند لذا متغیر انتخاب شده فقط براساس تابعی غیرخطی از سایر ویژگی‌ها می‌تواند به دقت برآورد شود [۷].

استراتژی کلی این است سایر ویژگی‌ها از داده‌ها را به یک فضای ویژگی جدید تصویر می‌کنند به طوری که رابطه‌ی تعیین شده، بتواند به صورت خطی ارائه شود. بنابراین الگوریتم رگرسیون ریبج توصیف شده در قبل قادر به کشف و آشکارسازی آن‌ها خواهد بود.

<sup>۱</sup>Gram matrix

نگاشت زیر را در نظر بگیرید:

$$\begin{aligned} \phi : \mathbf{R}^n &\longrightarrow F \subseteq \mathbf{R}^N \\ x &\longmapsto \phi(x) \end{aligned}$$

انتخاب نگاشت  $\phi$  به منظور تبدیل یک رابطه‌ی غیرخطی به یک رابطه‌ی خطی است. این نگاشت انتظار ما از رابطه‌ی  $y = g(x)$  که باید تعیین شود را نشان می‌دهد. نتیجه نگاشت  $\phi$  تبدیل مجموعه داده‌های  $s$  به صورت  $\hat{S} = \{(\phi(x_1), y_1), \dots, (\phi(x_l), y_l)\}$  است. حال مانند قبل هدف یافتن رابطه‌ای به فرم زیر است:

$$f((x, y)) = |y - g(x)| = |y - \langle \beta, \phi(x) \rangle| = |\xi|$$

اگرچه روش اصلی می‌تواند به کار گرفته شود اما مشکلی که به وجود خواهد آمد این است: اگر  $N$  خیلی بزرگ باشد؛ پیدا کردن جواب یک سیستم  $N \times N$  از معادلات هزینه‌بر است. از طرفی دیگر اگر جواب دوگان در نظر گرفته شود؛ نشان داده شد که همه اطلاعات مورد نیاز الگوریتم، حاصل ضرب داخلی بین نقاط داده  $\langle \phi(x), \phi(z) \rangle$  در فضای ویژگی  $F$  است. به ویژه تابع پیش‌بینی  $g(x) = Y^T(G + \lambda I_l)^{-1}K$  شامل ماتریس گرام  $G = \mathbf{X}\mathbf{X}^T$  با درایه‌های

$$G_{ij} = \langle \phi(x_i), \phi(x_j) \rangle \quad (14-1)$$

است. که سطرهای  $\mathbf{X}$  بردارهای ویژگی  $\phi(x_1)^T, \dots, \phi(x_l)^T$  می‌باشد. بردار  $K$  شامل مقادیر

$$k_i = \langle \phi(x_i), \phi(x) \rangle \quad (15-1)$$

است.

گاهی اوقات حاصل ضرب داخلی می‌تواند به طور مؤثرتری به عنوان یک تابع مستقیم از ویژگی‌های ورودی، بدون محاسبه صریح نگاشت  $\phi$  محاسبه شود. تابعی که این محاسبات مستقیم را انجام می‌دهد به عنوان تابع کرنل معرفی می‌شود.

تعریف ۱-۳-۱ (تابع کرنل<sup>۱</sup>).  $\kappa$  یک تابع کرنل است که برای همه  $x, z \in X$  رابطه‌ی زیر برقرار باشد:

$$\kappa(x, z) = \langle \phi(x), \phi(z) \rangle$$

مثال ۱-۳-۲. فضای دو بعدی ورودی  $x \subseteq \mathbb{R}^2$  را با نگاشت ویژگی زیر در نظر می‌گیریم:

$$\phi : x \mapsto \phi(x) = (x_1^2, x_2^2, \sqrt{2}x_1x_2) \in F = \mathbb{R}^3.$$

فضای فرضی تابع خطی  $F$  به صورت زیر خواهد بود:

$$g(x) = \omega_{11}x_1^2 + \omega_{22}x_2^2 + \omega_{12}\sqrt{2}x_1x_2$$

نگاشت ویژگی، داده‌ها را از فضای دو بعدی به فضای سه بعدی به گونه‌ای نگاشت می‌کند که روابط خطی در فضای ویژگی متناظر با روابط درجه‌ی دو در فضای ورودی هستند. ترکیب نگاشت طرح با حاصل ضرب داخلی در فضای طرح می‌تواند به صورت زیر محاسبه شود:

$$\begin{aligned} \langle \phi(x), \phi(z) \rangle &= \langle (x_1^2, x_2^2, \sqrt{2}x_1x_2), (z_1^2, z_2^2, \sqrt{2}z_1z_2) \rangle \\ &= x_1^2z_1^2 + x_2^2z_2^2 + 2x_1x_2z_1z_2 \\ &= (x_1z_1 + x_2z_2)^2 = \langle x, z \rangle^2. \end{aligned}$$

بنابراین تابع

$$\kappa(x, z) = \langle x, z \rangle^2$$

یک تابع کرنل با فضای ویژگی  $F$  است. این به این معناست که می‌توان حاصل ضرب داخلی بین تصویر دو نقطه در فضای ویژگی را بدون محاسبه صریح مختصات آن‌ها، محاسبه کرد.

---

<sup>۱</sup>Kernel function

## ۴-۱ روش لاسو

حداقل انقباض مطلق و انتخاب عملگر (Lasso) <sup>۱</sup> که به اختصار از آن با نام لاسو یاد می‌کنیم، یک روش منظم‌سازی مشابه رگرسیون ریج است، با این تفاوت که جریمه  $L_2$  در رگرسیون ریج یعنی  $\sum_1^n \beta_j^2$  با مقدار جریمه  $L_1$  یعنی  $\sum_1^n |\beta_j|$  جایگزین شده است. فرض کنید ماتریس  $X$  که سطرهای آن بردارهایی با اندازه  $n$  هستند، داده شده باشد. هدف پیش‌بینی بردار پاسخ  $y$  با اندازه  $1 \times l$  با یک مدل خطی می‌باشد [۱۰]. روش لاسو مسئله بهینه‌سازی منظم زیر را حل می‌کند [۱۰]:

$$\operatorname{argmin}_{\beta} h(\beta) = \frac{1}{2} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1, \quad \lambda \geq 0 \quad (16-1)$$

همان‌گونه که ذکر شد  $X$  ماتریس  $l \times n$  و  $y$  بردار  $l \times 1$  می‌باشد که فرض می‌شود  $n > l$  باشد.  $\beta$  بردار  $1 \times n$  است که باید تعیین شود.

رگرسیون ریج و لاسو را نیز می‌توان به فرم مسئله بهینه‌سازی مقید زیر نیز تعریف نمود [۱۱]:

$$\hat{\beta}_{ridge} = \operatorname{argmin}_{\beta} \|y - X\beta\|_2^2 \quad \text{st.} \quad \|\beta\|_2^2 \leq t$$

و

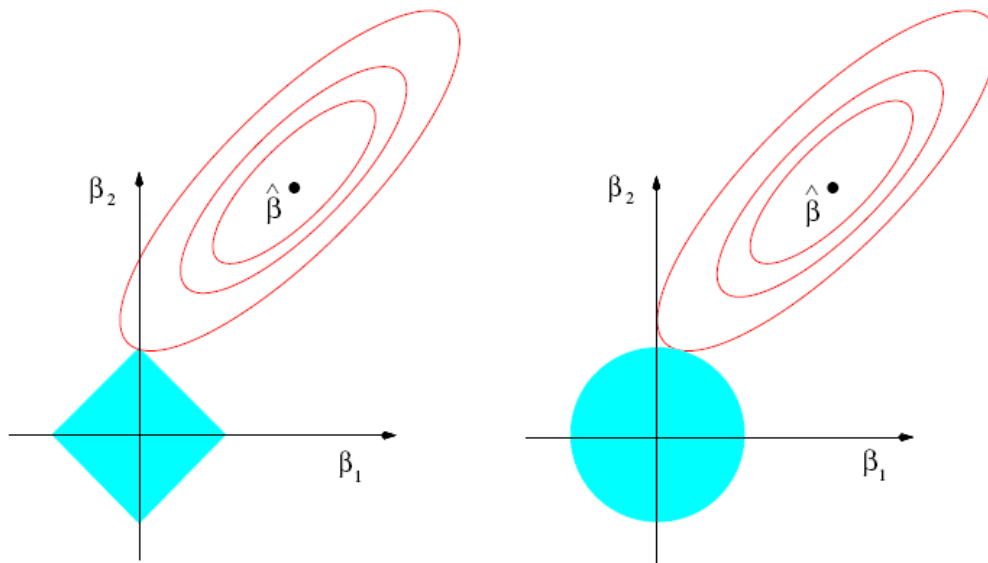
$$\hat{\beta}_{lasso} = \operatorname{argmin}_{\beta} \|y - X\beta\|_2^2 \quad \text{st.} \quad \|\beta\|_1 \leq t$$

که  $t$  پارامتر میزان‌ساز <sup>۲</sup> (مشابه  $\lambda$ ) است. برای هر  $\lambda$  و جواب متناظر با آن در شیوه نمایش قبل (فرم جریمه گذاری شده)، یک مقدار  $t$  وجود دارد به طوری که فرم مقید فوق نیز همان جواب را داشته باشد.

در شکل ۱-۲ تصویر سمت راست رگرسیون ریج و تصویر سمت چپ لاسو را نمایش می‌دهند [۱۱]. منحنی‌ها بیان‌گر خطا و ناحیه‌های خاکستری رنگ نمایان‌گر نواحی  $|\beta_1| + |\beta_2| \leq t$  و  $\beta_1^2 + \beta_2^2 \leq t$  که به ترتیب متناظر با قیود لاسو و رگرسیون ریج هستند، می‌باشد. همان‌طور که از شکل ۱-۲ نتیجه می‌شود، این جریمه باعث تنگ شدن <sup>۳</sup> مؤلفه‌های بردار جواب  $\beta$  و به طور خودکار منجر به انتخاب بعضی ویژگی‌ها می‌شود. به علاوه، جمله‌ی جریمه بردار جواب  $\beta$  را منظم می‌کند و بنابراین از بیش‌برازش جلوگیری می‌کند.

<sup>۱</sup>Least Absolute Shrinkage and Selection Operator (LASSO)      <sup>۲</sup>Tuning parameter      <sup>۳</sup>Sparsity





شکل ۱-۲: نمایش رگرسیون ریج و لاسو

### ۱-۴-۱ مقدمات لاسو

در این قسمت تعاریف و روابط بین مفاهیم تحدب، نیمه معین مثبت و مینیمم محلی و سراسری بیان می‌شوند، که برای فهم لاسو ضروری می‌باشد [۱۰].

تعریف ۱-۴-۱. تحدب<sup>۱</sup>: مجموعه‌ی  $D$  محدب<sup>۲</sup> است اگر

$$\forall x_1, x_2 \in D \quad \forall \alpha \in (0, 1) : \quad x = \alpha x_1 + (1 - \alpha)x_2 \in D$$

تابع  $f(x)$  محدب است اگر اولاً دامنه آن محدب باشد و ثانياً

$$f(x) = f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2)$$

باشد.

تعریف ۱-۴-۲. نیمه معین مثبت: ماتریس  $H$  با ابعاد  $n \times n$  را نیمه معین مثبت (PSD)<sup>۳</sup> گویند اگر برای هر بردار  $z$ ، با اندازه  $n \times 1$  رابطه‌ی  $z^T H z \geq 0$  برقرار باشد.

قضیه ۱-۴-۳ (نیمه معین مثبت بودن ماتریس هسیان تابع  $f(\cdot)$  محدب بودن آن را نتیجه می‌دهد). فرض کنید

---

<sup>۱</sup>Convexity    <sup>۲</sup>Convex    <sup>۳</sup>Positive Semi Definite (PSD)

$x$  یک بردار  $1 \times n$  و  $f(x)$  یک تابع اسکالر از  $n$  متغیر با مشتقات مرتبه دوم پیوسته با دامنه‌ی محدب  $D$  باشد. اگر ماتریس هسیان  $(\nabla^2 f(x))$  برای هر  $x \in D$  نیمه‌معین مثبت باشد، آنگاه  $f$  محدب است.

برهان. با توجه به قضیه تیلور<sup>۱</sup> برای هر  $x, x+h \in D$  و  $\theta \in (0, 1)$  می‌توان نوشت:

$$f(x+h) = f(x) + \nabla f(x)^T h + \frac{1}{2} h^T \nabla^2 f(x+\theta h) h \quad (17-1)$$

بنا به فرض  $\nabla^2 f(x+\theta h)$  نیمه‌معین مثبت است و لذا

$$h^T \nabla^2 f(x+\theta h) h \geq 0$$

بنابراین برای هر  $x, x+h \in D$  می‌توان نوشت:

$$f(x+h) \geq f(x) + \nabla f(x)^T h \quad (18-1)$$

$x+h = y$  قرار می‌دهیم و معادله بالا را به صورت زیر می‌نویسیم:

$$f(y) \geq f(x) + \nabla f(x)^T (y-x) \quad \forall x, y \in D \quad (19-1)$$

حال  $x_1$  و  $x_2$  را دو نقطه دلخواه در  $D$  و  $\alpha \in (0, 1)$  را یک اسکالر در نظر می‌گیریم، با توجه به محدب بودن  $D$ ،  $x = \alpha x_1 + (1-\alpha)x_2 \in D$  است. با توجه به (۱۹-۱) می‌توان نوشت:

$$f(x_1) \geq f(x) + \nabla f(x)^T (x_1 - x) \quad (20-1)$$

$$f(x_2) \geq f(x) + \nabla f(x)^T (x_2 - x) \quad (21-1)$$

با ضرب معادله‌ی (۲۰-۱) در  $\alpha$  و ضرب معادله‌ی (۲۱-۱) در  $(1-\alpha)$  و جمع آن‌ها نتیجه می‌شود:

$$\alpha f(x_1) + (1-\alpha)f(x_2) \geq f(x) + \nabla f(x)^T (\alpha x_1 + (1-\alpha)x_2 - x) = f(x) \quad (22-1)$$

---

<sup>۱</sup>Taylor's theorem

□ بنابراین  $f(x)$  محدب است.

قضیه ۱-۴-۴. اگر  $f(x)$  و  $g(x)$  دو تابع محدب با دامنه محدب  $D$  باشد، آن گاه  $r(x) = f(x) + g(x)$  نیز بر  $D$  محدب است.

برهان. فرض کنید  $x_1, x_2 \in D$  باشند، برای  $\alpha \in (0, 1)$  ،  $x = \alpha x_1 + (1 - \alpha)x_2$  قرار می دهیم. بنا به محدب بودن  $D$ ، نتیجه می شود  $x \in D$  است. حال

$$\begin{aligned} r(x) &= r(\alpha x_1 + (1 - \alpha)x_2) \\ &= f(\alpha x_1 + (1 - \alpha)x_2) + g(\alpha x_1 + (1 - \alpha)x_2) & (23-1) \\ &\leq \alpha f(x_1) + (1 - \alpha)f(x_2) + \alpha g(x_1) + (1 - \alpha)g(x_2) & (24-1) \\ &= \alpha r(x_1) + (1 - \alpha)r(x_2) & (25-1) \end{aligned}$$

□ بنابراین  $r(x)$  محدب است.

قضیه ۱-۴-۵. (تابع هدف لاسو محدب است). تابع هدف لاسو،  $h(\beta)$ ، در معادله (۱-۱۶) محدب است. برهان. یادآوری می شود، تابع هدف لاسو به صورت زیر تعریف می شود:

$$\operatorname{argmin}_{\beta} h(\beta) = \frac{1}{2} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1, \quad \lambda \geq 0.$$

تابع هدف لاسو را می توان به صورت زیر نوشت:

$$h(\beta) = f(\beta) + g(\beta) \quad (26-1)$$

که  $f(\beta) = \frac{1}{2} \|y - X\beta\|_2^2$  و  $g(\beta) = \lambda \|\beta\|_1$  می باشد.  $f(\beta)$  را نیز می توان به صورت زیر نمایش داد:

$$f(\beta) = \frac{1}{2} (y - X\beta)^T (y - X\beta)$$

بنابراین مشتق مرتبه اول  $f(\beta)$  به صورت  $\nabla f(\beta) = X^T y + X^T X\beta$  می باشد و هسیان تابع  $f(\beta)$ ، به صورت  $\nabla^2 f(\beta) = X^T X$  است. لذا برای هر بردار  $n \times 1$  مانند  $z$ ، خواهیم داشت:

$$z^T X^T X z = (Xz)^T Xz = \|Xz\|_2^2 \geq 0.$$

بنابراین  $\nabla^2 f(\beta)$  نیمه معین مثبت است. پس بنا به قضیه ۱-۴-۳،  $f(\beta)$  محدب است.

برای هر  $\beta_1, \beta_2$  و هر  $\alpha \in (0, 1)$  در نظر می‌گیریم، روابط زیر را خواهیم داشت:

$$g(\beta) = \lambda \|\alpha\beta_1 + (1 - \alpha)\beta_2\|_1 \quad (27-1)$$

$$\leq \lambda \|\alpha\beta_1\|_1 + \lambda \|(1 - \alpha)\beta_2\|_1 \quad (28-1)$$

$$= \lambda\alpha \|\beta_1\|_1 + \lambda(1 - \alpha) \|\beta_2\|_1 \quad (29-1)$$

$$= \alpha g(\beta_1) + (1 - \alpha)g(\beta_2) \quad (30-1)$$

بنابراین  $g(\beta)$  محدب است. چون  $f(\beta), g(\beta)$  هر دو محدب هستند. پس بنا به قضیه ۱-۴-۴،

$$\square \quad h(\beta) = f(\beta) + g(\beta) \text{ نیز محدب است.}$$

**قضیه ۱-۴-۶.** اگر  $f(x)$  یک تابع محدب بر مجموعه محدب  $D$  باشد آنگاه هر مینیمم محلی  $f$  بر  $D$ ، یک مینیمم سراسری  $f$  بر  $D$  است.

**برهان.** فرض کنید  $x^*$  یک مینیمم محلی باشد که مینیمم سراسری نیست، پس یک مینیمم سراسری مانند  $x_g^*$  وجود دارد به طوری که

$$f(x_g^*) < f(x^*) \quad (31-1)$$

به علاوه چون  $x^*$  یک مینیمم محلی است، لذا:

$$f(y) \geq f(x^*) \quad \forall y \in \text{nbhd}(x^*) \quad (32-1)$$

می‌باشد، که در این جا  $\text{nbhd}(x^*)$  همسایگی محلی  $x^*$  می‌باشد. بنا به محدب بودن  $f$  و  $D$  برای هر  $\alpha \in (0, 1)$  می‌توان نوشت:

$$f(\alpha x^* + (1 - \alpha)x_g^*) \leq \alpha f(x^*) + (1 - \alpha)f(x_g^*) \quad (33-1)$$

$$< \alpha f(x^*) + (1 - \alpha)f(x^*) \quad \text{بنا به معادله (31-1)} \quad (34-1)$$

$$= f(x^*) \quad (35-1)$$

برای مقدار  $\alpha$  به اندازه‌ی کافی کوچک به طوری که  $y = \alpha x^* + (1 - \alpha)x^* \in \text{nbhd}(x^*)$  باشد، نتیجه می‌شود:

$$f(y) < f(x^*) \quad y \in \text{nbhd}(x^*) \quad (36-1)$$

با مقایسه معادلات (32-1) و (36-1) به تناقض می‌رسیم. بنابراین بایستی  $f(x_g^*) \geq f(x^*)$  باشد. اما چون  $x_g^*$  یک مینیمم سراسری است پس باید  $f(x_g^*) \leq f(x^*)$  باشد. بنابراین  $f(x_g^*) = f(x^*)$  است. به عبارت دیگر مینیمم محلی  $x^*$ ، مینیمم سراسری نیز می‌باشد.  $\square$

ملاحظه 1-4-7. توجه داشته باشید که در قضیه 1-4-6، لزوماً برابر  $x_g^*$  نمی‌باشد. این کاملاً ممکن است که  $x_g^* \neq x^*$  باشد اما همزمان محدب بودن  $f$  روی دامنه‌ی  $D$  نتیجه دهد که  $f(x_g^*) = f(x^*)$  است.

## 5-1 الگوریتم شوتینگ برای لاسو

در این قسمت الگوریتمی به نام الگوریتم شوتینگ<sup>1</sup> برای حل مسئله لاسو بیان می‌کنیم. ابتدا نمونه‌ای از بهینه‌سازی یک متغیره را توضیح می‌دهیم، مثلاً هنگامی که  $n = 1$  باشد. سپس نشان خواهیم داد چگونه این نمونه ساده می‌تواند برای حالت‌های چند پارامتری با الگوریتم شوتینگ به کار گرفته شود.

### 1-5-1 نمونه یک متغیره: $n = 1$

مسئله بهینه‌سازی (16-1) به دلیل وجود جمله‌ی جریمه  $l_1$ ، ناهموار است. این مسئله را می‌توان با تعریف یک متغیر اسکالر  $t$  هموار نمود. گزاره بعد رابطه‌ای بین این دو مسئله بهینه‌سازی ایجاد می‌کند.

قضیه 1-5-1. فرض کنیم  $\beta \in \mathbb{R}$  یک اسکالر و  $x, y$  بردارهایی  $l \times 1$  باشند. مسئله بهینه‌سازی یک بعدی زیر مفروض است:

$$\operatorname{argmin}_{\beta} h(\beta) = \frac{1}{\gamma} \|y - x\beta\|_{\gamma}^{\gamma} + \lambda |\beta| \quad \lambda \geq 0 \quad (37-1)$$

<sup>1</sup>Shooting algorithm

فرض کنیم  $\beta_1^*$  جوابی برای معادله (۳۷-۱) باشد. مسئله بهینه‌سازی یک بعدی زیر را نیز در نظر گرفته:

$$\operatorname{argmin}_{\beta} \bar{h}(\beta) = \frac{1}{\gamma} \|y - x\beta\|_{\gamma}^2 + \lambda t \quad \lambda \geq 0 \quad (38-1)$$

$$t - \beta \geq 0 \quad (39-1)$$

$$t + \beta \geq 0 \quad (40-1)$$

فرض می‌کنیم  $(\beta^*, t^*)$  جوابی برای معادله (۳۸-۱) باشد. در این صورت  $\beta^* = \beta_1^*$  می‌باشد.

برهان. بنا به قضیه ۱-۴-۵ تابع هدف در معادله (۳۷-۱) محدب است. فرض کنیم  $(t_1, \beta_1)$  و  $(t_2, \beta_2)$  در محدودیت‌های (۳۸-۱) صدق کند. پس  $t_1 + \beta_1 \geq 0$  و  $t_1 - \beta_1 \geq 0$  و همچنین  $t_2 + \beta_2 \geq 0$  و  $t_2 - \beta_2 \geq 0$  می‌باشند.  $\alpha \in (0, 1)$  را در نظر گرفته و  $t = \alpha t_1 + (1 - \alpha)t_2$  و  $\beta = \alpha \beta_1 + (1 - \alpha)\beta_2$  قرار می‌دهیم. بنابراین

$$t - \beta = \alpha(t_1 - \beta_1) + (1 - \alpha)(t_2 - \beta_2) \geq 0$$

و به طور مشابه

$$t + \beta = \alpha(t_1 + \beta_1) + (1 - \alpha)(t_2 + \beta_2) \geq 0$$

می‌باشد. لذا  $(t, \beta)$  نیز در محدودیت‌ها صدق می‌کند، بنابراین محدودیت‌های تعریف شده یک مجموعه محدب است.

هسیان تابع هدف در (۳۸-۱) برابر

$$H(\beta, t) = \begin{bmatrix} (x^T x) & 0 \\ 0 & 0 \end{bmatrix} \quad (41-1)$$

می‌باشد. واضح است که  $H$  نیمه معین مثبت می‌باشد. بنابراین بنا به قضیه ۱-۴-۳ مسئله بهینه‌سازی (۳۸-۱) محدب است. لذا بنا به قضیه ۱-۴-۶ هر مینیمم محلی از (۳۷-۱) یا (۳۸-۱) نیز یک بهینه سراسری است. چون  $(\beta^*, t^*)$  یک جواب محلی (و بنابراین سراسری) از (۳۸-۱) است پس برای همه‌ی  $(\beta, t)$  هایی که  $t - \beta \geq 0$

و  $t + \beta \geq 0$  برقرار باشد، می توان نوشت:

$$\frac{1}{p} \|y - x\beta^*\|_p^q + \lambda t^* \leq \frac{1}{p} \|y - x\beta\|_p^q + \lambda t \quad (42-1)$$

به ویژه  $\beta = \beta_1^*$  و  $t = |\beta_1^*|$  در محدودیت های (38-1) صدق می کند. بنابراین خواهیم داشت:

$$\frac{1}{p} \|y - x\beta^*\|_p^q + \lambda t^* \leq \frac{1}{p} \|y - x\beta_1^*\|_p^q + \lambda |\beta_1^*| \quad (43-1)$$

چون  $\beta_1^*$  یک مینیمم سراسری از (37-1) است، پس می توان نوشت:

$$\frac{1}{p} \|y - x\beta_1^*\|_p^q + \lambda |\beta_1^*| \leq \frac{1}{p} \|y - x\beta^*\|_p^q + \lambda |\beta^*| \quad (44-1)$$

با جمع کردن (43-1) و (44-1) و ساده کردن آن نتیجه می شود:

$$t^* \leq |\beta^*| \quad (45-1)$$

اما  $t^* \geq \beta^*$  و  $t^* \geq -\beta^*$  صدق می کند، یعنی  $t^* \geq |\beta^*|$  می باشد. پس با توجه به (45-1) باید

$$t^* = |\beta^*| \quad (46-1)$$

باشد. با جای گذاری (46-1) در (43-1) نتیجه می شود:

$$\frac{1}{p} \|y - x\beta^*\|_p^q + \lambda |\beta^*| \leq \frac{1}{p} \|y - x\beta_1^*\|_p^q + \lambda |\beta_1^*| \quad (47-1)$$

با مقایسه (44-1) و (47-1) داریم:

$$\frac{1}{p} \|y - x\beta^*\|_p^q + \lambda |\beta^*| = \frac{1}{p} \|y - x\beta_1^*\|_p^q + \lambda |\beta_1^*| \quad (48-1)$$

با گسترش آن نتیجه می شود:

$$\frac{1}{p} y^T y + \frac{1}{p} (\beta^*)^T x^T x - y^T x \beta^* + \lambda |\beta^*| = \frac{1}{p} y^T y + \frac{1}{p} (\beta_1^*)^T x^T x - y^T x \beta_1^* + \lambda |\beta_1^*| \quad (49-1)$$

حالت اول :

$\lambda = 0$ : در این حالت  $\frac{1}{2} \|y - x\beta_1^*\|_2^2 + \lambda |\beta_1^*| = \frac{1}{2} \|y - x\beta_1^*\|_2^2$  که به ازای  $\beta_1^* = \beta^* = y^T x / x^T x$  مینیمم می شود. بنابراین در این حالت  $\beta^* = \beta_1^* = y^T x / x^T x$  می باشد.

حالت دوم :

$\lambda \neq 0, y^T x = 0$ : در این حالت

$\frac{1}{2} \|y - x\beta_1^*\|_2^2 + \lambda |\beta_1^*| = \frac{1}{2} y^T y + \frac{1}{2} (\beta_1^*)^2 x^T x + \lambda |\beta_1^*|$  مینیمم می شود.

به طور مشابه

$$\frac{1}{2} \|y - x\beta^*\|_2^2 + \lambda t^* = \frac{1}{2} \|y - x\beta^*\|_2^2 + \lambda |\beta^*| = \frac{1}{2} y^T y + \frac{1}{2} (\beta^*)^2 x^T x + \lambda |\beta^*|$$

می باشد، که به ازای  $\beta^* = \beta_1^* = 0$  مینیمم می شود. بنابراین در این حالت  $\beta^* = \beta_1^* = 0$  می باشد.

حالت سوم :

$\lambda \neq 0, y^T x \neq 0$ : معادله (۴۹-۱) برای همگی مقادیر  $\lambda, x, y$  برقرار است. با برابر قرار دادن جملات شامل  $\lambda$  باید

$$|\beta^*| = |\beta_1^*| \quad (۵۰-۱)$$

باشد. معادله (۵۰-۱) تضمین می کند که  $\frac{1}{2} (\beta_1^*)^2 x^T x = \frac{1}{2} (\beta^*)^2 x^T x$  است. با برابر قرار دادن ضرایب  $y^T x$  در هر دو طرف معادله (۴۹-۱) نتیجه می شود:

$$-y^T x \beta^* = -y^T x \beta_1^* \quad (۵۱-۱)$$

چون  $y^T x \neq 0$  است. بنابراین بایستی  $\beta^* = \beta_1^*$  باشد که با معادله (۵۰-۱) سازگار است.

بنابراین در همگی حالات همان طور که ادعا شد  $\beta^* = \beta_1^*$  می باشد.

□



در قضیه فوق نشان داده شد که هر دو مسئله بهینه‌سازی مطرح شده دارای جواب یکسان هستند. در قضیه بعد این جواب را به دست می‌آوریم.

قضیه ۱-۵-۲. مسئله بهینه‌سازی یک بعدی زیر مفروض است:

$$\operatorname{argmin}_{\beta} \bar{h}(\beta) = \frac{1}{\gamma} \|y - x\beta\|_{\gamma}^2 + \lambda t \quad \lambda \geq 0 \quad (52-1)$$

$$t - \beta \geq 0 \quad (53-1)$$

$$t + \beta \geq 0 \quad (54-1)$$

اگر  $x \neq 0$  و  $(\beta^*, t^*)$  جوابی برای معادله (۵۲-۱) باشد. آن‌گاه  $\beta^*$  به صورت زیر حاصل می‌شود:

$$\beta^* = \begin{cases} \frac{(y^T x - \lambda)}{x^T x} & ; \quad y^T x - \lambda > 0 \\ \frac{(y^T x + \lambda)}{x^T x} & ; \quad y^T x + \lambda < 0 \\ 0 & ; \quad -\lambda \leq y^T x \leq \lambda \end{cases} \quad (55-1)$$

برهان. فرم لاگرانژ برای مسئله بهینه‌سازی (۵۲-۱) به صورت زیر است:

$$L(\beta, t, \lambda_1, \lambda_2) = \frac{1}{\gamma} \|y - x\beta\|_{\gamma}^2 + \lambda t - \lambda_1(t - \beta) - \lambda_2(t + \beta) \quad (56-1)$$

شرایط بهینگی کروش‌کان‌تاکر (KKT) برای  $(\beta^*, t^*)$  به صورت زیر حاصل می‌شوند:

$$\frac{\partial L}{\partial \beta} = 0 \implies \beta x^T x = y^T x + \lambda_2 - \lambda_1 \quad (57-1)$$

$$\frac{\partial L}{\partial t} = 0 \implies \lambda_1 + \lambda_2 = \lambda$$

$$\left. \begin{array}{l} t - \beta \geq 0 \\ t + \beta \geq 0 \end{array} \right\} \text{محدودیت‌های نامساوی}$$

$$\left. \begin{array}{l} \lambda_1 \geq 0 \\ \lambda_2 \geq 0 \end{array} \right\} \text{مثبت بودن } \lambda_1, \lambda_2$$

$$\left. \begin{aligned} \lambda_1(t - \beta) &= 0 \\ \lambda_2(t + \beta) &= 0 \end{aligned} \right\} \text{محدودیت‌های مکمل}$$

اگر  $y^T x = 0$  باشد آن‌گاه همان‌گونه که در قضیه ۱-۵-۱ حالت اول و دوم نشان دادیم  $\beta^* = 0$  می‌باشد. بنابراین بدون از دست دادن کلیت فرض می‌شود  $y^T x \neq 0$  باشد.

حالت اول :

$$y^T x - \lambda > 0 \text{ باشد:}$$

از رابطه‌ی (۵۷-۱)،  $\beta x^T x = y^T x + \lambda_2 - \lambda_1$ ، و  $\lambda_1 = \lambda - \lambda_2$  می‌باشد. بنابراین

$$\beta x^T x = y^T x - \lambda + 2\lambda_2 \quad (58-1)$$

می‌باشد. چون  $\lambda_2 \geq 0$  و بنا به فرض حالت اول  $y^T x - \lambda > 0$  است. لذا

$$\beta x^T x = (y^T x - \lambda) + 2\lambda_2 > 0 \quad (59-1)$$

چون  $x \neq 0$  است، پس باید  $\beta > 0$  باشد. همچنین با جمع محدودیت‌های نامساوی در (۵۷-۱) نتیجه می‌شود  $t \geq 0$  است. بنابراین در حالت اول نتیجه می‌شود:  $(t + \beta) > 0$  است. بنابراین محدودیت‌های مکمل در (۵۷-۱) نتیجه می‌دهد  $\lambda_2 = 0$  است. لذا بنا به (۵۹-۱) داریم:

$$\beta = \frac{(y^T x - \lambda)}{x^T x} \quad (60-1)$$

حالت دوم :

$$y^T x + \lambda < 0 \text{ باشد:}$$

بنا به رابطه‌ی (۵۷-۱)  $\beta x^T x = y^T x + \lambda_2 - \lambda_1$  و  $\lambda_2 = \lambda - \lambda_1$  می‌باشد. بنابراین

$$\beta x^T x = y^T x + \lambda - 2\lambda_1 \quad (61-1)$$

می‌باشد. چون  $\lambda_1 \geq 0$  و بنا به فرض حالت دوم  $y^T x + \lambda < 0$  است، پس

$$\beta x^T x = (y^T x + \lambda) - 2\lambda_1 < 0 \quad (62-1)$$

می‌باشد. چون  $x \neq 0$  است. پس باید  $\beta < 0$  باشد. همچنین از محدودیت‌های نامساوی نتیجه می‌شود که  $0 \leq |t| \leq \beta$  است، پس در حالت دوم باید  $t > \beta$  باشد. بنابراین محدودیت‌های مکمل در (57-1) نتیجه می‌دهد  $\lambda_1 = 0$  است. لذا بنا به (62-1) داریم:

$$\beta = \frac{(y^T x + \lambda)}{x^T x} \quad (63-1)$$

می‌باشد.

حالت سوم :

$$\lambda \leq y^T x \leq -\lambda \text{ باشد:}$$

اگر  $\beta > 0$  باشد. پس  $t > \beta$  است و نتیجه می‌دهد:  $\lambda_2 = 0$  است (بنا به محدودیت‌های مکمل) و مطابق (60-1)  $\beta = \frac{(y^T x - \lambda)}{x^T x}$  می‌باشد. اما در حالت سوم  $y^T x - \lambda \leq 0$  است یعنی  $\beta \leq 0$  است که این تناقض است.

به طور مشابه اگر  $\beta < 0$  باشد، پس  $t > \beta$  است که نشان می‌دهد  $\lambda_1 = 0$  است (بنا به محدودیت مکمل) و مطابق (63-1)  $\beta = \frac{(y^T x + \lambda)}{x^T x}$  می‌باشد. بنا به فرض حالت سوم  $y^T x + \lambda \geq 0$  است که نتیجه می‌دهد:  $\beta \geq 0$  است و این تناقض است.

تنها راه برای جلوگیری از تناقض انتخاب  $\beta = 0$  می‌باشد که با استفاده از روابط (58-1) و (61-1) به انتخاب معتبر ضرایب لاگرانژ به فرم زیر منتهی می‌شود:

$$\lambda_1 = \frac{\lambda + y^T x}{2} \geq 0 \quad (64-1)$$

$$\lambda_2 = \frac{\lambda - y^T x}{2} \geq 0 \quad (65-1)$$

می‌توان بررسی کرد که  $\beta = 0$  و  $t = 0$  و  $\lambda_1, \lambda_2$  به دست آمده در (۶۴-۱)، (۶۵-۱) در شرایط بهینگی  
کروش کان تا کر در رابطه‌ی (۵۷-۱) صدق می‌کند. بنابراین در تمام حالت‌ها، مطابق آن چه ادعا کردیم  $\beta^*$   
توسط رابطه‌ی (۵۵-۱) تعیین می‌گردد.

□

## ۲-۵-۱ حالت چند متغیره: $n > 1$

در این قسمت روش بهینه‌سازی ارائه شده در [۱۲] که با عنوان 'الگوریتم شوتینگ' نیز از آن یاد می‌شود را بیان  
کرده و همگرایی این روش به مینیمم سراسری تابع هدف لاسورا نشان می‌دهیم.  
تابع هدف لاسو مجموع دو تابع محدب است که یکی از آن‌ها مشتق‌پذیر نیست. اما قسمت مشتق‌ناپذیر آن،  
تفکیک‌پذیر به مؤلفه‌های منفرد مثبتی بر مختصات<sup>۱</sup> می‌باشد.  
برای مسائل بهینه‌سازی با این ساختار، روش بهینه‌سازی مثبتی بر مختصات به مینیمم سراسری همگرا است  
[۱۲]. برای توضیحات بیش‌تر در خصوص روش مثبتی بر مختصات و همگرایی این روش به [۱۳] مراجعه شود.  
در الگوریتم (۱-۱)، الگوریتم شوتینگ برای حل مسئله بهینه‌سازی لاسورا مطرح می‌کنیم، و در قضیه ۱-۵-۳  
همگرایی این الگوریتم به مینیمم مثبتی بر مختصات تابع هدف لاسورا اثبات می‌نماییم [۱۴].

---

<sup>۱</sup>Coordinate wise

## الگوریتم ۱-۱ الگوریتم شویتینگ برای حل مسئله لاسو:

**ورودی:**  $X = [x_1, x_2, \dots, x_n]$ ،  $y$  و  $\lambda$  و  $X^{(-i)} = [x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n]$  و  $\beta^{(-i)} = [\beta_1, \dots, \beta_{i-1}, \beta_{i+1}, \dots, \beta_n]$  که به ترتیب از حذف  $i$  امین مؤلفه بردارهای  $X$  و  $\beta$  حاصل می‌شوند، و  $\beta_0$  (مقدار اولیه  $\beta$  که مثلاً از روش کمترین مربعات یا روش تصادفی حاصل می‌شود).  
**خروجی:**  $\beta$  که  $h(\beta)$  را کمینه کند:  $\lambda \geq 0$ ،  $\text{argmin}_{\beta} h(\beta) = \frac{1}{\sqrt{2}} \|y - X\beta\|_2 + \lambda \|\beta\|_1$   
 ۱:  $\beta = \beta_0$

۲: برای  $k = 0, 1, 2, \dots, m$  دو مرحله زیر را تکرار می‌کنیم:

$$1. f_k = h(\beta)$$

۲. برای  $i = 1, 2, \dots, n$  مراحل زیر را انجام می‌دهیم:

- مقدار فعلی  $\beta^{(-i)}$  را برای حل مسئله بهینه‌سازی یک بعدی زیر به کار می‌بریم.

$$(66-1) \quad \min_{\beta_i} h'(\beta) = \frac{1}{\sqrt{2}} \|y_i - x_i \beta_i\|_2 + \lambda |\beta_i| + \lambda \|\beta^{(-i)}\|_1, \quad \lambda \geq 0$$

که

$$(67-1) \quad y_i = y - X^{-i} \beta^{(-i)}$$

- فرض می‌کنیم  $\beta_i^*$  جواب (66-1) باشد، آن‌گاه  $i$  امین مؤلفه‌ی  $\beta$  را به‌روزرسانی کرده، برابر  $\beta_i^*$  قرار می‌دهیم. یعنی  $\beta_i = \beta_i^*$ .

**قضیه ۱-۵-۳.** الگوریتم ارائه شده در (۱-۱) را برای حل مسئله بهینه‌سازی لاسو در نظر بگیرید، آن‌گاه دنباله‌ای از تکرارهای  $f_1, f_2, \dots, f_m$  به مینیمم مبتنی بر مختصات تابع هدف لاسو همگرا می‌شود وقتی که  $m \rightarrow \infty$ .  
**برهان.** رابطه زیر برقرار است:

$$(68-1) \quad \frac{1}{\sqrt{2}} \|y - X\beta\|_2 + \lambda \|\beta\|_1 = \frac{1}{\sqrt{2}} \|y_i - x_i \beta_i\|_2 + \lambda |\beta_i| + \lambda \|\beta^{(-i)}\|_1$$

که  $y_i$  در (67-1) تعریف شده است. اگر  $\beta_i^*$  جواب مسئله بهینه‌سازی محدب (66-1) باشد، بنابراین خواهیم داشت:

$$(69-1)$$

$$\frac{1}{\sqrt{2}} \|y_i - x_i \beta_i^*\|_2 + \lambda |\beta_i^*| + \lambda \|\beta^{(-i)}\|_1 \leq \frac{1}{\sqrt{2}} \|y_i - x_i \beta_i\|_2 + \lambda |\beta_i| + \lambda \|\beta^{(-i)}\|_1 = h(\beta)$$

اگر  $\beta_{new}$  بردار جدید به دست آمده توسط به روز رسانی  $i$  امین مؤلفه  $\beta$  توسط جای‌گزینی  $\beta_i^*$  باشد، می‌توان (۶۹-۱) را به صورت زیر بازنویسی نمود:

$$h(\beta_{new}) \leq h(\beta) \quad (۷۰-۱)$$

بنابراین در هر تکرار در داخل حلقه‌ی  $(i = 1, 2, \dots, n)$  تابع هدف کاهش می‌یابد و این نتیجه می‌دهد که

$$f_{k+1} \leq f_k \quad \text{برای هر } k \quad (۷۱-۱)$$

به علاوه  $f_k$  از پایین به صفر محدود می‌شود. یعنی برای هر  $k$ ،  $f_k \geq 0$  می‌باشد. فرض کنید  $\hat{f}$  بزرگ‌ترین کران پایین دنباله  $\{f_k\}$  باشد. پس برای هر  $k$ ،  $\hat{f} \leq f_k$  است. با انتخاب هر  $\varepsilon > 0$  پس

$$f_k + \varepsilon > \hat{f} \quad (۷۲-۱)$$

هم‌چنین  $\hat{f} + \varepsilon$  بزرگ‌ترین کران پایین نمی‌باشد. بنابراین  $n$  وجود دارد به طوری که

$$f_{n_0} < \hat{f} + \varepsilon \quad (۷۳-۱)$$

چون  $k > n_0$  نتیجه می‌دهد  $f_k \leq f_{n_0}$ . پس می‌توان نتیجه گرفت که:

$$f_k \leq f_{n_0} < \hat{f} + \varepsilon \quad \text{اگر } k > n_0 \quad (۷۴-۱)$$

بنابراین برای هر  $k > n_0$  نتیجه می‌شود:

$$\hat{f} - \varepsilon < f_k < \hat{f} + \varepsilon \quad (۷۵-۱)$$

به عبارت دیگر دنباله  $\{f_k\}$  به  $\hat{f}$  همگرا است. اگر فرآیند فوق برای هر مؤلفه انجام شود  $\hat{f}$  مینیمم مبتنی بر مختصات  $h(\beta)$  خواهد بود.  $\square$

در قضیه فوق به همگرایی به سمت کمینه اشاره شد ولی مشخص نشد که کمینه نهایی، مینیمم سراسری است یا خیر. در قضیه بعد نشان داده می‌شود که کمینه به دست آمده، مینیمم سراسری  $h(\beta)$  است.

قضیه ۱-۵-۴. فرض کنید  $\hat{\beta}$  مینیمم مبتنی بر مختصات  $h(\beta)$  باشد، یعنی:

$$h(\hat{\beta} + \delta_i e_i) \geq h(\hat{\beta}) \quad \delta_i \neq 0 \quad (۷۶-۱)$$

که  $e_i$  برداری است که درایه‌ی  $i$  ام آن برابر یک و بقیه درایه‌های آن صفر می‌باشد. پس برای هر بردار  $P$  در همسایگی باز  $\hat{\beta}$  نتیجه می‌شود:

$$h(\hat{\beta} + P) \geq h(\hat{\beta}) \quad (۷۷-۱)$$

است، یعنی  $\hat{\beta}$  یک مینیمم محلی از  $h(\beta)$  است. بنا به محذب بودن  $h(\beta)$ ،  $\hat{\beta}$  یک مینیمم سراسری نیز می‌باشد.

برهان. یادآوری می‌شود که تابع هدف لاسورا می‌توان به صورت زیر نوشت:

$$h(\beta) = f(\beta) + g(\beta)$$

که

$$f(\beta) = \frac{1}{2} \|y - X\beta\|_2^2 \quad (۷۸-۱)$$

$$g(\beta) = \lambda \|\beta\|_1 = \lambda \sum_{i=1}^p |\beta_i| \quad (۷۹-۱)$$

می‌باشد، بنابراین روابط زیر را داریم:

$$\begin{aligned}
 h(\hat{\beta} + P) &= f(\hat{\beta} + P) + g(\hat{\beta} + P) & (۸۰-۱) \\
 &= f(\hat{\beta}) + P^T \nabla f(\hat{\beta}) + \frac{1}{\gamma} P^T X^T X P + \lambda \sum_{j=1}^p \left| \hat{\beta}_j + P_j \right| \\
 &= f(\hat{\beta}) + \lambda \sum_{j=1}^p \left| \hat{\beta}_j \right| + P^T \nabla f(\hat{\beta}) + \frac{1}{\gamma} P^T X^T X P \\
 &+ \lambda \sum_{j=1}^p \left| \hat{\beta}_j + P_j \right| - \lambda \sum_{j=1}^p \left| \hat{\beta}_j \right| \\
 &= f(\hat{\beta}) + g(\hat{\beta}) + P^T \nabla f(\hat{\beta}) + \frac{1}{\gamma} P^T X^T X P + \lambda \sum_{j=1}^p \left| \hat{\beta}_j + P_j \right| - \lambda \sum_{j=1}^p \left| \hat{\beta}_j \right| \\
 &= h(\hat{\beta}) + P^T \nabla f(\hat{\beta}) + \frac{1}{\gamma} P^T X^T X P + \lambda \sum_{j=1}^p \left| \hat{\beta}_j + P_j \right| - \lambda \sum_{j=1}^p \left| \hat{\beta}_j \right| & (۸۱-۱)
 \end{aligned}$$

فرض کنیم در (۸۰-۱)  $P = \delta_i e_i$  با  $\delta_i \neq 0$  باشد، لذا خواهیم داشت:

$$\begin{aligned}
 h(\hat{\beta} + \delta_i e_i) &= h(\hat{\beta}) + \delta_i e_i^T \nabla f(\hat{\beta}) + \frac{1}{\gamma} \delta_i^2 e_i^T X^T X e_i + \lambda \sum_{\substack{j=1 \\ j \neq i}}^p \left| \hat{\beta}_j \right| \\
 &+ \lambda \left| \hat{\beta}_i + \delta_i \right| - \lambda \sum_{\substack{j=1 \\ j \neq i}}^p \left| \hat{\beta}_j \right| - \lambda \left| \hat{\beta}_i \right| & (۸۲-۱) \\
 &= h(\hat{\beta}) + \delta_i e_i^T \nabla f(\hat{\beta}) + \frac{1}{\gamma} \delta_i^2 e_i^T X^T X e_i + \lambda \left| \hat{\beta}_i + \delta_i \right| - \lambda \left| \hat{\beta}_i \right| & (۸۳-۱)
 \end{aligned}$$

بنا به فرض  $h(\hat{\beta} + \delta_i e_i) \geq h(\hat{\beta})$  است. لذا

$$\delta_i e_i^T \nabla f(\hat{\beta}) + \frac{1}{\gamma} \delta_i^2 e_i^T X^T X e_i + \lambda \left| \hat{\beta}_i + \delta_i \right| - \lambda \left| \hat{\beta}_i \right| \geq 0 \quad (۸۴-۱)$$

است. رابطه فوق برای همگی  $\delta_i$  برقرار است. با انتخاب  $|\delta_i|$  به اندازه کافی کوچک می‌توان جمله  $\frac{1}{\gamma} \delta_i^2 e_i^T X^T X e_i$  را به اندازه دلخواه به صفر نزدیک کرد. بنابراین  $\theta_i > 0$  وجود دارد به طوری که برای همگی  $\delta_i \in (-\theta_i, \theta_i)$  رابطه‌ی زیر برقرار است.

$$\delta_i e_i^T \nabla f(\hat{\beta}) + \lambda \left| \hat{\beta}_i + \delta_i \right| - \lambda \left| \hat{\beta}_i \right| \geq 0 \quad (۸۵-۱)$$



اکنون قرار می‌دهیم:

$$P = \sum_{i=1}^n \delta_i e_i. \quad (۸۶-۱)$$

از معادله (۸۰-۱) نتیجه می‌شود:

$$h(\hat{\beta} + P) = h(\hat{\beta}) + \sum_{i=1}^n \delta_i e_i^T \nabla f(\hat{\beta}) + \frac{1}{\lambda} P^T X^T X P + \lambda \sum_{i=1}^n |\hat{\beta}_i + \delta_i| - \lambda \sum_{i=1}^n |\hat{\beta}_i| \quad (۸۷-۱)$$

با توجه به رابطه (۸۵-۱) خواهیم داشت:

$$\sum_{i=1}^n \delta_i e_i^T \nabla f(\hat{\beta}) + \lambda \sum_{i=1}^n |\hat{\beta}_i + \delta_i| - \lambda \sum_{i=1}^n |\hat{\beta}_i| \geq 0. \quad (۸۸-۱)$$

بنابراین از (۸۷-۱) و (۸۸-۱) نتیجه می‌شود:

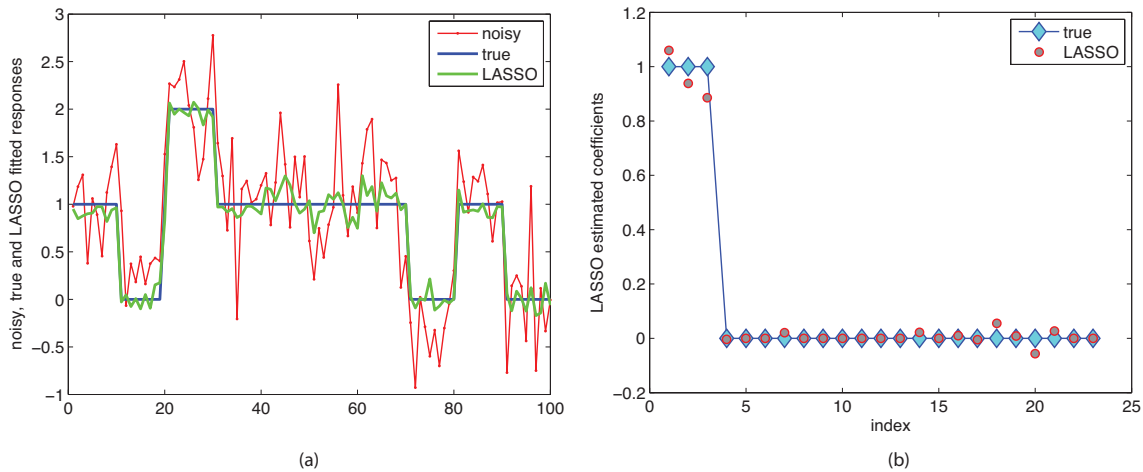
$$\begin{aligned} h(\hat{\beta} + P) &= h(\hat{\beta}) + \sum_{i=1}^n \delta_i e_i^T \nabla f(\hat{\beta}) + \frac{1}{\lambda} P^T X^T X P + \lambda \sum_{i=1}^n |\hat{\beta}_i + \delta_i| - \lambda \sum_{i=1}^n |\hat{\beta}_i| \\ &\geq h(\hat{\beta}) + \frac{1}{\lambda} P^T X^T X P \\ &\geq h(\hat{\beta}) \quad \text{بنا به نیمه معین مثبت بودن } X^T X \end{aligned}$$

به عبارت دیگر یک همسایگی باز با  $\delta_i \in (-\theta_i, \theta_i)$  که  $\theta_i \geq 0$  است پیدا شد به طوری که برای همه  $P$  به فرم (۸۶-۱)،  $h(\hat{\beta} + P) \geq h(\hat{\beta})$  است. این نتیجه می‌دهد که  $\hat{\beta}$  دقیقاً یک مینیمم کننده محلی و بنا به قضیه (۶-۴-۱) یک مینیمم کننده سراسری  $h(\beta)$  می‌باشد.  $\square$

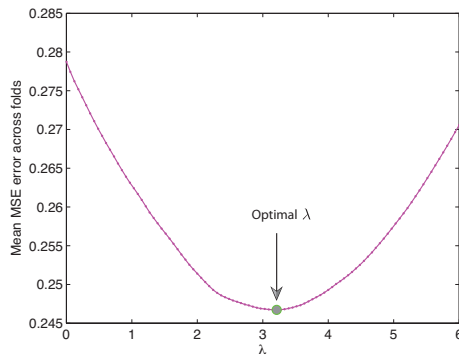
## ۶-۱ چگونگی انتخاب $\lambda$

پارامتر منظم ساز  $l_1$  برای لاسو را می‌توان با استفاده از اعتبارسنجی متقابل<sup>۲</sup> انتخاب کرد. به طور مختصر، داده‌های  $(X, y)$  را در نظر بگیرید، به این منظور سطرهای  $X, y$  را به  $k$  قسمت افراز<sup>۳</sup> می‌کنند که  $k$  جفت داده-پاسخ ایجاد می‌کند:  $(X_1, y_1), (X_2, y_2), \dots, (X_k, y_k)$ . فرض کنید  $(X^{(-i)}, y^{(-i)})$  زوج داده-پاسخی باشند که از حذف  $i$  امین قسمت  $(X_i, y_i)$  از  $(X, y)$  حاصل شده است، و  $\beta_{lasso}^{(-i)}$  جواب لاسو به دست آمده

<sup>۱</sup>Regularization parameter    <sup>۲</sup>Cross validation    <sup>۳</sup>Partition



(I)



(ب)

شکل ۱-۳: انتخاب  $\lambda$  با روش اعتبارسنجی متقابل

برای  $(X^{(-i)}, y^{(-i)})$  باشد. و  $n_i$  تعداد نقاط داده در  $i$  امین زوج داده-پاسخ  $(X_i, y_i)$  باشد. برای به دست آوردن  $\lambda$  مقدار متوسط میانگین مربعات خطای اعتبارسنجی متقابل<sup>۱</sup> به صورت زیر تعریف می شود:

$$\bar{C}V_{MSE}(\lambda) = \frac{1}{k} \sum_{i=1}^k \frac{1}{n_i} \left\| (y_i - X_i \beta_{lasso}^{(-i)}) \right\|_2^2 \quad (۸۹-۱)$$

با داشتن محدوده‌ی قابل قبول برای  $\lambda$ ، مقدار بهینه  $\lambda$  به صورت زیر محاسبه می شود.

$$\lambda^* = \arg \min_{\lambda} \bar{C}V_{MSE}(\lambda) \quad (۹۰-۱)$$

شکل ۱-۳ فرآیند انتخاب  $\lambda$  را برای یک مجموعه از داده‌ها نشان می دهد.

<sup>۱</sup>Average cross validated mean squared error

## فصل ۲

# برش هوشمند تصویر بر اساس روش یادگیری

همان‌گونه که پیش از این اشاره شد، استخراج پس‌زمینه<sup>۱</sup> و پیش‌زمینه<sup>۲</sup> تصویر یکی از مسائل مهم پردازش تصویر است. منظور از برش هوشمند تصویر تخمین دقیق پیش‌زمینه در تصویر و ویدیو است. در فصل ۱ روش‌های مختلف مواجهه با این مسئله ذکر شد، در این فصل به بیان یکی از مشهورترین روش‌ها در این حوزه با نام روش مبتنی بر یادگیری خواهیم پرداخت.

در شکل ۱-۲ (آ) یک عروسک را مشاهده می‌کنید، هدف استخراج دقیق عروسک است. تصویر ۱-۲ (ب) خروجی حاصل از استخراج این عروسک با نرم‌افزار Picture Cut Out را نشان می‌دهد که عدم کارایی آن کاملاً مشخص است. شکل ۱-۲ (ج) خروجی حاصل از روش برش هوشمند مبتنی بر یادگیری را نمایش می‌دهد.



(آ) تصویر اصلی (ب) خروجی نرم‌افزار PictureCutOut (ج) خروجی روش برش هوشمند

شکل ۱-۲: خروجی حاصل از نرم‌افزار Picture Cut Out و روش مبتنی بر یادگیری بر روی تصویر عروسک

مقاله [۱۵] و [۱] یک روش مبتنی بر یادگیری نیمه نظارتی ارائه شده است. در روش‌های نیمه نظارتی<sup>۳</sup> در حالتی که قرار است تصویر به دو زیر مجموعه افراز شود، کاربر قسمت‌هایی از هر دو زیر مجموعه را مشخص می‌کند و الگوریتم مربوطه تصویر را به دو زیر مجموعه افراز می‌کند.

<sup>۱</sup>Background <sup>۲</sup>Foreground <sup>۳</sup>Semi-supervised

به طور مثال اگر هدف جداسازی تصویر به دو قسمت شیء و زمینه باشد، کاربر از طریق ایجاد یک نقشه سه گانه<sup>۱</sup> (ماسک)، سیستم را آموزش می دهد. ماسک به نحوی ایجاد می شود که برای نواحی از تصویر که مطمئناً متعلق به شیء هستند، سفید، نواحی که متعلق به زمینه هستند، سیاه و نواحی دارای عدم قطعیت با خاکستری مشخص می شوند. نقاط متناظر با نواحی سیاه و سفید ماسک در تصویر اصلی به عنوان داده های آموزشی در نظر گرفته می شوند و هدف این روش تخمین برچسب سایر نقاط (پیکسل های متناظر با نقاط خاکستری) با استفاده از این داده های آموزشی است.

شکل ۲-۲: تصویر ورودی نمونه، ماسک ایجاد شده توسط کاربر و خروجی الگوریتم را نشان می دهد.



(ج) خروجی

(ب) ماسک

(ا) تصویر اصلی

شکل ۲-۲: اجرای فرآیند نیمه نظارتی برش هوشمند بر روی تصویر عروسک

عموما در مواجهه با این مسئله فرض می شود هر پیکسل ترکیب خطی از زمینه تصویر و شیء مورد نظر است. بنابراین فرض کنید تصویر  $I$  ترکیب خطی از پیش زمینه  $F$  و پس زمینه  $B$  با ضریب مشارکت  $\alpha$  باشد، یعنی:

$$I = \alpha F + (1 - \alpha)B \quad (1-2)$$

که  $\alpha$  در بازه صفر و یک است. اگر در یک پیکسل مشخص  $\alpha$  برابر یک یا صفر باشد پیکسل به صورت قطعی متعلق به شیء یا زمینه خواهد بود، در غیر این صورت ترکیبی از پیش زمینه و پس زمینه می باشد. این فرآیند معمولاً برای ترکیب به منظور خلق تصویر جدید به وسیله ترکیب خطی تصویر پیش زمینه استخراج شده و پس زمینه جدید با ضریب  $\alpha$  به دست آمده، به کار گرفته می شود. برش تصویر ابزاری مفید در ویرایش تصویر، فیلم و تولید تصویر متحرک و ... می باشد.

با در دست داشتن یک تصویر تمام مقادیر  $\alpha$ ،  $F$ ،  $B$  نامعلوم هستند و باید در هر پیکسل مقادیرشان مشخص شود. اطلاعات موجود بردار سه بعدی  $I = (I_r, I_g, I_b)$  در تصاویر رنگی  $RGB$  و مقادیر نامشخص بردارهای سه بعدی  $\alpha$ ،  $F$ ،  $B$  می باشند. لذا مسئله برش هوشمند ذاتاً یک مسئله فرومعین<sup>۲</sup> می باشد. زیرا برای سه سطح

<sup>۱</sup>Trimap      <sup>۲</sup> Under-constrained

تصاویر RGB داریم:

$$\begin{aligned} I_r &= \alpha F_r + (1 - \alpha) B_r \\ I_g &= \alpha F_g + (1 - \alpha) B_g \\ I_b &= \alpha F_b + (1 - \alpha) B_b \end{aligned} \quad (2-2)$$

بنابراین در هر پیکسل ۷ مقدار مجهول باید از ۳ مقدار معلوم به دست آورده شوند. پس در معادلات (۲-۲) تعداد مجهولات  $(\alpha, F, B)$  بیش‌تر از تعداد محدودیت‌ها است. همان‌گونه که در فصل پیش اشاره شد، مسائل بد طرح<sup>۱</sup>، عموماً با افزودن اطلاعات و محدودیت‌های بیش‌تر، مورد بررسی قرار می‌گیرند. در مسئله برش هوشمند تصویر اطلاعات اضافی، به وسیله نشان‌گذاری‌ها<sup>۲</sup> یا نقشه سه‌گانه یعنی برچسب‌گذاری بعضی پیکسل‌ها به عنوان قطعا پیش‌زمینه و برخی پیکسل‌ها به عنوان قطعا پس‌زمینه، عرضه می‌شوند. یک تصویر ساده حتی با مقادیر مشخص  $\alpha$  برای پیکسل‌های برچسب‌گذاری شده، هنوز هم بد طرح است، بنابراین چند محدودیت اضافی پیشنهاد می‌شود: برای مثال  $F, B$  با یک سری از نزدیکترین پیکسل‌های مشخص‌شده‌ی پیش‌زمینه و پس‌زمینه، می‌توانند به خوبی برآورد شوند.

رنگ پیش‌زمینه و پس‌زمینه به طور محلی یکنواخت در نظر گرفته می‌شوند. مقدار  $\alpha$ ، به طور محلی پیوسته در نظر گرفته می‌شود.

برای به‌کارگیری اطلاعات اضافی و محدودیت‌ها برای حل مسئله‌ی برش یک عنصر مهم در تولید نتایج دقیق، چگونگی ارتباط بین مقادیر آلفا و رنگ پیکسل‌های متناظرش می‌باشد. برای سادگی این ارتباط را رابطه آلفا-رنگ می‌نامیم.

مدل آلفا-رنگ در نظر گرفته شده، متعاقباً برای پیش‌بینی مقدار  $\alpha$  برای پیکسل‌های برچسب‌گذاری نشده، به کار گرفته می‌شود.

## ۱-۲ برآورد آلفا با یادگیری

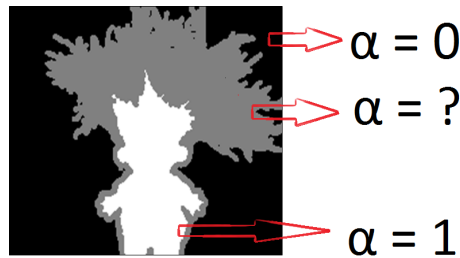
برش هوشمند شامل ۲ وظیفه مهم است: برآورد آلفا و محاسبه رنگ پیش‌زمینه و پس‌زمینه می‌باشد. فرض کنید تصویر  $I$  داده شده است، مجموعه همه پیکسل‌هایش را با  $\Omega = \{1, \dots, n\}$  نشان داده که  $n$  تعداد کل پیکسل‌ها است. مجموعه پیکسل‌های برچسب‌گذاری شده با  $\Omega_l \subset \Omega$  نمایش می‌دهیم، که مقدار  $\alpha$  برای آن‌ها معلوم

<sup>۱</sup>Ill-posed

<sup>۲</sup>Scribbles

است. برآورد آلفا به عنوان محاسبه مقدار  $\alpha$  برای مجموعه‌ای از پیکسل‌های نامعلوم (برچسب گذاری نشده)  $\Omega_u = \Omega - \Omega_l$  تعریف می‌شود، که در آن  $\Omega_l$  مجموعه پیکسل‌های برچسب گذاری شده، مرکب از ۲ زیرمجموعه است:

- برچسب گذاری شده به عنوان قطعا پیش‌زمینه که برای آن‌ها  $\alpha = 1$  است.
  - برچسب گذاری شده به عنوان قطعا پس‌زمینه که برای آن‌ها  $\alpha = 0$  است.
- که در تصویر ۲-۳،  $\Omega_l^B$  و  $\Omega_l^F$  به ترتیب با رنگ‌های سفید و سیاه نشان داده شده است.



شکل ۲-۳: نقشه سه‌گانه

هر پیکسل  $i \in \Omega$  به عنوان یک نقطه داده تعبیر می‌شود که با  $x_i \in R^d$  نمایش داده می‌شود.  $x_i$  می‌تواند همان  $I_i$  یا سایر ویژگی‌های به دست آمده برای پیکسل  $i$  باشد. در این بخش  $x_i = I_i$  در نظر گرفته می‌شود، که همان شدت رنگ خاکستری پیکسل یا یک بردار مرکب از مولفه رنگ برای تصاویر رنگی است.

مسئله یادگیری می‌تواند به صورت زیر فرمول‌بندی شود:

مجموعه نقاط  $\chi \subseteq R^d$  که  $\chi = \{x_i\}_{i \in \Omega}$  مفروض‌اند و  $\{\alpha_i\}_{i \in \Omega_l}$  مقادیر آلفا برای داده‌های برچسب گذاری شده  $\chi_l = \{x_i\}_{i \in \Omega_l}$  می‌باشد. هدف محاسبه دقیق مقدار آلفا  $\{\alpha_i\}_{i \in \Omega_u}$  برای داده‌های برچسب گذاری نشده،  $\Omega_u = \Omega - \Omega_l$ ، از روش یادگیری است.

## ۲-۲ برآورد آلفا با یادگیری محلی

روش برش هوشمند براساس یادگیری محلی<sup>۱</sup> یک مدل آلفا-رنگ را برای هر پیکسل در تصویر فقط براساس نزدیک‌ترین پیکسل‌ها که تصور می‌شود بیش‌ترین ارتباط را دارند، ارائه می‌دهد [۱۵].

<sup>۱</sup>Local learning

سپس برآورد آلفا با این فرض که مقدار آلفای هر پیکسل ترکیب خطی از مقدار آلفای پیکسل های مجاورش است، فرمول بندی می شود. اکنون تعیین ضرایب ترکیب خطی با فرآیند یادگیری محلی براساس مدل خطی آلفا-رنگ را شرح می دهیم.

## ۱-۲-۲ برآورد آلفا با به کارگیری رگرسیون ساده

در برآورد آلفا با یادگیری محلی، برای هر پیکسل  $i \in \Omega$ ، فرض می شود که مقدار  $\alpha_i$ ، می تواند به وسیله ترکیب خطی از مقادیر آلفای پیکسل های مجاورش،  $\{\alpha_i\}_{i \in N_i}$ ، پیش بینی شود. که  $N_i \subseteq \Omega$ ، پیکسل های مجاور پیکسل  $i$  ام می باشند و معمولاً پیکسل های موجود در یک پنجره محلی  $7 \times 7$ ، به مبدا  $i$  به عنوان همسایه های پیکسل  $i$  ام انتخاب می شود. سپس مقدار  $\alpha$  برای همه پیکسل ها از طریق کمینه سازی تابع هزینه درجه دوم، به طور هم زمان برآورد می شود.

فرض کنیم  $N_i = \{T_1, \dots, T_m\}$  مجموعه همسایگی پیکسل  $i$  و  $\alpha_i = [\alpha_{T_1}, \dots, \alpha_{T_j}, \dots, \alpha_{T_m}]$  بردار مقادیر آلفای  $N_i$  باشد که  $T_j \in N_i$  اگر  $f_i = [f_{iT_1}, \dots, f_{iT_j}, \dots, f_{iT_m}]$  برای نمایش بردار ضرایب ترکیب خطی در نظر گرفته شود، آن گاه ترکیب خطی ضرایب برای پیکسل  $i$ ، را می توان به صورت زیر نوشت:

$$\alpha_i = f_i^T \alpha_i \quad (۳-۲)$$

همچنین می توان  $\alpha_i$  در معادله (۳-۲) را به صورت ترکیب خطی از مقادیر آلفای تمام پیکسل ها بازنویسی کرد. اگر مقادیر آلفای همه پیکسل ها با بردار  $\alpha = [\alpha_1, \dots, \alpha_n]$  و ضرایب با  $\xi_i = [f_{i1}, \dots, f_{in}]$  نمایش داده شوند، آن گاه

$$\alpha_i = \xi_i^T \alpha \quad (۴-۲)$$

با تعریف ماتریس جدید  $F$  به صورت  $F = [\xi_1, \dots, \xi_n]_{\{\xi_i\}_{i \in \Omega}}$  معادله (۳-۲) را می توان به صورت

$$\alpha = F^T \alpha \quad (۵-۲)$$

بازنویسی نمود. اگر  $F$  معلوم باشد،  $\alpha$  از طریق کمینه‌سازی تابع درجه دوم زیر برآورد می‌شود:

$$\arg \min_{\alpha} h(\alpha) = \|\alpha - F^T \alpha\|^2 + c \|\alpha_l - \alpha_l^*\|^2 \quad (6-2)$$

که در آن  $\alpha_l^*$  برچسب از قبل مشخص برای داده‌های آموزشی است (نواحی سیاه و سفید در تصویر آموزشی) و  $\alpha_l$  تخمین همین نقاط در  $\alpha$  کمینه است.  $C$  را ماتریس قطری  $n \times n$  در نظر گرفته که  $j$  امین عنصر قطری مقدار ثابت  $c$  را اختیار می‌کند اگر  $j \in \Omega_l$  باشد و بقیه عناصر قطری صفر می‌باشند. اگر بردار  $\alpha^*$  به طول  $n$  در نظر گرفته شود، که  $j$  امین عنصر آن مقدار معلوم  $\alpha_j$  باشد؛ اگر  $j \in \Omega_l$ ، می‌توان معادله (6-2) را به صورت زیر فرمول‌بندی نمود:

$$\arg \min_{\alpha \in \mathbb{R}^n} h(\alpha) = \alpha^T (I_n - F)(I_n - F)^T \alpha + (\alpha - \alpha^*)^T C (\alpha - \alpha^*) \quad (7-2)$$

که  $I_n$  ماتریس همانی  $n \times n$  می‌باشد. با مشتق‌گیری نسبت به  $\alpha$  و برابر صفر قرار دادن آن داریم:

$$\begin{aligned} \frac{\partial h(\alpha)}{\partial \alpha} &= 0 \\ \frac{\partial h(\alpha)}{\partial \alpha} &= (I_n - F)(I_n - F)^T \alpha + (I_n - F)(I_n - F)^T \alpha + C(\alpha - \alpha^*) + C^T(\alpha - \alpha^*) \end{aligned}$$

ماتریس  $C$  یک ماتریس قطری است، لذا  $C = C^T$  می‌باشد. بنابراین

$$\begin{aligned} \frac{\partial h(\alpha)}{\partial \alpha} &= 2(I_n - F)(I_n - F)^T \alpha + 2C(\alpha - \alpha^*) = 0 \\ ((I_n - F)(I_n - F)^T + C)\alpha &= C\alpha^* \\ \Rightarrow \alpha &= ((I_n - F)(I_n - F)^T + C)^{-1} C\alpha^* \end{aligned} \quad (8-2)$$

اگر ضرایب ترکیب خطی در معادله (8-2) معلوم باشند. مقدار  $\alpha$  برآورد می‌شود.

حال برای تعیین ضرایب ترکیب خطی در معادله (8-2) به صورت زیر عمل می‌کنیم:

برای بردار معلوم  $x = [x^T \ 1]^T$ ، قرار می‌دهیم و سپس یک مدل خطی آلفا-رنگ به صورت زیر در نظر



می‌گیریم:

$$\alpha = x^T \beta + \beta_0 = \hat{x}^T \begin{bmatrix} \beta \\ \beta_0 \end{bmatrix} \quad (9-2)$$

که  $\beta = [\beta_1, \dots, \beta_d]$  و  $\beta_0$  ضرایب مدل هستند. با فرض  $X_i = [\hat{x}_{T1}, \dots, \hat{x}_{Tm}]$  که یک ماتریس  $m \times (d+1)$  است، می‌توان با روش رگرسیون ساده، برآوردی از  $\beta$  و  $\beta_0$  را توسط حل مسئله بهینه‌سازی درجه دوم زیر به دست آورد:

$$\arg \min_{\beta, \beta_0} \left\| \alpha_i - X_i \begin{bmatrix} \beta \\ \beta_0 \end{bmatrix} \right\|^2 \quad (10-2)$$

که با مشتق‌گیری از رابطه‌ی (10-2) و برابر صفر قرار دادن آن برآوردی از  $\beta$  و  $\beta_0$  به صورت زیر حاصل می‌شود:

$$\begin{bmatrix} \hat{\beta} \\ \hat{\beta}_0 \end{bmatrix} = (X_i^T X_i)^{-1} X_i^T \alpha_i \quad (11-2)$$

با جای‌گذاری معادله‌ی (11-2) در معادله‌ی (9-2) رابطه‌ی زیر حاصل می‌شود:

$$\alpha_i = \hat{x}_i^T (X_i^T X_i)^{-1} X_i^T \alpha_i \quad (12-2)$$

از طرفی بنا به معادله‌ی (3-2)،  $\alpha_i = f_i^T \alpha_i$  می‌باشد. لذا با مقایسه‌ی دو معادله‌ی (3-2) و (12-2)،  $f_i$  در معادله (3-2) به صورت زیر حاصل می‌شود، که مستقل از  $\{\alpha_j\}_{j \in N_i}$  بوده و فقط به  $\{X_j\}_{j \in N_i}$  وابسته می‌باشد.

$$f_i = X_i (X_i^T X_i)^{-1} \hat{x}_i \quad (13-2)$$

حال با معلوم شدن این ضرایب ترکیب خطی، آلفای پیکسل‌های تصویر به راحتی قابل محاسبه می‌باشد. با توجه به بدطرح بودن مسئله، ماتریس  $(X_i^T X_i)$  معکوس‌ناپذیر یا بدمقیاس است که موجب برآورد نادقیق آلفا می‌گردد.

در بخش بعد نتایج حاصل از پیاده‌سازی این روش را بر روی تصاویر مختلف نشان خواهیم داد.

## ۲-۲-۲ برآورد آلفا با به کارگیری رگرسیون ریج

در این قسمت نیز مشابه قسمت قبل عمل کرده با این تفاوت که این بار با به کارگیری روش رگرسیون ریج، برآوردی از  $\beta$  و  $\beta_0$  را توسط حل مسئله بهینه سازی درجه دوم زیر به دست می آوریم [۱۵]:

$$\arg \min_{\beta, \beta_0} \left\| \alpha_i - X_i \begin{bmatrix} \beta \\ \beta_0 \end{bmatrix} \right\|^2 + \lambda_r \begin{bmatrix} \beta \\ \beta_0 \end{bmatrix}^T \begin{bmatrix} \beta \\ \beta_0 \end{bmatrix} \quad (۱۴-۲)$$

جواب بهینه‌ی معادله (۱۴-۲) بنا به روابط (۱۱-۱) و (۱۳-۱) براساس نمایش اصلی و دوگان به صورت زیر حاصل می‌شوند:

$$\begin{bmatrix} \hat{\beta} \\ \hat{\beta}_0 \end{bmatrix} = (X_i^T X_i + \lambda_r I)^{-1} X_i^T \alpha_i \quad \text{جواب حاصل از نمایش اصلی} \quad (۱۵-۲)$$

$$= X_i^T (X_i X_i^T + \lambda_r I)^{-1} \alpha_i \quad \text{جواب حاصل از نمایش دوگان} \quad (۱۶-۲)$$

با جای‌گذاری معادله (۱۶-۲) در معادله (۹-۲)، رابطه‌ی زیر حاصل می‌شود:

$$\alpha_i = \hat{x}_i^T X_i^T (X_i X_i^T + \lambda_r I)^{-1} \alpha_i \quad (۱۷-۲)$$

از طرفی بنا به معادله‌ی (۳-۲)،  $\alpha_i = f_i^T \alpha_i$  می‌باشد. با مقایسه‌ی دو معادله (۳-۲) و (۱۷-۲)،  $f_i$  در معادله (۳-۲) به صورت زیر حاصل می‌شود که مستقل از  $\{\alpha_j\}_{j \in N_i}$  بوده و فقط به  $\{X_j\}_{j \in N_i}$  وابسته می‌باشد.

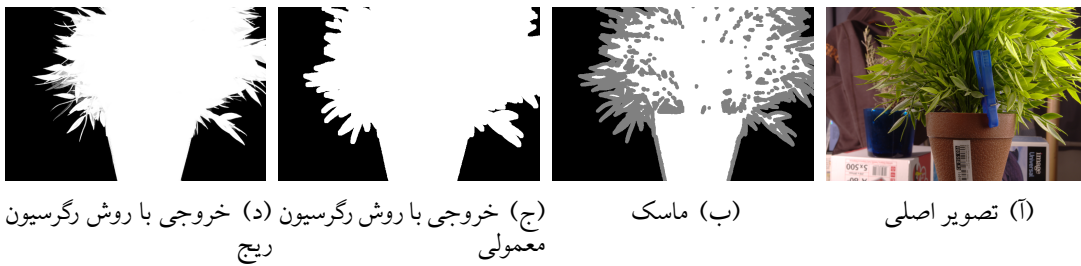
$$f_i = (X_i X_i^T + \lambda_r I_{(m)})^{-1} X_i \hat{x}_i \quad (۱۸-۲)$$

حال با معلوم شدن این ضرایب ترکیب خطی، آلفای پیکسل‌های تصویر به راحتی قابل محاسبه می‌باشد. شکل ۲-۵ و ۲-۴ نتایج حاصل از پیاده‌سازی این دو روش را با نرم‌افزار Matlab بر روی تصاویر ۲-۴ (آ) و ۲-۵ (ب) نشان می‌دهد.

با مقایسه خروجی‌های حاصل از اجرای این دو روش بر روی این تصاویر مشاهده می‌شود خروجی حاصل از روش رگرسیون ریج دقیق‌تر از خروجی حاصل از روش رگرسیون ساده می‌باشد.



شکل ۲-۴: خروجی حاصل از برش هوشمند تصویر براساس روش یادگیری با به کارگیری رگرسیون ساده و رگرسیون ریبج بر روی تصویر عروسک



شکل ۲-۵: خروجی حاصل از برش هوشمند تصویر براساس روش یادگیری با به کارگیری رگرسیون ساده و رگرسیون ریبج بر روی تصویر گیاه

### ۲-۲-۳ برآورد آلفا با به کارگیری تکنیک کرنل

مدل خطی آلفا-رنگ در معادله (۲-۹) می‌تواند با استفاده از تکنیک کرنل با جای‌گزین کردن  $x \in \mathbb{R}^d$  با بردار ترکیب  $\phi(x) \in \mathbb{R}^p$  که  $\phi$  یک نگاشت غیرخطی است، به صورت غیرخطی بسط داده می‌شود. معمولاً  $p > d$  مدل غیرخطی در فضای با بعد پایین‌تر می‌تواند با یک مدل خطی در فضای با بعد بالاتر تعریف شود. همان‌گونه که در زیر نشان داده شده است:

$$\alpha = \phi(x)^T \beta + \beta. \quad (۲-۱۹)$$

که برخلاف معادله (۲-۹)،  $\beta = [\beta_1, \dots, \beta_p]^T$  و  $\phi(x) = [\phi_1(x), \dots, \phi_p(x)]^T$  می‌باشد. با تکنیک کرنل فقط کافی است حاصل ضرب داخلی هر دو بردار معلوم  $x'_i$  و  $x'_j$  در معادله (۲-۱۸)، با مقدار تابع کرنل  $k(x'_i, x'_j)$  جایگزین شود. فرض کنید

$$\mathbf{k}_i = [k(x'_{T_1}, x'_i), \dots, k(x'_{T_m}, x'_i)]^T \quad (۲-۲۰)$$

$$\mathbf{K}_i = \begin{bmatrix} k(x'_{T1}, x'_{T1}) & \dots & k(x'_{T1}, x'_{Tm}) \\ \vdots & \ddots & \vdots \\ k(x'_{Tn}, x'_{T1}) & \dots & k(x'_{Tn}, x'_{Tm}) \end{bmatrix} \quad (21-2)$$

باشد، آنگاه  $f_i$  براساس مدل غیرخطی به صورت زیر حاصل می شود:

$$f_i = (\mathbf{K}_i + \lambda_r I_{(m)})^{-1} \mathbf{k}_i \quad (22-2)$$

که هم چنان فقط به  $x_i$  و  $\{x_j\}_{j \in N_i}$  وابسته است.

مقدار تابع کرنل می تواند به عنوان پیش پردازش برای برآورد آلفا محاسبه شود. معمولاً از کرنل گاوسین<sup>۱</sup> که به صورت زیر تعریف می شود:

$$k(x'_i, x'_j) = \exp\left(-\frac{1}{\vartheta} \|x'_i - x'_j\|^2\right)$$

استفاده می کنیم که در آن پارامتر  $\vartheta$  مقدار واریانس تصویر خاکستری تصویر داده شده، می باشد.

## ۳-۲ برآورد آلفا با یادگیری سراسری

یادگیری سراسری<sup>۲</sup> در برآورد آلفا، برآورد مقدار آلفای پیکسل های مجهول با یک مدل سراسری آلفا-رنگ آموزش یافته از برخی پیکسل های برچسب گذاری انتخابی، می باشد. پیکسل های انتخاب شده نزدیک تر به پیکسل های مجهول می تواند نقش مؤثرتری در برآورد آلفا داشته باشند. این کار به ویژه برای نمونه هایی که نقشه سه گانه آنها فراهم است و ناحیه مجهول کم عرض (نازک تر) است، مناسب است. برای هر پیکسل مجهول دو زیرمجموعه، انتخاب کرده و آنها را برای یادگیری مدل سراسری آلفا - رنگ به کار رفته در روش رگرسیون ریج وزن دهی شده، وزن دهی می کنند.

برای انتخاب زیرمجموعه هایی از پیکسل های برچسب گذاری شده، ابتدا دو زیرمجموعه  $Q_l^F \subseteq \Omega_l^F$  و  $Q_l^B \subseteq \Omega_l^B$  را انتخاب می کنند که برای هر پیکسل  $j$ ،  $D_j < D_{th}$  می باشد که  $D_j$ ، میانگین کوتاهترین فاصله اقلیدسی از  $j$ ، برای پیکسل های موجود در  $\Omega_u$ ، در شبکه های منظم است. و  $D_{th}$  فاصله آستانه ای است.

<sup>۱</sup>Gaussian kernel      <sup>۲</sup>Global learning



(آ) تصویر اصلی (ب) ماسک (ج) زیرمجموعه‌های انتخاب شده از هر ناحیه (د) آلفای حاصل



(ه) برشی از الفا

شکل ۲-۶: نمونه‌ای از برش هوشمند براساس روش یادگیری سراسری

برای هر پیکسل مجهول  $i$ ، دو زیرمجموعه  $Q_i^{F'} \subset Q_i^F$  و  $Q_i^{B'} \subset Q_i^B$  را انتخاب می‌کنند که پیکسل‌های موجود در آن کوتاهترین فاصله را نسبت به پیکسل  $i$  دارند، و هر دو مجموعه دارای تعداد پیکسل‌های یکسان باشند. (مثلاً ۸۰ پیکسل).

$D_{th}$  به صورت زیر تعیین می‌شود:  $D_{th} = \frac{(\gamma_d |\Omega_u|)}{|\Omega|} + \sqrt{2}$  که  $\gamma_d$  یک ثابت است و به طور تجربی  $\gamma_d = 1/2$  قرار داده می‌شود، و طوری طراحی شده که هنگامی که ناحیه مجهول ضخیم‌تر است، بزرگ‌تر است و بیش‌تر پیکسل‌های برچسب گذاری شده مجاور پیکسل مورد برآورد، می‌توانند انتخاب شوند. در غیر این صورت کوچک‌تر است و بیش‌تر پیکسل‌های برچسب گذاری شده با فاصله (دورتر) می‌توانند انتخاب شوند. یک مثال در شکل ۲-۶ نشان داده شده است.

برای هر پیکسل  $j$  در زیرمجموعه‌های  $Q_i^{F'}$  و  $Q_i^{B'}$ ، وزن به صورت  $W_j = \frac{1}{(D_j)^{\gamma_w}}$  تنظیم می‌شود که  $\gamma_w = 0.25$  یک ثابت است که به طور تجربی تعیین می‌شود. به علاوه یک ماتریس قطری  $W_{Q_i}$  با مقادیر وزن پیکسل‌های موجود در  $Q_i^{F'} \cup Q_i^{B'}$  ایجاد می‌کنند که یک ماتریس  $t \times t$  است، که  $t = |Q_i^{F'} \cup Q_i^{B'}|$  می‌باشد.

روش رگرسیون ریج وزن‌دهی شده [۱۶] برای آموزش مدل سراسری به کار گرفته می‌شود. مطابق با روش یادگیری محلی، ابتدا نتایج را برای مدل خطی آلفا - رنگ ارائه می‌دهند، سپس آن را برای نمونه غیرخطی بسط می‌دهند.

$\alpha_{Q_i}$  برای مشخص کردن بردار ترکیب مقادیر آلفای پیکسل‌های موجود در  $Q_i^{F'} \cup Q_i^{B'}$  و  $X_{Q_i}$  برای معرفی ماتریس ایجاد شده مشابه  $X_i$  در قسمت ۲-۲-۱ می‌باشد، اما با مقادیر معلوم پیکسل‌ها در  $Q_i^{F'} \cup Q_i^{B'}$ ، به جای

$N_i$  تعریف می‌شوند. با روش استقرای ریاضی برای یادگیری محلی، برای هر پیکسل  $i \in \Omega_u$  مقدار آلفا با مدل خطی به صورت زیر برآورد می‌شود:

$$\alpha_i = x_i' X_{Q_i}^T W_{Q_i} (W_{Q_i} X_{Q_i} X_{Q_i}^T W_{Q_i} + \lambda_r I(t)) \alpha_{Q_i} \quad (2-23)$$

با بسط دادن مدل خطی به مدل غیرخطی با استفاده از تکنیک کرنل مقدار آلفا به صورت زیر برآورد می‌شود:

$$\alpha_i = k_{Q_i}^T W_{Q_i} (W_{Q_i} \mathbf{K}_{Q_i} W_{Q_i} + \lambda_r I(t)) \alpha_{Q_i} \quad (2-24)$$

که  $k_{Q_i}$  مشابه  $\mathbf{k}_i$  در معادله (2-20) و  $\mathbf{K}_{Q_i}$  مشابه  $\mathbf{K}_i$  در معادله (2-21) اما با مقادیر معلوم پیکسل‌ها در  $Q_i^{F'} \cup Q_i^{b'}$  ایجاد می‌شوند. هم‌چنین از تابع کرنل مشابه تابع کرنل تعریف شده در یادگیری محلی استفاده می‌شود.

## فصل ۳

# استفاده از فرم بسته برای برش تصاویر

در این قسمت روش رسیدن به یک جواب فرم بسته برای به دست آوردن ضریب آلفا از رابطه‌ی (۱-۲) ارائه می‌شود [۱]. برای روشن شدن مطلب، ابتدا یافتن یک جواب فرم بسته مسئله تخمین آلفا برای تصاویر خاکستری را توضیح می‌دهیم. سپس این جواب‌ها برای تصاویر رنگی تعمیم خواهیم داد.

### ۱-۳ تصاویر خاکستری

یادآوری می‌شود هر پیکسل ترکیب خطی از زمینه تصویر و شیء مورد نظر است، یعنی

$$I = \alpha F + (1 - \alpha)B \quad (1-3)$$

که  $\alpha$  در بازه صفر و یک است.

همان‌گونه که ذکر شد برای حل مسئله برش برخی فرضیات معمول بر  $F$ ،  $B$ ،  $\alpha$  مورد نیاز است. برای رسیدن به جواب برای ناحیه مجهول فرض می‌شود  $F$ ،  $B$ ، هر دو روی پنجره کوچکی اطراف هر پیکسل تقریباً ثابت هستند.

این فرض که  $F$ ،  $B$  به طور موضعی هموار است به این معنی نیست که تصویر ورودی ( $I$ ) به طور موضعی هموار است زیرا ناپیوستگی در  $\alpha$  می‌تواند ناپیوستگی در  $I$  را نتیجه دهد. فرض این که  $F$ ،  $B$ ، هر دو روی پنجره کوچکی اطراف هر پیکسل تقریباً ثابت هستند، اجازه می‌دهد تا رابطه‌ی (۱-۳) را به صورت زیر بازنویسی کرده و

$\alpha$  به عنوان یک تابع خطی از تصویر  $I$  به صورت زیر در نظر گرفته شود.

$$\alpha_i = aI_i + b \quad \forall i \in w \quad (2-3)$$

که در آن  $a, b$  به صورت زیر حاصل می شود: بنا به رابطه ی (۱-۳) برای هر پیکسل  $i$  در پنجره ی  $w$  رابطه ی زیر برقرار است:

$$I_i = \alpha_i F_i + (1 - \alpha_i) B_i$$

$$\Rightarrow \alpha_i (F_i - B_i) = I_i - B_i$$

$$\Rightarrow \alpha_i = \frac{1}{F_i - B_i} I_i + \frac{-B_i}{F_i - B_i}$$

$$\text{لذا } a = \frac{1}{F_i - B_i}, \quad b = \frac{-B_i}{F_i - B_i} \text{ می باشد.}$$

هدف پیدا کردن  $\alpha, a, b$  و مینیمم سازی تابع هزینه زیر است.

$$J(\alpha, a, b) = \sum_{k \in I} \left( \sum_{i \in w_k} (\alpha_i - a_k I_i - b_k)^2 + \varepsilon a_k^2 \right) \quad (3-3)$$

که  $w_k$  پنجره کوچکی اطراف پیکسل  $k$  می باشد.

تابع هزینه بالا شامل یک عبارت منظم سازی بر  $a$  می باشد. یک دلیل افزودن این عبارت پایداری عددی<sup>۱</sup> است. برای مثال اگر تصویر در  $k$  امین پنجره ثابت باشد،  $a_k$  و  $b_k$  نمی تواند بدون داشتن یک دانش پیشین به صورت منحصر به فرد مشخص شوند. همچنین مینیمم کردن نرم  $a$  باعث پیدا شدن مقادیر آلفای هموارتر خواهد شد. ( زیرا  $a_k = 0$  به این معنا است که  $\alpha$  سراسر پنجره  $k$  ام ثابت است. )

در عمل به طور نمونه از پنجره های  $3 \times 3$  پیکسل استفاده می شود. چون اطراف هر پیکسل یک پنجره قرار داده می شود؛ پنجره های  $w_k$  در رابطه ی (۳-۳) هم پوشانی دارند. این ویژگی است که انتقال و توسعه اطلاعات بین دو پیکسل مجاور را ممکن می سازد. تابع هزینه بالا نسبت به  $b, a, \alpha$  درجه دوم است. برای تصویری با  $N$  پیکسل شامل  $3N$  مجهول می باشد.

خوشبختانه همانگونه که در قضیه بعد نشان می دهیم، ممکن است  $a, b$  در رابطه ی (۳-۳) نادیده گرفته شود و تابع هزینه فقط با  $N$  مجهول باقی بماند.

<sup>۱</sup>Numerical stability



قضیه ۱-۱-۳. فرض کنید  $J(\alpha)$  به صورت زیر تعریف شود:

$$J(\alpha) = \min_{a,b} J(\alpha, a, b)$$

آن‌گاه

$$J(\alpha) = \alpha^T L \alpha \quad (۴-۳)$$

می‌باشد. که  $L$  یک ماتریس  $N \times N$  است که  $(i, j)$  امین درایه‌ی آن به صورت زیر است:

$$\sum_{k|(i,j) \in w_k} \left( \delta_{ij} - \frac{1}{|w_k|} \left( 1 + \frac{1}{\frac{\varepsilon}{|w_k|} + \sigma_k^2} (I_i - \mu_k)(I_j - \mu_k) \right) \right) \quad (۵-۳)$$

که در آن  $\delta_{ij}$  دلتای کرونیگر<sup>۱</sup> و  $\mu_k, \sigma_k^2$  به ترتیب میانگین<sup>۲</sup> و واریانس<sup>۳</sup> شدت روشنایی<sup>۴</sup> پیکسل‌های پنجره‌ی  $w_k$  حول پیکسل  $k$  و  $|w_k|$  تعداد پیکسل‌های این پنجره است.

برهان. رابطه‌ی (۳-۳) را می‌توان به صورت زیر بازنویسی کرد:

$$\begin{aligned} J(\alpha, a, b) &= \sum_{k \in I} \left( \sum_{i \in w_k} (\alpha_i - a_k I_i - b_k)^2 + (\sqrt{\varepsilon} a_k + b_k)^2 \right) \\ &= \sum_{k \in I} \left\| \underbrace{\begin{bmatrix} -\alpha_i + a_k \\ b_k \end{bmatrix}}_{i \in w_k} + \begin{bmatrix} a_k \\ b_k \end{bmatrix} \begin{bmatrix} I_i & 1 \end{bmatrix} + \begin{bmatrix} a_k \\ b_k \end{bmatrix} \begin{bmatrix} \sqrt{\varepsilon} & 0 \end{bmatrix} \right\|^2 \end{aligned} \quad (۶-۳)$$

ماتریس  $G_k$  را یک ماتریس  $(|w_k| + 1) \times 2$  در نظر می‌گیریم که برای هر  $i \in w_k$  شامل یک سطر به صورت  $\begin{bmatrix} I_i & 1 \end{bmatrix}$  و سطر آخر آن به صورت  $\begin{bmatrix} \sqrt{\varepsilon} & 0 \end{bmatrix}$  است.

و  $\bar{\alpha}_k$  را برداری  $1 \times (|w_k| + 1)$  در نظر گرفته که مؤلفه‌هایش  $\alpha_i$  و آخرین مؤلفه آن صفر است.

---

<sup>۱</sup>Kronecker delta    <sup>۲</sup>Mean    <sup>۳</sup>Variance    <sup>۴</sup>Intensities

لذا رابطه‌ی (۳-۶) را می‌توان به صورت زیر نمایش داد:

$$J(\alpha, a, b) = \sum_{k \in I} \left\| G_k \begin{bmatrix} a_k \\ b_k \end{bmatrix} - \bar{\alpha}_k \right\|^2 \quad (۷-۳)$$

برای  $\alpha$  داده شده، فرض کنید  $(a_k^*, b_k^*)$  زوج بهینه داخل هر پنجره، یک جواب برای مسئله کم‌ترین مربعات<sup>۱</sup> باشد:

$$\begin{aligned} (a_k^*, b_k^*) &= \arg \min \left\| G_k \begin{bmatrix} a_k \\ b_k \end{bmatrix} - \bar{\alpha}_k \right\|^2 \\ &= (G_k^T G_k)^{-1} G_k^T \bar{\alpha}_k \end{aligned} \quad (۸-۳)$$

با جای‌گذاری رابطه‌ی (۳-۸) در رابطه‌ی (۳-۷) داریم:

$$\begin{aligned} J(\alpha) &= \min_{a,b} J(\alpha, a, b) \\ &= \sum_k J(\alpha, a_k^*, b_k^*) \\ &= \sum_{k \in I} \left\| G_k \begin{bmatrix} a_k^* \\ b_k^* \end{bmatrix} - \bar{\alpha}_k \right\|^2 \\ &= \sum_k \left\| G_k (G_k^T G_k)^{-1} G_k^T \bar{\alpha}_k - \bar{\alpha}_k \right\|^2 \\ &= \sum_k \left\| (I - G_k (G_k^T G_k)^{-1} G_k^T) (-\bar{\alpha}_k) \right\|^2 \end{aligned} \quad (۹-۳)$$

حال

$$\bar{G}_k = I - G_k (G_k^T G_k)^{-1} G_k^T \quad (۱۰-۳)$$

---

<sup>۱</sup>Least squares problem

قرار می دهیم و رابطه‌ی (۳-۹) به صورت زیر حاصل می شود:

$$J(\alpha) = \sum_k \|\bar{G}_k \bar{\alpha}_k\|^2$$

که آن را به صورت زیر بسط می دهیم:

$$\begin{aligned} J(\alpha) &= \sum_k (\bar{G}_k \bar{\alpha}_k)^T (\bar{G}_k \bar{\alpha}_k) \\ &= \sum_k \bar{\alpha}_k^T \bar{G}_k^T \bar{G}_k \bar{\alpha}_k \end{aligned}$$

ماتریس  $\bar{G}_k^T \bar{G}_k$  به صورت زیر حاصل می شود:

$$\begin{aligned} \bar{G}_k^T \bar{G}_k &= \left( I - G_k (G_k^T G_k)^{-1} G_k^T \right)^T \left( I - G_k (G_k^T G_k)^{-1} G_k^T \right) \\ &= \left( I - G_k (G_k^T G_k)^{-1} G_k^T \right) \left( I - G_k (G_k^T G_k)^{-1} G_k^T \right) \\ &= I - G_k (G_k^T G_k)^{-1} G_k^T - (G_k (G_k^T G_k)^{-1} G_k^T) + \left( G_k (G_k^T G_k)^{-1} G_k^T \right) G_k (G_k^T G_k)^{-1} G_k^T \\ &= I - 2(G_k (G_k^T G_k)^{-1} G_k^T) + (G_k (G_k^T G_k)^{-1} G_k^T) \\ &= \left( I - (G_k (G_k^T G_k)^{-1} G_k^T) \right) \\ &= \bar{G}_k \end{aligned} \tag{۳-۱۱}$$

لذا بنا به معادله (۳-۱۰) خواهیم داشت:

$$\bar{G}_k^T \bar{G}_k = \bar{G}_k = I - G_k (G_k^T G_k)^{-1} G_k^T \tag{۳-۱۲}$$

با بسط دادن معادله (۳-۱۲) و انجام عملیات جبری بیش تر و با توجه به این که  $\mu_k, \sigma_k^2$  به ترتیب میانگین و واریانس

شدت روشنایی پیکسل‌های موجود در پنجره‌ی  $w_k$  حول پیکسل  $k$  به صورت زیر تعریف می‌شوند [۱۷]:

$$\mu_k = \frac{\sum_{i \in w_k} I_i}{|w_k|} \quad (۱۳-۳)$$

$$\sigma_k^2 = \frac{\sum_{i \in w_k} I_i^2}{|w_k|} - \left( \frac{\sum_{i \in w_k} I_i}{|w_k|} \right)^2 \quad (۱۴-۳)$$

$(i, j)$  امین عنصر از  $\bar{G}_k^T \bar{G}_k$  به صورت زیر حاصل می‌شود:

$$\left( \delta_{ij} - \frac{1}{|w_k|} \right) \left( 1 + \frac{1}{\frac{\varepsilon}{|w_k|} + \sigma_k^2} (I_i - \mu_k)(I_j - \mu_k) \right) \quad (۱۵-۳)$$

با جمع بستن رابطه فوق روی هر پیکسل  $k$  حکم نتیجه می‌شود.

□

یک روش دیگر اثبات این قضیه در پیوست (آ-۱) آورده شده است.

## ۲-۳ تصاویر رنگی

یک روش ساده برای به کارگیری تابع هزینه برای تصاویر رنگی، به کارگیری تابع هزینه سطح خاکستری<sup>۱</sup> برای هر کانال به صورت جداگانه<sup>۲</sup> می‌باشد.

مجدداً می‌توان مدل خطی (۲-۳) را با مدل خطی زیر جایگزین کرد.

$$\alpha_i \approx \sum_c a^c I_i^c + b \quad \forall i \in w \quad (۱۶-۳)$$

که  $c$  نمایانگر کانال‌های رنگی می‌باشد. مزیت این مدل خطی ترکیب شده این است که فرض قبلی را که  $B, F$  روی هر پنجره ثابت باشند را تخفیف می‌دهد. در عوض همان‌طور که در پایین نشان داده می‌شود کافی است فرض شود که در هر پنجره کوچک هر کدام از  $B, F$  ترکیب خطی از دو رنگ می‌باشد. به عبارت دیگر مقدار  $F_i$  در هر پنجره کوچک روی یک خط واحد در فضای رنگ RGB قرار می‌گیرد:

$$F_i = \beta_i F_1 + (1 - \beta_i) F_2 \quad (۱۷-۳)$$

<sup>۱</sup>Gray level

<sup>۲</sup>Channel separately

به طور مشابه این فرض برای مقدار پس‌زمینه  $B_i$  نیز صادق می‌باشد.

**قضیه ۳-۲-۱.** اگر رنگ پیش‌زمینه و پس‌زمینه در هر پنجره در مدل خطی رنگ (۳-۱۷) صدق کند می‌توان نشان داد:

$$\alpha_i = \sum_c a^c I_i^c + b \quad \forall i \in w$$

$c$  نشان‌دهنده‌ی کانال‌های رنگ می‌باشد.

برهان. همان‌طور که گفته شد در تصاویر رنگی فرض می‌شود در هر پنجره کوچک رنگ هر کدام از  $F$  و  $B$  ترکیب خطی از دو رنگ باشد یعنی:

$$F_i^c = \beta_i^f F_\gamma^c + (1 - \beta_i^f) F_\gamma^c$$

$$B_i^c = \beta_i^b B_\gamma^c + (1 - \beta_i^b) B_\gamma^c$$

که  $B_\gamma, B_\gamma, F_\gamma, F_\gamma$  روی هر پنجره کوچک ثابت هستند.

با جای‌گذاری روابط بالا در معادله (۳-۱) روابط زیر حاصل می‌شوند:

$$\begin{aligned} I_i^c &= \alpha_i F_i^c + (1 - \alpha_i) B_i^c \\ &= \alpha_i \left( \beta_i^f F_\gamma^c + (1 - \beta_i^f) F_\gamma^c \right) + (1 - \alpha_i) \left( \beta_i^b B_\gamma^c + (1 - \beta_i^b) B_\gamma^c \right) \\ &= \alpha_i \beta_i^f F_\gamma^c + \alpha_i F_\gamma^c - \alpha_i \beta_i^f F_\gamma^c + \beta_i^b B_\gamma^c + B_\gamma^c - \beta_i^b B_\gamma^c - \alpha_i \beta_i^b B_\gamma^c - \alpha_i B_\gamma^c + \alpha_i \beta_i^b B_\gamma^c \\ &= \alpha_i (F_\gamma^c - B_\gamma^c) + \alpha_i \beta_i^f (F_\gamma^c - F_\gamma^c) + \beta_i^b (B_\gamma^c - B_\gamma^c - \alpha_i B_\gamma^c) + B_\gamma^c \\ &= \alpha_i (F_\gamma^c - B_\gamma^c) + \alpha_i \beta_i^f (F_\gamma^c - F_\gamma^c) + \beta_i^b (1 - \alpha_i) (B_\gamma^c - B_\gamma^c) + B_\gamma^c \end{aligned}$$

$$\Rightarrow I_i^c - B_\gamma^c = \begin{bmatrix} F_\gamma^c - B_\gamma^c & F_\gamma^c - F_\gamma^c & B_\gamma^c - B_\gamma^c \end{bmatrix} \begin{bmatrix} \alpha_i \\ \alpha_i \beta_i^f \\ \beta_i^b (1 - \alpha_i) \end{bmatrix} \quad (۳-۱۸)$$

$H$  را ماتریس  $3 \times 3$  در نظر می‌گیریم که  $c$  امین سطر آن به صورت زیر است:

$$\begin{bmatrix} F_{\gamma}^c - B_{\gamma}^c & F_{\gamma}^c - F_{\gamma}^c & B_{\gamma}^c - B_{\gamma}^c \end{bmatrix}$$

و  $I_i, B_{\gamma}$  بردارهایی  $3 \times 1$  می‌باشند. با استفاده از نمادگذاری بالا رابطه‌ی (۳-۱۸) به صورت زیر حاصل می‌شود:

$$H \begin{bmatrix} \alpha_i \\ \alpha_i \beta_i^f \\ \beta_i^b (1 - \alpha_i) \end{bmatrix} = I_i - B_{\gamma}$$

اگر عناصر سطر اول  $H^{-1}$  را با

$$\begin{bmatrix} a^1 & a^2 & a^3 \end{bmatrix}$$

و ضرب اسکالر سطر اول  $H^{-1}$  با بردار  $B_{\gamma}$  را با  $b$  مشخص کنیم، آنگاه

$$\alpha_i = \begin{bmatrix} a^1 & a^2 & a^3 \end{bmatrix} \begin{bmatrix} I_i^1 \\ I_i^2 \\ I_i^3 \end{bmatrix} - \begin{bmatrix} a^1 & a^2 & a^3 \end{bmatrix} \begin{bmatrix} B_{\gamma}^1 \\ B_{\gamma}^2 \\ B_{\gamma}^3 \end{bmatrix}$$

است. یعنی

$$\alpha_i = \sum_c a^c I_i^c + b$$

□

با استفاده از مدل خطی (۳-۱۶) تابع هزینه زیر برای برش تصاویر رنگی (RGB) تعریف می‌شود:

$$J(\alpha, a, b) = \sum_{k \in I} \left( \sum_{i \in w_k} (\alpha_i - \sum_c a_k^c I_i^c - b_k)^2 + \varepsilon \sum_c a_k^c \right) \quad (۳-۱۹)$$

مشابه نمونه‌ی خاکستری  $a^c, b$  می‌توانند از تابع هزینه حذف شوند. در این صورت تابع هزینه درجه دوم فقط

نسبت به  $\alpha$  مجهول بوده و به صورت زیر حاصل می‌شود:

$$J(\alpha) = \alpha^T L \alpha \quad (۳-۲۰)$$

که  $L$  یک ماتریس  $N \times N$  است که  $(i, j)$  درایه‌ی آن به صورت زیر می‌باشد:

$$\sum_{k|(i,j) \in w_k} \left( \delta_{ij} - \frac{1}{|w_k|} (1 + (I_i - \mu_k)(\Sigma_k + \frac{\varepsilon}{|w_k|} I_3)^{-1}(I_j - \mu_k)) \right) \quad (21-3)$$

که در آن  $\Sigma$  ماتریس کواریانس  $3 \times 3^1$  و  $\mu_k$  بردار  $1 \times 3$  نشان‌دهنده‌ی میانگین رنگ پیکسل‌های پنجره‌ی  $w_k$  است و  $I_3$  ماتریس همانی  $3 \times 3$  است.

ماتریس  $L$  در معادلات (3-4) و (3-20) با عنوان ماتریس لاپلاسیان<sup>۲</sup> معرفی می‌شوند. ماتریس لاپلاسیان برش تصویر ( $L$ ) را نیز می‌توان به صورت زیر

$$L = D - W \quad (22-3)$$

بازنویسی کرد، که  $D$  یک ماتریس قطری به صورت  $D(i, i) = \sum_j W(i, j)$  است و  $W$  یک ماتریس متقارن است که درایه‌های غیر قطری آن با استفاده از رابطه (3-21) حاصل می‌شود [۱].

در واقع ماتریس ( $L$ )، ماتریس لاپلاسیان گراف به کار رفته در روش قطعه بندی طیفی است که براساس تابع معرفی شده در معادله (3-21) محاسبه می‌شود.

شکل ۳-۱ نمونه‌ای ساده از یک گراف و ماتریس لاپلاسیان آن را نمایش می‌دهد.

جدول ۳-۱: گراف شش راسی و ماتریس لاپلاسیان آن

گراف برچسب گذاری شده	ماتریس درجه ( $D$ )	ماتریس مجاورت	ماتریس لاپلاسیان
	$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 2 & -1 & 0 & 0 & -1 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 \\ -1 & -1 & 0 & -1 & 3 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}$

ملاحظه ۳-۲-۲. ماتریس لاپلاسیان برش تصویر ( $L$ ) دارای ویژگی‌های زیر می‌باشد:

- ماتریس  $L$  متقارن است. (بنا به قضیه ۳-۱-۱ و اثباتش بدیهی است).
- ماتریس  $L$  نیمه معین مثبت است. (به این منظور کافی است نشان دهیم

$$\forall \alpha \quad J(\alpha) = \alpha^T L \alpha \geq 0$$

<sup>۱</sup>Covariance matrix

<sup>۲</sup>Laplacian matrix

در روند اثبات قضیه ۳-۱-۱ نشان دادیم  $J(\alpha) = \sum_k \|\bar{G}_k \bar{\alpha}_k\|^2 \geq 0$

لذا ماتریس لاپلاسین نیمه معین مثبت می باشد.

- مجموع هر سطر و هرستون از ماتریس  $L$  برابر صفر است. ( زیرا درجه هر راس با  $(-1)$  به ازای هر راس مجاور جمع زده می شود. )
- کوچکترین مقدار ویژه  $(\lambda_0)$  ماتریس لاپلاسین، صفر می باشد. یعنی  $\lambda_0 = 0$  است. ( زیرا برای هر بردار ثابت مانند  $V_0 = (1, 1, \dots, 1)$  رابطه  $Lv_0 = 0$  برقرار است. ) بنابراین فضای پوچ  $L$  شامل بردارهای ثابت است.

### ۳-۳ راهنمایی ها و محدودیت های کاربر

کاربر برخی نواحی زمینه و پیش زمینه را با خطوط رنگی روی تصویر مشخص می کند. کاربر از نشانه گذاری با نقاط مشکی برای نشان دادن پیکسل های پس زمینه ( $\alpha = 0$ ) و نشانه گذاری با نقاط سفید برای نشان دادن پیکسل های پیش زمینه ( $\alpha = 1$ ) استفاده می کند. برای به دست آوردن ضریب  $\alpha$  مطابق با این نشان گذاری ها سیستم زیر تعریف می شود [۱]:

$$\alpha = \operatorname{argmin} \alpha^T L \alpha$$

st

$$\alpha_i = \alpha_{\Omega_i} \quad \forall i \in \Omega_l \quad (23-3)$$

که  $\Omega_l$  مجموعه پیکسل های نشان گذاری شده توسط کاربر و  $\alpha_{\Omega_i}$  مقدار مشخص شده به وسیله نشان گذاری ها می باشد. در واقع به ازای هر  $i$ ،  $\alpha_{\Omega_i} = 0$  یا  $\alpha_{\Omega_i} = 1$  می باشد. با ترکیب تابع هدف با محدودیت های مشخص شده توسط کاربر نظیر نقشه سه گانه یا نشان گذاری ها  $\Omega_l$ ، کل تابع هزینه که باید مینیمم شود به صورت زیر حاصل می شود [۱۸]:

$$\operatorname{argmin} F(\alpha) = \alpha^T L \alpha + \lambda \|\alpha_{\Omega_l} - \alpha_{\Omega_l}^*\|_1^2 \quad (24-3)$$

که در آن  $\alpha_{\Omega_l}^*$  برچسب از قبل مشخص برای داده های آموزشی است (نواحی سیاه و سفید در تصویر آموزشی)



و  $\alpha_{\Omega_l}$  تخمین همین نقاط در  $\alpha$  کمینه است.  $D_{\Omega_l}$  را ماتریس قطری  $n \times n$  در نظر گرفته که  $j$  امین عنصر قطری مقدار ثابت ۱ را اختیار می‌کند اگر  $j \in \Omega_l$  باشد و بقیه عناصر قطری صفر می‌باشند. بردار  $\alpha^*$  به طول  $n$  در نظر گرفته می‌شود، که  $j$  امین عنصر آن مقدار معلوم  $\Omega_{lj}$  را اختیار می‌کند، اگر  $j \in \Omega_l$  باشد. بنابراین معادله (۲۴-۳) را می‌توان به صورت زیر فرمول‌بندی نمود:

$$\arg \min F(\alpha) = \alpha^T L \alpha + \lambda (\alpha - \alpha^*)^T D_{\Omega_l} (\alpha - \alpha^*) \quad (25-3)$$

$\lambda$  نشان دهنده میزان اهمیت و تاثیر اطلاعات تعیین شده توسط کاربر می‌باشد. معادله (۲۵-۳) یک مسئله بهینه‌سازی نامقید است که به آسانی قابل حل است و می‌توان با انتخاب مقدار بزرگ‌تر منظم‌ساز  $\lambda$  بر اطلاعات تعیین شده توسط کاربر، تاکید کرد.

برای کمینه‌سازی تابع  $F(\alpha)$  به صورت زیر عمل می‌شود:

از تابع  $F(\alpha)$  نسبت به پارامتر مجهول  $\alpha$  مشتق گرفته و برابر صفر قرار می‌دهیم:

$$\begin{aligned} \frac{\partial F(\alpha)}{\partial \alpha} &= 0 \\ \frac{\partial F(\alpha)}{\partial \alpha} &= L\alpha + L^T \alpha + \lambda D_{\Omega_l} (\alpha - \alpha^*) + D_{\Omega_l}^T \lambda (\alpha - \alpha^*) \end{aligned} \quad (26-3)$$

همان‌طور که قبلاً بیان شد، ماتریس لاپلاسیان  $L$  یک ماتریس متقارن است لذا  $L^T = L$  می‌باشد و ماتریس  $D_{\Omega_l}$  قطری است لذا  $D_{\Omega_l}^T = D_{\Omega_l}$  می‌باشد. بنابراین معادله ۲۶-۳ به صورت زیر حاصل می‌شود:

$$2L\alpha + 2\lambda D_{\Omega_l} (\alpha - \alpha^*) = 0$$

$$2(L + \lambda D_{\Omega_l})\alpha = 2\lambda D_{\Omega_l} \alpha^* \quad (27-3)$$

$$\alpha = \frac{\lambda D_{\Omega_l} \alpha^*}{L + \lambda D_{\Omega_l}} \quad (28-3)$$

معادله ۲۸-۳ آلفای به دست آمده براساس روش فرم بسته را ارائه می‌دهد. تصویر ۱-۳ نتایج حاصل از پیاده‌سازی این روش را با استفاده از نرم افزار Matlab بر روی تصاویر مختلف، نشان می‌دهد.

تصویر ۱-۳ (آ) یک مثال ترکیبی است که با کامپیوتر شبیه‌سازی شده است و یک مه غلیظ روی یک پس‌زمینه ساده مرکب از چند نوار رنگی را نشان می‌دهد که در مدل خطی رنگ صدق می‌کند. نقاط سیاه و سفید محدودیت‌های

ورودی را نشان می‌دهد. تصویر ۱-۳ (ب) خروجی حاصل از این روش را نشان می‌دهد.

تصویر ۱-۳ (ج) یک تصویر واقعی با رنگ‌های پیش‌زمینه و پس‌زمینه نسبتاً یکسان می‌باشد. تصویر ۱-۳ (د)

نیز خروجی این روش را نشان می‌دهد. تنها با مشخص کردن دو نقطه سیاه و سفید یک برش با کیفیت بالا به دست آمده است.



(ب) خروجی



(آ) تصویر اصلی



(د) خروجی



(ج) تصویر اصلی

شکل ۳-۱: خروجی حاصل از پیاده‌سازی روش فرم بسته

**قضیه ۳-۳-۱.** فرض کنید  $I$  تصویری با پیش‌زمینه  $F$  و پس‌زمینه  $B$  مطابق معادله (۳-۱) باشد و  $\alpha^*$  مقدار درست ضریب  $\alpha$  باشد. اگر  $F$  و  $B$  در هر پنجره‌ی محلی  $w_k$  در مدل خطی رنگ صدق کند و محدودیت‌های تعیین شده توسط کاربر با  $\alpha^*$  هماهنگ باشد، آن‌گاه  $\alpha^*$  یک جواب بهینه سیستم (۳-۲۳) است که  $L$  با  $\varepsilon = 0$  ایجاد شده است.

برهان. از آن‌جا که  $\varepsilon = 0$  است، اگر هر پنجره محلی در مدل خطی رنگ صدق کند، بنا به رابطه‌ی (۳-۱۹)

$$J(\alpha^*, a, b) = 0$$

است و بنابراین

$$J(\alpha) = \min_{a,b} J(\alpha^*, a, b) = \alpha^{*T} L\alpha^* = 0$$

□

می باشد.

### ۴-۳ برش تصویر با به کارگیری لاسو

از روش بهینه‌سازی لاسو نیز در مسئله برش هوشمند تصویر استفاده شده است و نتایج خوبی داشته است [۱۹]، [۲۰]. چارچوب کلی کار در این روش‌ها نیز شامل حل یک مسئله بهینه‌سازی است که مقید به یک نرم یک است. از جمله روش‌های حل چنین مسائلی می‌توان به الگوریتم‌های LASSO و Nesterov's اشاره نمود که به ترتیب در [۱۹] و [۲۰] مورد استفاده قرار گرفته‌اند. که در این پایان‌نامه فقط روش لاسو را مورد بررسی قرار دادیم. شکل ۲-۳ مراحل کار این روش را روی تصویر عروسک نمایش می‌دهد. ۲-۳ (آ) تصویر ورودی و ۳-۳ (ب) نقشه سه گانه متناظر با آن را نمایش می‌دهد. در این روش ابتدا با استفاده از یک الگوریتم یک مجموعه از ناحیه معلوم به عنوان نمونه‌ای از پیکسل‌های معلوم پیش‌زمینه و پس‌زمینه اطراف ناحیه مجهول انتخاب می‌شود (شکل ۲-۳ ج)، سپس با استفاده از یک تابع احتمال، زیرمجموعه‌ای از مجموعه‌ی ایجاد شده، که بیش‌ترین ارتباط را با پیکسل مجهول موردنظر دارد، انتخاب می‌شود (شکل ۲-۳ د). در مرحله بعد با استفاده از مسئله بهینه‌سازی زیر ضریب بتا مشخص می‌شود [۱۹].

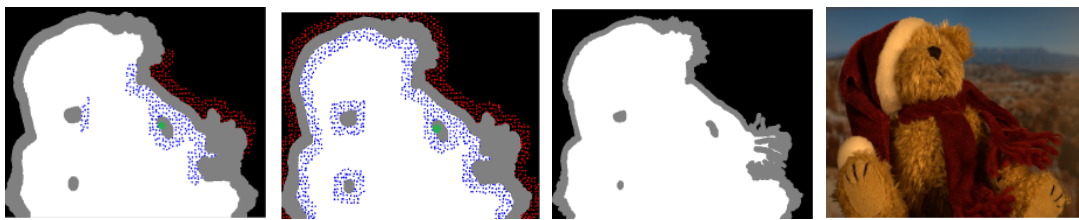
$$\beta = \arg \min \|\mathbf{v}_i - D\beta_i\|_2^2 \quad \text{st} \quad \|\beta_i\|_1 \leq 1; \quad \beta_i \geq 0 \quad (29-3)$$

در پایان با استفاده از رابطه‌ی (۳-۳۰) ضریب آلفا محاسبه می‌شود (شکل ۲-۳ د).

$$\alpha = \sum_{p \in F} \beta^{(p)} \quad (30-3)$$

شکل ۳-۳ خرجی حاصل از روش لاسو و روش فرم بسته را نمایش می‌دهد. برای مقایسه بهتر قسمت‌هایی از تصویر اصلی را با کادر نمایش داده‌ایم، همان‌طور که ملاحظه می‌شود این نقاط در خروجی حاصل از روش لاسو بهتر نمایان شده است.

در [۲۱] نیز از لاسو در برش هوشمند تصویر استفاده شده است. همان‌طور که در بخش ۱-۳ ذکر شد در

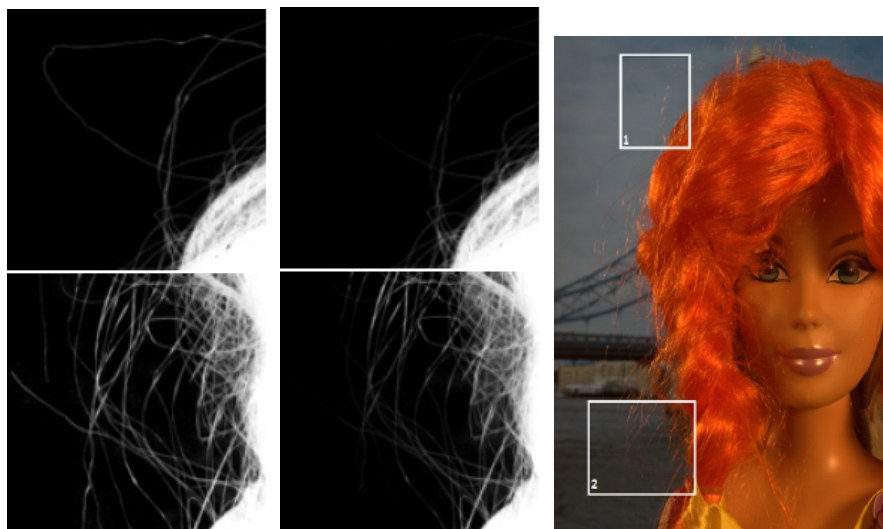


(آ) تصویر اصلی (ب) نقشه سه گانه (ج) مجموعه‌ی انتخاب شده از (د) زیرمجموعه انتخابی نواحی معلوم در نقشه سه گانه براساس تابع احتمال مطرح شده در [۱۹]



(ه)

شکل ۳-۲: خروجی حاصل از برش هوشمند تصویر با به کارگیری لاسو بر روی تصویر عروسک



(آ) تصویر اصلی (ب) خروجی فرم بسته (ج) خروجی روش لاسو

شکل ۳-۳: مقایسه خروجی حاصل از برش هوشمند تصویر با به کارگیری لاسو و روش فرم بسته

روش فرم بسته تابع هدف اصلی به صورت زیر است:

$$J(\alpha, a, b) = \sum_{i \in w_j} (\alpha_i - a_j I_i - b_j)^2 \quad (31-3)$$

در [۲۱] تابع هدف به صورت زیر تعریف می شود:

$$J(\alpha, a, b) = \sum_{i \in w_j} |\alpha_i - a_j I_i - b_j| \quad (32-3)$$

همان طور که ملاحظه می شود در این فرمول بندی از لاسو استفاده شده است که مبانی آن را در فصل ۱ ذکر کردیم.

## فصل ۴

### روش پیشنهادی

#### ۱-۴ آشکارسازی بهتر شریان‌های کرونری قلب با یادگیری نیمه‌نظارتی خودکار

تخمین درست میزان گرفتگی شریان‌های کرونری اهمیت به‌سزایی در تشخیص و درمان بیماری‌های قلبی دارد. تصاویر آنژیوگرافی مرتبط با این شریان‌ها عموماً به نحوی هستند که در برخی نواحی رگ‌ها به خوبی مشهود نیستند. در این فصل با استفاده از روشی که در فصل ۲ بیان شد، آشکارسازی بهتر شریان‌ها انجام شده است. به این منظور ابتدا با یک روش خودکار قسمت‌هایی از تصویر که به احتمال زیاد متعلق به شریان‌های قلب هستند و همچنین آن نواحی که به احتمال زیاد متعلق به شریان‌های قلب نیستند، شناسایی می‌شوند. سپس تعیین تعلق نواحی باقیمانده تصویر به هر یک از دو دسته شریان یا غیر آن با استفاده از نواحی شناسایی شده مرحله قبل، به عنوان داده‌های آموزشی، به عهده‌ی تابع هدف (۲-۶) می‌باشد. نتایج آزمایشات انجام شده کارایی روش مورد استفاده را نشان داده است.

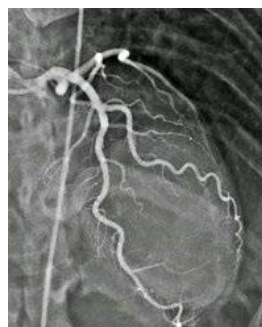
شکل ۴-۱ یک نمونه از تصاویر آنژیوگرافی و خروجی این روش را نمایش می‌دهد.

#### ۲-۴ ایجاد نقشه‌ی سه‌گانه خودکار

هدف اصلی نقشه سه‌گانه مشخص کردن  $\alpha^*$  در رابطه (۲-۶) است. هر چه تخمین نقاط آموزشی دقیق‌تر باشد، آلفای حاصل از کمینه‌سازی رابطه (۲-۶) دقیق‌تر خواهد بود. در روش‌های معمول مورد استفاده در برش



شکل ۴-۱ (ب) خروجی روش برش هوشمند



شکل ۴-۱ (آ) تصویر اصلی

شکل ۴-۱: نمونه‌ای از تصاویر آنژیوگرافی و نمایش خروجی حاصل از روش برش هوشمند با استفاده از نقشه سه‌گانه پیشنهادی

هوشمند، نقشه سه‌گانه به صورت دستی ایجاد می‌شود. از آنجا که ایجاد دستی چنین ماسکی برای تصاویر شریان‌های کرونری کار سختی می‌باشد، در این بخش شیوه خودکاری برای تخمین آن بیان شده و کارایی آن در شناسایی شریان‌های کرونری و بهبود دادن تصویر اصلی نشان داده شده است.

مراحل زیر را برای تولید نقشه سه‌گانه خودکار پیشنهاد می‌کنیم:

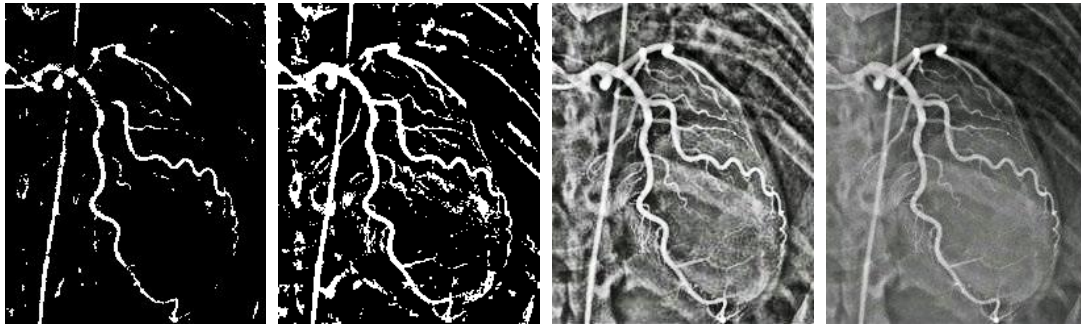
- ابتدا برای ایجاد بهبود نسبی اولیه روی تصویر، یک روش افزایش کنتراست روی تصویر اعمال می‌کنیم تا تقطیع تصویر بهتری داشته باشیم.

- برای مشخص کردن نواحی از تصویر اصلی که متعلق به شریان‌های کرونری هستند، با توجه به این‌که شریان‌های کرونری در تصاویر آنژیوگرافی از شدت روشنایی بیشتری برخوردارند، کافی است یک حد آستانه مناسب که آن را  $\theta_1$  می‌نامیم در نظر بگیریم. استفاده از حد آستانه برای دوبخشی کردن تصویر روشی معمول در تقطیع تصویر است [۲۲]. که ما در این قسمت دو حد آستانه مختلف  $\theta_1, \theta_2$  را روی تصویر اعمال نموده‌ایم.

شکل ۴-۲ (ج) تصویر آستانه‌گذاری شده با حد آستانه  $\theta_1 = 0.7$  را نمایش می‌دهد که علاوه بر رگ‌ها نواحی دیگری از تصویر را نیز شامل شده است.

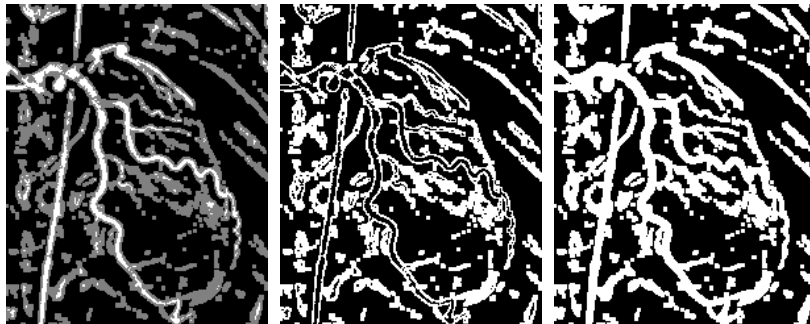
شکل ۴-۲ (د) تصویر آستانه‌گذاری شده با حد آستانه  $\theta_2 = 0.85$  را نمایش می‌دهد که با خطای کمتری فقط رگ‌ها را نمایش می‌دهد.

به دلیل وجود نواحی دیگری به جز شریان‌های کرونری که آن‌ها هم روشن هستند (مانند استخوان‌های قفسه سینه و ستون فقرات)، استفاده از این شیوه و یا نسخه‌های محلی آن خروجی مناسبی را باعث نخواهد شد.



(آ) تصویر ورودی (ب) افزایش کنتراست تصویر (ج) تصویر آستانه گذاری شده (د) تصویر آستانه گذاری شده با حد آستانه  $\theta_1$  با حد آستانه  $\theta_2$

شکل ۲-۴: مراحل تولید نقشه‌ی سه‌گانه خودکار متناظر با تصویر شریان‌های کرونری



(آ) اتساع یافته تصویر آستانه (ب) تفاضل تصاویر ۴-۳ (ج) ترکیب تصاویر ۴-۲ (د) و گذاری شده با حد آستانه  $\theta_1$  (د) ۲ و (آ) ۳-۴ (ب) ایجاد نقشه سه‌گانه خودکار

شکل ۳-۴: ادامه مراحل تولید نقشه‌ی سه‌گانه خودکار متناظر با تصویر شریان‌های کرونری

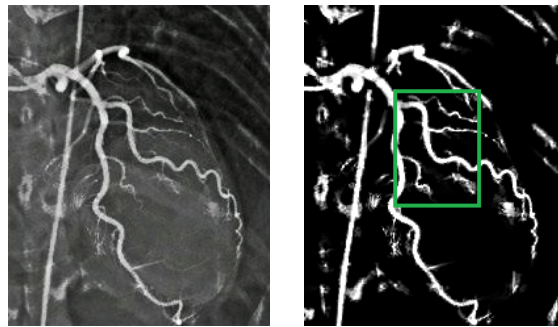
- در مرحله بعد برای ایجاد ناحیه‌ی خاکستری در نقشه سه‌گانه، اتساع یافته تصویر آستانه‌گذاری شده با حد آستانه  $\theta_1$  را ایجاد می‌کنیم.
- سپس تفاضل تصویر ۴-۲ (د) و ۳-۴ (آ) را به عنوان ناحیه خاکستری در نظر می‌گیریم.
- سرانجام از ترکیب تصویر ۴-۲ (ج) و ۳-۴ (ب) نقشه سه‌گانه خودکار حاصل می‌شود.

شکل ۲-۴ و ۳-۴ مراحل تولید نقشه سه‌گانه خودکار را نمایش می‌دهند.

## ۳-۴ پیاده‌سازی روش

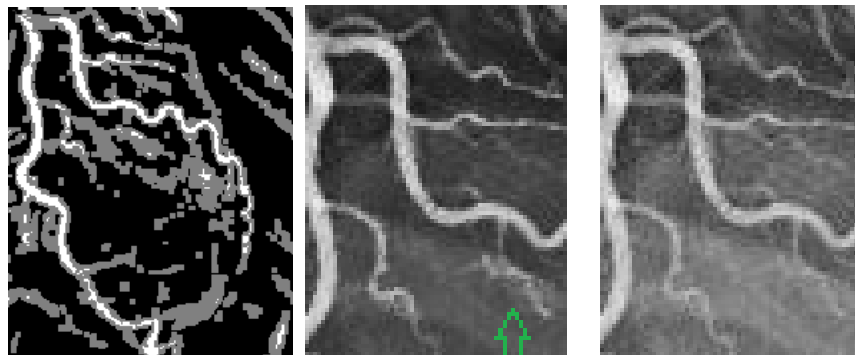
تصویر ۲-۴ (آ) را به عنوان تصویر ورودی و نقشه سه‌گانه تولید شده به صورت خودکار (شکل ۳-۴ (ج)) به عنوان  $\alpha^*$  در رابطه (۷-۲) قرار می‌دهیم و  $\alpha$  بهینه را محاسبه می‌کنیم.





شکل ۴-۴: (آ) آلفای بدست آمده از کمینه سازی (ب) تصویر بهبود یافته

شکل ۴-۴: خروجی روش برش هوشمند بر روی عروق کرونری



شکل ۴-۵: (آ) برشی از تصویر ورودی (ب) برش متناظر از تصویر خودکار (ج) برش متناظر از نقشه سه گانه

شکل ۴-۵: نمایش بهتر نتایج حاصل از روش برش هوشمند برای تصاویر آنژیوگرافی قلب

تصویر ۴-۴ (آ) آلفای محاسبه شده را نشان می دهد. ترکیب آلفای به دست آمده و تصویر ورودی، تصویر ۴-۴ (ب) است، که نمایش بهتر رگ ها در آن مشهود است. برای مقایسه بهتر قسمتی از تصویر ورودی و تصویر بهبود یافته را بزرگ تر کرده و در شکل ۴-۵ نمایش داده ایم. ناحیه متناظر با این قسمت در تصویر آلفا با یک کادر در میانه تصویر نشان داده شده است. اگر به نقشه سه گانه دقت شود مشاهده خواهید کرد که این رگ جزو داده های آموزشی نبوده است و آشکارسازی آن نتیجه فرآیند بهینه سازی (۲-۶) بوده است. مشهود بودن بهتر رگ ها در برخی موارد و من جمله رگی که در آخرین تصویر با یک فلش نمایانده شده است، کاملاً واضح است. ادعا نمی شود که روش مورد استفاده در این بخش به تنهایی برای امر آشکارسازی رگ های قلب کفایت می کند؛ بلکه نظر به خروجی های مناسب آن، می تواند در کنار سایر روش های مورد استفاده در این حوزه به بهبود نتیجه کلی کمک نماید.

# فهرست منابع

- [1] Levin, Anat, Lischinski, Dani, and Weiss, Yair. A closed form solution to natural image matting. in *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1*, CVPR '06, pp. 61–68, Washington, DC, USA, 2006. IEEE Computer Society.
- [2] Wang, Jue and Cohen, Michael F. Image and video matting: A survey. *Foundations and Trends in Computer Graphics and Vision*, 3(2):97–175, 2007.
- [3] Hefferon, Jim. *Linear Algebra*. Virginia Commonwealth University Mathematics, 2006.
- [4] Chong, E.K.P., Zak, S.H., and Žak, S.H. *An Introduction to Optimization*. Wiley Series in Discrete Mathematics and Optimization. Wiley, 2013.
- [5] لیتهد، لوئیس. حساب دیفرانسیل و انتگرال. ترجمه‌ی ناظمی، بهزاد، رزاقی، کاظمی. مرکز نشر دانشگاهی، ویرایش بیست و چهارم، ۱۳۸۶.
- [6] Bishop, Christopher M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
- [7] Shawe-Taylor, John and Cristianini, Nello. *Kernel Methods for Pattern Analysis*. Cambridge University Press, New York, NY, USA, 2004.
- [8] چنی، دیوید کینکید، وارد. آنالیز عددی. ترجمه‌ی صائمی، فائزه توتونیان، منصوره. دانشگاه امام رضا (ع)، ویرایش ششم، ۱۳۸۸.
- [9] Neumaier, Arnold. Solving ill-conditioned and singular linear systems: A tutorial on regularization. *SIAM Review*, 40:636–666, 1998.
- [10] Pendse, Gautam V. A tutorial on the lasso and the “shooting algorithm”. tech. rep., P.A.I.N Group, Imaging and Analysis Group - McLean Hospital, Harvard Medical School, 2011.

- [11] Hastie, Trevor, Tibshirani, Robert, and Friedman, Jerome. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc., New York, NY, USA, 2008.
- [12] Fu, W. J. Penalized regressions: the bridge versus the lasso. *Journal of Computational and Graphical Statistics*, 7(3):397–416, 1998.
- [13] Tseng, P., for Intelligent Control Systems (U.S.), Center, of Technology. Laboratory for Information, Massachusetts Institute, and Systems, Decision. *Nondifferentiable Concave Functions, Coordinate Ascent for Maximizing*. CICS (Series). Center for Intelligent Control Systems, M.I.T., 1988.
- [14] Tibshirani, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society (Series B)*, 58:267–288, 1996.
- [15] Zheng, Yuanjie and Kambhamettu, Chandra. Learning based digital matting. in *ICCV*, pp. 889–896. IEEE, 2009.
- [16] Holland, Paul W. Weighted ridge regression: Combining ridge and robust regression methods. Working Paper 11, National Bureau of Economic Research, September 1973.
- [17] بهبودیان، جواد. آمار و احتمال مقدماتی. دانشگاه امام رضا (ع)، ویرایش سی ام، ۱۳۸۶.
- [18] Gao, Junbin, Paul, Manoranjan, and Liu, Jun. The image matting method with regularized matte. in *ICME*, pp. 550–555. IEEE Computer Society, 2012.
- [19] Johnson, Jubin, Rajan, Deepu, and Cholakkal, Hisham. Sparse codes as alpha matte. in Valstar, Michel François, French, Andrew P., and Pridmore, Tony P., eds., *BMVC*. BMVA Press, 2014.
- [20] Gao, Junbin, Paul, Manoranjan, and Liu, Jun. The image matting method with regularized matte. in *Proceedings of the 2012 IEEE International Conference on Multimedia and Expo, ICME 2012, Melbourne, Australia, July 9-13, 2012*, pp. 550–555. IEEE Computer Society, 2012.
- [21] Lee, Philip Gregory and Wu, Ying. L1 matting. in *ICIP*, pp. 4665–4668. IEEE, 2010.
- [22] Gonzalez, Rafael C. and Woods, Richard E. *Digital Image Processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 3rd ed., 2006.

- [23] He, Kaiming, Sun, Jian, and Tang, Xiaoou. Guided image filtering. in *Proceedings of the 11th European Conference on Computer Vision: Part I*, ECCV'10, pp. 1–14, Berlin, Heidelberg, 2010. Springer-Verlag.

# پیوست آ

## آ-۱ روش دیگر اثبات قضیه (۱-۱-۳)

در این قسمت ابتدا مقدماتی که برای اثبات قضیه (۱-۱-۳) ضروری است، مطرح می‌کنیم و سپس روش دیگری از اثبات قضیه (۱-۱-۳) را که در [۲۳] آمده است، بیان می‌کنیم. در بسیاری از روش‌ها، لازم است که یک تابع هزینه درجه دوم بهینه‌سازی شود و یک سیستم خطی حل گردد که به طور ضمنی معادل فیلترسازی یک تصویر توسط ماتریس معکوس می‌باشد. در برش هوشمند تصویر، ماتریس لاپلاسیان به منظور طراحی ضریب آلفا به عنوان یک تبدیل خطی محلی از رنگ‌های تصویر، ایجاد می‌شود [۲۳].

## آ-۲ فیلتر راهنما

ابتدا یک فرآیند فیلتر کردن وابسته به انتقال خطی کلی<sup>۱</sup> را تعریف می‌کنیم، که شامل تصویر راهنما<sup>۲</sup>  $(I)$ ، یک تصویر ورودی فیلترسازی<sup>۳</sup>  $(p)$  و تصویر خروجی<sup>۴</sup>  $(q)$  است.  $I$  و  $p$  قبل از انجام عملیات معلوم هستند و می‌توانند یکسان باشند. خروجی فیلترسازی در پیکسل  $i$ ، به عنوان میانگین وزن‌دهی شده به صورت زیر تعریف می‌شود:

$$q_i = \sum_j W_{ij}(I)p_j \quad (1-A)$$

<sup>۱</sup>Translation-Variant

<sup>۲</sup>Guidance image

<sup>۳</sup>Filtering input image

<sup>۴</sup>Output image

که  $i$  و  $j$  اندیس پیکسل‌ها می‌باشد. فیلتر کرنل  $W_{ij}$  یک تابع از تصویر راهنما ( $I$ ) می‌باشد و مستقل از  $p$  است. این فیلتر نسبت به  $p$  خطی است.

فیلترهای میانگین‌گیر وزنی یک تابع درجه دوم را بهینه‌سازی می‌کند و یک سیستم خطی به فرم زیر را حل می‌کند:

$$Aq = p \quad (2-آ)$$

که  $p$  و  $q$  بردارهای  $1 \times N$  با مؤلفه‌های  $\{q_i\}$   $\{p_i\}$  هستند و  $A$  یک ماتریس  $N \times N$  فقط وابسته به  $I$  است. جواب معادله‌ی (2-آ) به صورت  $q = A^{-1}P$  است، که با توجه به معادله‌ی (1-آ)،  $W_{ij} = (A^{-1})_{ij}$  است. اکنون به معرفی فیلتر راهنما می‌پردازیم:

فرض کلیدی در فیلتر راهنما، مدل خطی محلی بین راهنمای  $I$  و خروجی  $q$  است. فرض می‌کنیم که  $q$  یک تبدیل خطی از  $I$  در پنجره‌ی  $w_k$  به مرکزیت پیکسل  $k$  ام باشد:

$$q_i = a_k I_i + b_k \quad \forall i \in w_k \quad (3-آ)$$

که  $(a_k, b_k)$  ضرایب خطی هستند که فرض می‌شود روی پنجره‌ی  $w_k$  ثابت باشند. این مدل خطی محلی، نتیجه می‌دهد که  $q$  دارای لبه است اگر فقط  $I$  دارای لبه باشد زیرا  $\nabla q = a \nabla I$  است. برای تعیین ضرایب خطی  $(a_k, b_k)$ ، قیدهایی از ورودی  $p$  مورد نیاز است. خروجی  $q$  را به عنوان اختلاف بین ورودی  $p$  و برخی مؤلفه‌های ناخواسته نظیر نویز در نظر می‌گیریم. یعنی:

$$q_i = p_i - n_i \quad (4-آ)$$

اکنون به دنبال یافتن جوابی هستیم که اختلاف بین  $p$  و  $q$  کمینه باشد، و مدل خطی (3-آ) نیز برقرار باشد. به طور ویژه تابع هزینه زیر را کمینه می‌کنیم:

$$E(a_k, b_k) = \sum_{i \in w_k} ((a_k I_i + b_k - p_i)^2 + \epsilon a_k^2) \quad (5-آ)$$

که  $\epsilon$  پارامتر منظم‌ساز است. معادله‌ی (5-آ) همان رگرسیون ریب می‌باشد که جواب آن به صورت زیر حاصل

می‌شود:

$$a_k = \frac{\frac{1}{|w|} \sum_{i \in w_k} I_i p_i - \mu_k \bar{p}_k}{\sigma_k^2 + \epsilon} \quad (6-\bar{A})$$

$$b_k = \bar{p}_k - a_k \mu_k \quad (7-\bar{A})$$

که  $\mu_k$  و  $\sigma_k^2$  به ترتیب میانگین و واریانس تصویر  $I$  در پنجره  $w_k$  است،  $|w|$  تعداد پیکسل‌های موجود در پنجره  $w_k$  می‌باشد.  $\bar{p}_k = \frac{1}{|w|} \sum_{i \in w_k} p_i$  میانگین  $p$  در پنجره  $w_k$  است. با معلوم شدن  $(a_k, b_k)$ ، می‌توان با استفاده از معادله  $(3-\bar{A})$  خروجی فیلتر سازی یعنی  $q_i$  را محاسبه نمود.

پیکسل  $i$  در همه پنجره‌های هم‌پوشانی شده  $w_k$  که  $i$  را می‌پوشانند، موجود می‌باشد، بنابراین مقدار  $q_i$  در معادله  $(3-\bar{A})$  هنگامی که در پنجره‌های مختلف محاسبه می‌شود، یکسان نیست. یک استراتژی ساده، میانگین‌گیری همه مقادیر ممکن  $q_i$  می‌باشد. بنابراین بعد از محاسبه  $(a_k, b_k)$  برای همه پنجره‌های  $w_k$  در تصویر، خروجی فیلتر سازی به صورت زیر محاسبه می‌شود:

$$q_i = \frac{1}{|w|} \sum_{k|i \in w_k} (a_k I_i + b_k) \quad (8-\bar{A})$$

از طرفی بنا به تقارن پنجره‌های مربعی،  $\sum_{k|i \in w_k} a_k = \sum_{k \in w_i} a_k$  می‌باشد. بنابراین معادله  $(8-\bar{A})$  را می‌توان به صورت زیر بازنویسی نمود:

$$q_i = \bar{a}_i I_i + \bar{b}_i \quad (9-\bar{A})$$

که  $\bar{a}_i = \frac{1}{|w|} \sum_{k \in w_i} a_k$  و  $\bar{b}_i = \frac{1}{|w|} \sum_{k \in w_i} b_k$ ، میانگین ضرایب همه پنجره‌های هم‌پوشانی شده  $i$  می‌باشد.

## ۱-۲-آ فیلتر کرنل

به راحتی می‌توان نشان داد که روابط بین  $p, I, q$  در معادلات  $(6-\bar{A})$ ،  $(7-\bar{A})$  و  $(9-\bar{A})$  به فرم میانگین وزن‌دهی شده مطابق معادله  $(1-\bar{A})$  می‌باشد. در حقیقت معادله  $(6-\bar{A})$ ، را می‌توان به عنوان مجموع وزن‌دهی شده  $p$  بازنویسی کرد، یعنی:  $a_k = \sum_j A_{kj}(I) p_j$  که  $A_{ij}$  نشان دهنده وزن‌ها است و فقط به  $I$  وابسته است. به صورت مشابه بنا به معادله  $(7-\bar{A})$  داریم:  $b_k = \sum_j B_{kj}(I) p_j$  و همچنین بنا به معادله  $(9-\bar{A})$  خواهیم داشت:

$$q_i = \sum_j W_{ij}(I)p_j$$

قضیه آ-۲-۱. وزن‌های کرنل  $(W_{ij})$ ، به صورت زیر محاسبه می‌شود:

$$W_{ij}(I) = \frac{1}{|w|} \sum_{k:(i,j) \in w_k} \left( 1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{\sigma_k^2 + \epsilon} \right) \quad (10-\bar{A})$$

برهان. بنا به وابستگی خطی بین  $p$  و  $q$ ، فیلتر کرنل به صورت زیر حاصل می‌شود:

$$W_{ij} = \frac{\partial q_i}{\partial p_j}$$

با قرار دادن معادله‌ی (۷-آ) در معادله‌ی (۹-آ) و حذف  $b$ ، خواهیم داشت:

$$q_i = \frac{1}{|w|} \sum_{k \in w_i} (a_k(I_i - \mu_k) + \bar{p}_k) \quad (11-\bar{A})$$

که با مشتق‌گیری از آن رابطه‌ی زیر حاصل می‌شود:

$$\frac{\partial q_i}{\partial p_j} = \frac{1}{|w|} \sum_{k \in w_i} \left( \frac{\partial a_k}{\partial p_j} (I_i - \mu_k) + \frac{\partial \bar{p}_k}{\partial p_j} \right) \quad (12-\bar{A})$$

در این معادله

$$\frac{\partial \bar{p}_k}{\partial p_j} = \frac{1}{|w|} \delta_{j \in w_k} = \frac{1}{|w|} \delta_{k \in w_j} \quad (13-\bar{A})$$

است، که  $\delta_{j \in w_k}$  هنگامی که  $j$  در پنجره  $w_k$  قرار دارد، برابر یک است و در غیر این صورت صفر می‌باشد. از طرفی دیگر  $\frac{\partial a_k}{\partial p_j}$  در معادله‌ی (۱۲-آ) را می‌توان با استفاده از معادله (۶-آ) به صورت زیر محاسبه نمود:

$$\begin{aligned} \frac{\partial a_k}{\partial p_j} &= \frac{1}{\sigma_k^2 + \epsilon} \left( \frac{1}{|w|} \sum_{i \in w_k} \frac{\partial p_i}{\partial p_j} I_i - \frac{\partial \bar{p}_k}{\partial p_j} \mu_k \right) \\ &= \frac{1}{\sigma_k^2 + \epsilon} \left( \frac{1}{|w|} I_j - \frac{1}{|w|} \mu_k \right) \delta_{k \in w_j} \end{aligned} \quad (14-\bar{A})$$



با قرار دادن معادلات (آ-۱۳) و (آ-۱۴) در معادله (آ-۱۲) خواهیم داشت:

$$\frac{\partial q_i}{\partial p_j} = \frac{1}{|w|^2} \sum_{k:(i,j) \in w_k} \left( 1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{\sigma_k^2 + \epsilon} \right) \quad (\text{آ-۱۵})$$

که نشان‌دهنده فیلتر کرنل  $W_{ij}$  می‌باشد.

□

## آ-۲-۲ ارتباط فیلتر کرنل و روش فرم بسته

فیلتر راهنما ارتباط تنگاتنگی با ماتریس لاپلاسین معرفی شده در روش فرم بسته دارد. در روش فرم بسته برای برش هوشمند تصویر، ماتریس لاپلاسین براساس یک مدل خطی محلی حاصل می‌شود. برخلاف فیلتر راهنما که بهینه محلی را برای هر پنجره محاسبه می‌کند، روش فرم بسته بهینه سراسری را محاسبه می‌کند. برای به دست آوردن ضریب مجهول آلفا، این روش تابع هزینه زیر را کمینه می‌کند:

$$E(q) = (q - p)^T \Lambda (q - p) + q^T L q \quad (\text{آ-۱۶})$$

که در اینجا  $q$  یک بردار  $1 \times N$  است که ضریب مجهول آلفا را نمایش می‌دهد و  $p$  نشان‌دهنده محدودیت‌هایی نظیر نقشه سه‌گانه می‌باشد.  $L$  ماتریس لاپلاسین با بعد  $N \times N$  است، و  $\Lambda$  ماتریس قطری ایجاد شده براساس محدودیت‌ها می‌باشد. جواب این مسئله بهینه‌سازی به وسیله حل سیستم خطی زیر حاصل می‌شود:

$$(L + \Lambda)Q = \Lambda P \quad (\text{آ-۱۷})$$

مؤلفه‌های ماتریس لاپلاسین به صورت زیر است:

$$L_{ij} = \sum_{k:(i,j) \in w_k} \left( \delta_{ij} - \frac{1}{|w|} \left( 1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{\sigma_k^2 + \epsilon} \right) \right) \quad (\text{آ-۱۸})$$

که  $\delta_{ij}$  دلتای کرونیکر است. با مقایسه معادله‌ی (آ-۱۸) با معادله‌ی (آ-۱۰) دژمی‌یابیم که مؤلفه‌های ماتریس لاپلاسین مستقیماً توسط فیلتر راهنمای کرنل به صورت زیر حاصل می‌شود:

$$L_{ij} = |w| (\delta_{ij} - W_{ij}) \quad (\text{آ-۱۹})$$

## برنامه‌های MATLAB روش‌های مطرح شده در پایان نامه ۳-آ

در این قسمت بخش اصلی برنامه‌های نوشته شده برای پایان نامه را ذکر می‌نمایم.

برنامه آ-۱: کد تخمین پارامتر منظم‌ساز  $\lambda$  با به کارگیری اعتبارسنجی متقابل

```
function z = estimateLassoLambda( y, X, kfold, lambda_vec )
if ( nargin ~= 4 )
    disp('Usage: z = estimateLassoLambda( y, X, kfold, lambda_vec )
        ');
    z = [];
    return;
end
% check size of y
[n1, p1] = size(y);
% is y a column vector?
if ( p1 ~= 1 )
    disp('y must be a n by 1 vector!!');
    z = [];
    return;
end
% check size of X
[n2, p2] = size(X);
% does X have the same number of rows as y?
if ( n2 ~= n1 )
    disp('X must have the same number of rows as y!!');
    z = [];
    return;
end
% make sure kfold is positive
kfold = floor(kfold);
if ( kfold < 2 )
    disp('kfold must be >= 2!!');
    z = [];
    return;
end
% make sure lambda_vec(i) > 0
for r = 1:length(lambda_vec)
    if ( lambda_vec(r) < 0 )
        disp('all entries of lambda_vec must be >= 0!!');
        z = [];
        return;
    end
end
% get size of X
[n, p] = size(X);
% partition input data for cross-validation
CVO = cvpartition( n, 'kfold' , kfold );
% compute cross-validation error for each regularization parameter
% initialize a variable to hold average MSEerror across folds
MSEerror = zeros( length(lambda_vec), 1 );
for k = 1:length( lambda_vec )
    % vector to store MSE error for each fold
    err = zeros(CVO.NumTestSets,1);
```

```

for i = 1:CVO.NumTestSets
    % get ith training set
    trIdx = CVO.training(i);
    % get ith test set
    teIdx = CVO.test(i);
    % train LASSO using training set
    sL = solveLasso( y(trIdx), X(trIdx,:), lambda_vec(k) );
    % test using testing set
    ypred = X(teIdx,:) * sL.beta;
    % calculate MSE error
    temp = y(teIdx) - ypred;
    % clear ypred
    clear ypred;
    % calculate average error for the ith test set
    err(i) = (temp'*temp)/length(temp);
end
% calculate the mean error over all test sets
MSEError(k) = mean(err);
end
% where do we get the smallest MSE?
min_index = find( MSEError == min(MSEError) );
min_index = min_index(1);
% set outputs
z.lambda = lambda_vec( min_index );
z.lambda_vec = lambda_vec;
z.MSEError = MSEError;
z.min_index = min_index;
z.CVO = CVO;
z.y = y;
z.X = X;
z.kfold = kfold;
end

```

### برنامه آ-۲: کد الگوریتم شوتینگ

```

beta = (X'*X + 2*lambda) \ (X'*y);
% start while loop
% convergence flag
found = 0;
% convergence tolerance
TOL = 1e-6;
while( found == 0 )
    % save current beta
    beta_old = beta;
    % optimize elements of beta one by one
    for i = 1:p
        % optimize element i of beta
        % get ith col of X
        xi = X(:,i);
        % get residual excluding ith col
        yi = (y - X*beta) + xi*beta(i);
        % calculate xi'*yi and see where it falls
        deltai = (xi'*yi); % 1 by 1 scalar
        if ( deltai < -lambda )
            beta(i) = ( deltai + lambda )/(xi'*xi);
        end
    end
    % check convergence
    found = norm(beta - beta_old) < TOL;
end

```

```

elseif ( deltai > lambda )
    beta(i) = ( deltai - lambda )/(xi'*xi);
else
    beta(i) = 0;
end
end
% check difference between beta and beta_old
if ( max(abs(beta - beta_old)) <= TOL )
    found = 1;
end
end
end

```

برنامه آ-۳: کد برش هوشمند تصویر براساس روش یادگیری با به کارگیری رگرسیون خطی ساده

```

function lapcoeff=compLapCoeff(winI,lambda)
% winI: nxc matrix where n is the number of pixel, c number of features
if ~exist('lambda','var')
    lambda=0.0000001;
end
Xi=winI;
Xi=[Xi ones(size(Xi,1),1)];
I=eye(size(Xi,1)); I(end,end)=0;
% coefficient, i.e. F
fenmu=Xi'*Xi
fenmub=Xi/fenmu
F=fenmub*Xi'
% laplacian coefficnets, i.e. (I-F)*(I-F)'
I_F=eye(size(F,1))-F;
lapcoeff=I_F'*I_F;

```

برنامه آ-۴: کد برش هوشمند تصویر براساس روش یادگیری با به کارگیری رگرسیون ریب

```

function lapcoeff=compLapCoeff(winI,lambda)
% winI: nxc matrix where n is the number of pixel, c
number of features
if ~exist('lambda','var')
    lambda=0.0000001;
end
Xi=winI;
Xi=[Xi ones(size(Xi,1),1)];
I=eye(size(Xi,1)); I(end,end)=0;
% coefficient, i.e. F
fenmu=(Xi*Xi'+(lambda)*I);
F=(Xi*Xi')/fenmu;
% laplacian coefficnets, i.e. (I-F)*(I-F)'
I_F=eye(size(F,1))-F;
lapcoeff=I_F'*I_F;

```

برنامه آ-۵: کد برش هوشمند تصویر براساس روش فرم بسته

```

function alpha=solveAlpha(I,consts_map,consts_vals,varargin)

```

```
[h,w,c]=size(I);  
img_size=w*h;  
A=getLaplacian1(I,consts_map,varargin{:});  
D=spdiags(consts_map(:),0,img_size,img_size);  
lambda=100;  
x=(A+lambda*D)\(lambda*consts_map(:).*consts_vals(:));  
alpha=max(min(reshape(x,h,w),1),0);
```

## واژه نامه فارسی به انگلیسی

Threshold . . . . .	آستانه‌گیری
Collective discrepancy . . . . .	اختلاف جمعی
Partition . . . . .	افراز
Cross validation . . . . .	اعتبارسنجی متقابل
Ill-posed . . . . .	بدطرح
Ill-conditioned . . . . .	بدوضع
Image matting . . . . .	برش هوشمند تصویر
Linear optimization . . . . .	بهینه‌سازی خطی
Nonlinear optimization . . . . .	بهینه‌سازی غیرخطی
Overfitting . . . . .	بیش‌برازش
Numerical stability . . . . .	پایداری عددی
Background . . . . .	پس‌زمینه
Foreground . . . . .	پیش‌زمینه
Continuous . . . . .	پیوسته
Labeled pixels . . . . .	پیکسل‌های برچسب‌گذاری
Unlabeled pixels . . . . .	پیکسل‌های برچسب‌گذاری نشده
Neighboring pixels . . . . .	پیکسل‌های مجاور
Pattern function . . . . .	تابع پیش‌بینی
Prediction function . . . . .	تابع پیش‌بینی
Objective function . . . . .	تابع هدف
Cost function . . . . .	تابع هزینه

Convert . . . . .	تبدیل کردن . . . . .
Extracting . . . . .	تخمین . . . . .
Convexity . . . . .	تحدب . . . . .
Transpose . . . . .	ترانهاده . . . . .
Compositing . . . . .	ترکیب . . . . .
Linear combination . . . . .	ترکیب خطی . . . . .
Approximation . . . . .	تقریب . . . . .
Contradiction . . . . .	تناقض . . . . .
Propagation . . . . .	تعمیم . . . . .
Constant . . . . .	ثابت . . . . .
Substituting . . . . .	جای گذاری . . . . .
Inner producte . . . . .	حاصل ضرب داخلی . . . . .
Eliminated . . . . .	حذف کردن . . . . .
Scribble . . . . .	نشان گذاری . . . . .
Numerical stability . . . . .	پایداری عددی . . . . .
Coefficients . . . . .	ضرایب . . . . .
Training data . . . . .	داده های آموزشی . . . . .
Domain . . . . .	دامنه . . . . .
Interpolates . . . . .	درون یابی . . . . .
Kronecker delta . . . . .	دلتهای کرونیکر . . . . .
Dual . . . . .	دوگان . . . . .
Linear regression . . . . .	رگرسیون خطی . . . . .
Ridge regression . . . . .	رگرسیون ریج . . . . .
Data-response pairs . . . . .	زوج داده-پاسخ . . . . .
Construct . . . . .	ساخت . . . . .
Gray level . . . . .	سطح خاکستری . . . . .
Pseudo-inverse . . . . .	شبه معکوس . . . . .
Intensities . . . . .	شدت روشنایی . . . . .

Necessary conditions . . . . .	شرایط لازم
Assumption . . . . .	فرض
Image segmentation . . . . .	قطعه بندی تصویر
Least squares errors . . . . .	کمترین مربعات خطا
Covariance . . . . .	کواریانس
Minimizing . . . . .	کمینه سازی
Lagrange . . . . .	لاگرانژ
Diagonal matrix . . . . .	ماتریس قطری
Training set . . . . .	مجموعه آموزشی
Data set . . . . .	مجموعه داده ها
Convex set . . . . .	مجموعه محدب
Unknown . . . . .	مجهول
Evaluation . . . . .	محاسبه
Constraint . . . . .	محدودیت
Linear model . . . . .	مدل خطی
Compar . . . . .	مقایسه
Independent . . . . .	مستقل
Restrict . . . . .	مقید کردن
Normal equation . . . . .	معادلات نرمال
Criterion . . . . .	معیار
Uniquely . . . . .	منحصرا
Singular . . . . .	منفرد
Regularisation . . . . .	منظم سازی
Invertible . . . . .	معکوس پذیر
Local minimizer . . . . .	مینیم کننده محلی
Global minimizer . . . . .	مینیم کننده سراسری
Non smooth . . . . .	ناهموار
Region . . . . .	ناحیه



Consequence	نتیجه
Feature mapping	نگاشت طرح
Trimap	نقشه سه گانه
Notation	نماد
Positive semidefinite	نیمه معین مثبت
Representation	نمایش
Semi-supervised	نیمه نظارتی
Diagnal element	عناصر قطری
Image editing	ویرایش تصویر
Goal	هدف
Overlap	هم پوشانی
Machine learning	یادگیری ماشین
Variance	واریانس

## واژه نامه انگلیسی به فارسی

Approximation . . . . .	تقریب
Assumption . . . . .	فرض
Background . . . . .	پس زمینه
Criterion . . . . .	معیار
Cross validation . . . . .	اعتبارسنجی متقابل
Coefficients . . . . .	ضرایب
Collective discrepancy . . . . .	اختلاف جمعی
Compar . . . . .	مقایسه
Compositing . . . . .	ترکیب
Consequence . . . . .	نتیجه
Constant . . . . .	ثابت
Constraint . . . . .	محدودیت
Construct . . . . .	ساخت-ایجاد
Continuous . . . . .	پیوسته
Contradiction . . . . .	تناقض
Convert . . . . .	تبدیل کردن
Convex set . . . . .	مجموعه‌ی محدب
Convexity . . . . .	تحدب
Cost function . . . . .	تابع هزینه
Covariance . . . . .	کواریانس
Data-response pairs . . . . .	زوج داده-پاسخ

Data set . . . . .	مجموعه‌ی داده‌ها . . . . .
Diagonal element . . . . .	عناصر قطری . . . . .
Diagonal matrix . . . . .	ماتریس قطری . . . . .
Domain . . . . .	دامنه . . . . .
Dual . . . . .	دوگان . . . . .
Editing . . . . .	ویرایش . . . . .
Eliminated . . . . .	حذف کردن . . . . .
Evaluation . . . . .	محاسبه . . . . .
Extracting . . . . .	استخراج . . . . .
Feature mapping . . . . .	نگاشت طرح . . . . .
Foreground . . . . .	پیش‌زمینه . . . . .
Global minimizer . . . . .	مینیمم‌کننده‌ی سراسری . . . . .
Gray level . . . . .	سطح خاکستری . . . . .
Goal . . . . .	هدف . . . . .
Ill-conditioned . . . . .	بد وضع . . . . .
Ill-posed . . . . .	بد طرح . . . . .
Image matting . . . . .	برش هوشمند تصویر . . . . .
Independent . . . . .	مستقل . . . . .
Inner producte . . . . .	حاصل ضرب داخلی . . . . .
Intensities . . . . .	شدت روشنایی . . . . .
Interpolates . . . . .	درون‌یابی . . . . .
Inverse . . . . .	معکوس . . . . .
Invertible . . . . .	معکوس‌پذیر . . . . .
Kernel function . . . . .	تابع کرنل . . . . .
Kronecker delta . . . . .	دلته‌ای کرونیکر . . . . .
Label . . . . .	برچسب . . . . .
Lagrange . . . . .	لاگرانژ . . . . .
Least squares . . . . .	کمترین مربعات . . . . .

Linear combination . . . . .	ترکیب خطی
Linear model . . . . .	مدل خطی
Linear regression . . . . .	رگرسیون خطی
Local learning . . . . .	یادگیری محلی
Local minimizer . . . . .	مینیمم کننده‌ی محلی
Machine learning . . . . .	یادگیری ماشین
Mean . . . . .	میانگین
Minimizing . . . . .	کمینه‌سازی
Mixture . . . . .	ترکیب
Neighboring pixels . . . . .	پیکس‌های مجاور
Nonlinear optimization . . . . .	بهینه‌سازی غیرخطی
Non smooth . . . . .	ناهموار
Normal equation . . . . .	معادلات نرمال
Notation . . . . .	نماد
Numerical stability . . . . .	پایداری عددی
Objective function . . . . .	تابع هدف
Optimal . . . . .	بهینه
Optimal solution . . . . .	جواب بهینه
Overlap . . . . .	هم‌پوشانی
Overfitting . . . . .	بیش‌برازش
Partition . . . . .	افراز
Pattern function . . . . .	تابع الگو
Positive semidefinite . . . . .	نیمه‌معین مثبت
Predict . . . . .	پیش‌بینی
Prediction function . . . . .	تابع پیش‌بینی
Propagation . . . . .	تعمیم
Pseudo-inverse . . . . .	شبه معکوس
Quadratic optimization problem . . . . .	مسئله بهینه‌سازی درجه دوم

Region . . . . .	ناحیه . . . . .
Regularisation . . . . .	منظم‌سازی . . . . .
Representation . . . . .	نمایش . . . . .
Restrict . . . . .	مقید کردن . . . . .
Ridge regression . . . . .	ریج رگرسیون . . . . .
Scribble . . . . .	نشان‌گذاری . . . . .
Segmentation . . . . .	قطعه‌بندی . . . . .
Singular . . . . .	منفرد . . . . .
Substituting . . . . .	جای‌گذاری . . . . .
Taylor theorem . . . . .	قضیه تیلور . . . . .
Training set . . . . .	مجموعه‌ی آموزشی . . . . .
Transpose . . . . .	ترانهاد . . . . .
Threshold . . . . .	آستانه‌گیری . . . . .
Trimap . . . . .	نقشه سه‌گانه . . . . .
Uniquely . . . . .	منحصرا . . . . .
Unlabeled pixels . . . . .	پیکسل‌های برچسب‌گذاری نشده . . . . .
Variance . . . . .	واریانس . . . . .
Vector . . . . .	بردار . . . . .

Hakim Sabzevari University

An Outline of MSc. Thesis



Surname: Rashidabadi

Name: Farzane

Student No.: 9213133042

Supervisor: Dr. Mahmood Amintoosi

Advisor: Dr. Alireza Ghodsi

Faculty of Mathematics and  
Computer Science

Applied Mathematics

Operational Research

Title of thesis: Image Matting

Keywords: Image matting, Ridge regression, Regularisation, Semi-supervised, Training data

Abstract: Interactive digital matting, the process of extracting a foreground object from an image based on limited user input, is an important task in image and video editing. From a computer vision perspective, this task is extremely challenging because it is massively ill-posed at each pixel.

In this thesis, some famous methods in this context including “Learning Based” and “Closed Form” are investigated and the results on various images is shown.

As a new application of the image matting, good performance of this method on enhancing coronary heart images is proposed and demonstrated here.



**Hakim Sabzevari University**  
**Faculty of Mathematics and Computer Science**

**A Thesis Submitted in Partial Fulfilment of the Requirement for the  
Degree of Master of Science in Applied Mathematics**

# **Image Matting**

**Supervisor:**  
**Dr. Mahmood Amintoosi**

**Advisor:**  
**Dr. Alireza Ghodsi**

**By:**  
**Farzane Rashidabadi**

**Jan 2016**