الباب الثامن

طريقة الحاجز اللوغاريتمية الأساسية
Primal Logarithmic Barrier Method

• مقدمة • طمر المسألة ذاتية الثنائية • استراتيجية الطمر وحلها • الخوارزمية الشكلية ودالة الأوسطية • إسقاط اتجاه نيوتن لدالة الحاجز الأساسية • اتجاه الموازنة التآلفي • السلوك قرب المسار الأوسط • تحديث وسيط المسار الأوسط • خوارزمية الثنائية • طرق التحديث الكبيرة

1.8 مقدمة Introduction

إذا كانت مسألة البرمجة الموجبة شبه المعرفة هي ذاتية الثنائية self dual أي أن المسألتان الأساسية والثنائية متطابقتان، فإن علاقات الثنائية تكون أبسط. وبشكل خاص إذا كانت هذه المسألة التي لها نفس الثنائية ولها حلول مسموح بها فعليّة فإن هذه المسألة قابلة للحل، وقيمة الحل الأمثل هي الصفر وهذا نستنتجه مباشرة من النظرية الثنائية القوية (نظرية 6.6). في هذا

الباب سوف نوضح كيف يمكننا تحويل مسألتي الأساسية والثنائية بالصيغة القياسية إلى مسألة أكبر ولها نفس الثنائية. وحيث الحلول المسموح بها الفعلية على المسار الأوسط معلومة، فمن الممكن حل المسألة الكبرى. وحلها سيقودنا إلى حل منطقة الحلول المسموح بها للمسألتين الأساسية و الثنائية وبالتالي إلى حلول لهما.

بعد ذلك سوف نقوم بتحليل طريقة الحاجز الأساسية للبرمجة الموجبة شبه المعرفة وهذه الطريقة أيضاً تُعرف بطريقة تابع المسار الأساسية، وذلك لأن طريقة هذه الخوارزمية هي بمتابعة المسار الأوسط تقريباً للوصول للحل الأمثل. وبشكل خاص سوف نتطرق هنا إلى تحليل بسيط لما يسمى طريقة تقريب وسط البرنامج الموجب شبه المعرف [HDRT].

إن الطرق التي تدرس حل المسألة الأساسية بشكل بحت تستخدم معلومات عن المسألة الأساسية ومنطقة الحلول المسموح بها، وكذلك الحال بالنسبة للمسألة الثنائية. ولكن طرق الأساسية الثنائية تستخدم معلومات عن كلتا المسألتين وذلك في صياغة اتجاه البحث. وهناك الكثير من مسائل البرمجة الموجبة شبه المعرفة وخاصة في مجال الأمثلة التركيبية combinatorial optimization. حيث الحلول المسموح بها الأساسية X أو الحلول المسموح بها الثنائية S تكون إحداهما متناثرة والأخرى ليست كذلك. وللاستفادة من هذا التناثر لابد من استخدام طرق المسألة الأساسية البحتة أو طرق المسألة الثنائية البحتة.

2.8 طمر المسألة ذاتية الثنائية Self-Dual Embeddings

أكثر الطرق المستخدمة في حل مسائل البرمجة الموجبة شبه المعرفة تتطلب وجود حلول مسموح بها فعلية، أي أن $X^{\circ} \succ 0 \,,\, S^{\circ} \succ 0$ لكلٍ من

المسألة الأساسية والمسألة الثنائية. على سبيل المثال: اعتبر المسألة الثنائية للبرمجة الموجبة شبه المعرفة بشكلها القياسي التالي:

$$\begin{aligned} \text{maximize} \quad & b^T y \\ \text{subject to} \quad & \sum_{i=1}^{m} y_i A_i + S = C \\ & S \succeq 0 \ , \ y \in \mathbb{R}^m. \end{aligned} \tag{1.8}$$

ولنفترض أن حلاً مسموحاً به فعلياً للمسألة الثنائية موجود ومعروف، وليكن (y°, S°) ، ولكن لا يوجد مصفوفة أساسية مسموح به فعلياً لثنائية لاجرانج أي (1.8). وهي على الشكل القياسي التالي:

$$\begin{aligned} \text{minimize} \quad & \text{tr}\,(CX) \\ \text{subject to} \quad & \text{tr}\,(A_i X) = b_i \ , \qquad i = 1, \ldots, m \\ & X \succeq 0 \end{aligned} \tag{2.8}$$

ولكي نطبق الطرق الأساسية الثنائية يجب أولاً أن نجد حلاً مسموحاً به لكلتا المسألتين، بعد ذلك يجب علينا حل المسألة المعدّلة التالية:

$$\begin{aligned} \text{maximize} \quad & b^T y \\ \text{subject to} \quad & \sum_{i=1}^{m} y_i A_i + S = C \\ & \text{tr}\, S \le M, \ S \succeq 0 \end{aligned} \tag{3.8}$$

وبشكل غير دقيق فإن (3.8) لها نفس حل (1.8)، إذا كانت M كبيرة بشكل كافٍ، وإذا كانت (1.8) ليس لها حلٌ مسموح به فبالمثل (3.8)، وإذا كانت (1.8) قابلة للحل فكذلك (3.8).

مسألة أخرى تختلف قليلاً عن (3.8) هي

$$\begin{aligned} \text{minimize} \quad & \text{tr}\,(C(X - \kappa I)) + \kappa M \\ \text{subject to} \quad & \text{tr}\,(A_i(X - \kappa I)) = b_i, \quad i = 1, \ldots m \\ & \kappa \ge 0, \ X \succeq 0 \end{aligned} \tag{4.8}$$

الآن نستطيع أن نبني حلاً مسموحاً به فعلياً كنقطة بداية للمسألة (4.8) وذلك باختيار κ كبيرة بشكل كافٍ، بينما لا يوجد مثل هذا الحل للمسألة الأساسية.

نستطيع أن نقوم باستراتيجية مشابهة إذا كانت المسألة الأساسية لها حلٌ مسموح به فعلياً. أما إذا كان كلٌ من المسألة الأساسية والثنائية ليس لهما حلٌ مسموح به فإن استراتيجية مشابهة بإضافة المصفوفة M لكلتا المسألتين. الصعوبة هنا عدم وجود اختيار مسبق للمصفوفة M بشكل عام. على سبيل المثال إذا لم تستطع جعل الوسيط κ يقترب من الصفر في المسألة (4.8) فهذا يعني عدم وجود حلٍ أمثلياً S للمسألة الثنائية حيث $\operatorname{tr} S^* \leq M$. وهذا يعني أننا نحتاج إلى حد لـ $\operatorname{tr} S^*$ وذلك لمعرفة حالة المسألة الثنائية. وهذه المعلومات ليست معروفة مقدماً بشكل عام.

في حالة البرمجة الخطية يوجد حل أنيق للمسألة الابتدائية initialization problem وذلك بطمر المسألة الأصلية بمسألة ذاتية الثنائية متماثلة تخالفياً، skew-symmetric والتي لها حل مسموح به داخلي معلوم يقع على المسار الأوسط. إن حل المسألة المطمورة يعطينا الحل الأمثل للمسألة الأصلية، ويتضح منه أنه إما المسألة الأساسية أو المسألة الثنائية ليس لها حلول مسموح بها. وفي هذه الحالة لا بد من استنتاج معلومات مفصلة أكثر عن الحل.

رغم وجود هذه الخصائص النظرية المشجعة لطمر ذاتية الثنائية، إلا أن الفكرة لم تلق قبولاً كبيراً في التطبيق، وذلك لأن مسألة الطمر لها عمود كثيف في مصفوفة المعاملات. وهذا يؤدي إلى امتلاء مفكوك شولسكي choleski factorization خلال الحسابات. ورغم كل هذا إلا أن [XHY] استطاع أن يطبق بنجاح طريقة طمر ذاتية الثنائية على البرمجة الخطية،

وكان هذا التطبيق واسع النجاح، حتى أن البرنامج الذي استنتجوه استُخدم تجارياً في برامج مثل MOSEK، XPRESSMP و CPLEX.

طُوِرت فكرة الطمر المتجانسة homogeneous embedding للبرمجة الموجبة شبه المعرفة بواسطة [PS]. وبعد ذلك قام [DRT] و [LSZ] بتمديد استراتيجية الطمر ليستنتج مسائل الطمر ذاتية الثنائية حيث تكون داخل منطقة الحل المسموح بها غير خالية. وعلى عكس الطمر المتجانس مسألة الطمر الناتجة، لها نقطة بداية على المسار الأوسط. وهذا يبسط التحليل، لأن المسار الأوسط معرّف جيداً. هذه الطريقة اُستخدمت في برنامج SeDuMi المطور بواسطة [St].

إن الحل إذا كان نقطة نهاية على المسار الأوسط لمسألة طمر، فإنه يكون لدينا إحدى الحالات التالية للمسألة الأصلية:

- زوج من الحلول المتتامة $(X^*, S^*) \in \mathcal{P}^* \times \mathcal{D}^*$.
- سهم محسن للمسألة الأساسية أو المسألة الثنائية أو كلتاهما.
- لا يوجد زوج من الحلول وبالتالي كلٌ من المسألتين الأساسية والثنائية لا يوجد لهم حلول مسموح بها قوية.

بشكل عام وغير دقيق المسألتان قابلتان للحل إذا وجد زوج من الحلول المتتامة أو كان لهما حلول مسموح بها قوية.

تختلف البرمجة الموجبة شبه المعرفة عن البرمجة الخطية، لأنه في حالة البرمجة الموجبة شبه المعرفة من الممكن وقوع الحالات التالية:

- وجود قيمة موجبة للفجوة الثنائية عند زوج الحلول المثلى للمسألتين.
- وجود فجوة ثنائية صغيرة ومع ذلك لا يوجد زوج من الحلول المثلى للمسألتين.

- قد يوجد للمسألة الأساسية قيمة مثلى، مع أن المسألة الثنائية ليس لها حل مسموح به.

وهذه الحالات لا يمكن إكتشافها بطريقة الطمر مالم نضيف فرضية جديدة ألا وهي أن المسألتين الأساسية والثنائية لهما الثنائية الكاملة perfect duality. وهذه الفرضية تتحقق إذا كانت المسألة الأساسية لها حلول مسموح بها فعلياً وكذلك المسألة الثنائية لها حلول مسموح بها أيضاً.

3.8 استراتيجية الطمر وحلها The Embedding Strategy

معطى لدينا مسألة الطمر المتجانس للمسألتين الأساسية والثنائية التالية:

$$\begin{array}{llll} \operatorname{tr}(A_i X) - \tau b_i & & = 0, & i = 1, \ldots, m \\ -\sum_{i=1}^{m} y_i A_i & + \tau C - S & = 0 & \\ b^T y \quad - \operatorname{tr}(CX) & \quad - \rho & = 0 & \\ y \in \mathbb{R}^m, \quad X \succeq 0, \quad \tau \geq 0, \quad S \succeq 0, \quad \rho \geq 0 & & & \end{array} \tag{5.8}$$

الحل المسموح به لهذا النظام مع $\tau \geq 0$ يؤدي إلى حل مسموح به $\frac{1}{\tau}X$ و $\frac{1}{\tau}S$ للمسألتين الأساسية والثنائية. والمعادلة الأخيرة تضمن لنا الحل الأمثل، وذلك بطلب وجود فجوة ثنائية غير موجبة. ولهذا السبب لا يوجد حلول مسموح بها فعلياً للمسألة (5.8).

سوف نشرح الآن تمديد لمسألة الطمر ذاتية الثنائية، وذلك لنتمكن من الحصول على حلول مسموح بها فعلياً، وكذلك لنحصل أيضاً على نقطة بداية للمسألة ذاتية الثنائية تكون على المسار الأوسط.

نحصل على مسألة الطمر المسموح بها فعلياً بتمديد مجموعة القيود في (5.8)، وإضافة متغيرات كالتالي:

$$
\begin{array}{llllll}
\text{minimize} & \theta\beta & & & & \\
\text{subject to} & & & & & \\
& \operatorname{tr}(A_i X) - \tau b_i & + \theta \bar{b}_i & & & = 0 \\
-\sum_{i=1}^{m} y_j A_j & + \tau C & - \theta \bar{C} & - S & & = 0 \\
b^T y - \operatorname{tr}(CX) & & + \theta\alpha & & - \rho & = 0 \\
-\bar{b}^T y - \operatorname{tr}(\bar{C}X) + \tau\alpha & & & & - \upsilon & = -\beta \\
y \in \mathbb{R}^m, \quad X \succeq 0, \ \tau \geq 0, \ \theta \geq 0, \ S \succeq 0, \ \rho \geq 0, \ \upsilon \geq 0 & & & & &
\end{array}
\tag{8.6}
$$

حيث

$$
\begin{aligned}
&\bar{b}_i = b_i - \operatorname{tr} A_i \\
&\bar{C} = C - I \\
&\alpha = 1 + \operatorname{tr} C \\
&\beta = n + 2 \\
&i = 1, \ldots, n
\end{aligned}
$$

من الواضح التحقق من وجود حل مسموح به داخلياً كنقطة بداية، وذلك باختيار:

$$y^{\circ} = 0, \ X^{\circ} = S^{\circ} = I, \ \theta^{\circ} = \rho^{\circ} = \tau^{\circ} = \upsilon^{\circ} = 1$$

لاحظ أن الحل $\upsilon = \beta$ ، وباقي المتغيرات صفر هو حل أمثل، لأن دالة الهدف دائماً غير سالبة. وبمعنى آخر نجد أن $\theta = 0$ في أي حل أمثل. وبالتالي فإنه من السهل الحصول على حل أمثل ولكننا هنا نهتم فقط بالحلول المثلى ذات المتممة العظمى.

إن الحلول ذات المتممة العظمى إذا وجدت تضمن لنا حلاً أمثلاً لمسألة الطمر ذاتية الثنائية حيث $\tau^* > 0$ ، وسنبين ذلك لاحقاً. إن المتممة العظمى هي

نقطة نهاية المسار الأوسط لمسألة الطمر، وسنعطي الآن نظرية تبين ثنائية مسألة الطمر.

نظرية 1.8

لتكن $\mathcal{K} \subset \mathbb{R}^k$ مخروطاً محدباً مغلقاً closed convex cone حيث ثنائي المخروط $\mathcal{K}$ هو $\mathcal{K}^*$ ولتكن $A \in \mathbb{R}^{k \times k}$ هي مصفوفة متماثلة تخالفياً. إن ثنائية لاجرنج لمسألة الأمثلة التالية:

$$\begin{aligned} q(x) = \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax - s = c \\ & x \in \mathcal{K},\ s \in \mathcal{K}^* \end{aligned} \tag{7.8}$$

هي

$$\begin{aligned} & -q(x) \\ \text{subject to} \quad & Ax - s = -c \\ & x \in \mathcal{K},\ s \in \mathcal{K}^* \end{aligned}$$

وإذا كانت (7.8) لها حل مسموح به فعلياً، فإن القيمة المثلى لـ $q(x)$ هي صفر.

البرهان:

إن مسألة لاجرانج المرافقة لـ(7.8) هي

$$\begin{aligned} L(x,s,y) &= c^T x + y^T (Ax - s + c) \\ &= (A^T y + c)^T x - y^T s + y^T c \end{aligned}$$

وثنائية لاجرانج للمسألة (7.8) معرفة كالتالي:

$$\underset{y \in \mathbb{R}^k}{\text{maximize}} \quad c^T y + \underset{x \in \mathcal{K},\ s \in \mathcal{K}^*}{\text{minimize}} \ \{(A^T y + c)^T x - y^T s\}$$

إن مسألة التصغير الداخلية تكون محدودة من أسفل فقط، إذا كانت $A^T y + c \in K^*$ وكانت $-y \in K$ ففي هذه الحالة يكون الحل الأمثل لمسألة التصغير الداخلية هو صفر. وبالتالي نستطيع تبسيط ثنائية لاجرانج كالتالي:

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & A^T y + c \in \mathcal{K}^* \\ & -y \in \mathcal{K} \end{array}$$

باستخدام $A^T = -A$ ، لأنها متماثلة تخالفياً وبالتعويض في المتغير الجديد $\upsilon = A^T y + c = Au + c, \quad u = -y$ سنحصل على:

$$\begin{array}{ll} \text{maximize} & -c^T u \\ \text{subject to} & Au - \upsilon = -c \\ & u \in \mathcal{K}, \ \upsilon \in \mathcal{K}^* \end{array}$$

وبإخراج السالب نحصل على المطلوب. □

الآن لدينا النتيجة التالية:

نتيجة 2.8

مسألة الطمر (6.8) ذاتية الثنائية.

البرهان:

يتم البرهان باستخدام النظرية 1.8 وبناء المصفوفة المتماثلة تخالفياً من المعادلة (7.8). □

إن كون مسألة الطمر (6.8) ذاتية ثنائية ولها حلول مسموح بها فعلياً تقتضي أن الفجوة الثنائية تساوي $2\ \theta\beta$. ومن السهل التحقق من ذلك

$$\theta\beta = \operatorname{tr}(XS) + \tau\rho + \theta\upsilon \,.$$

وهذا يوضح أن الحل الأمثل لابد أن يحقق شروط التمام التالية:

$$\left.\begin{aligned} XS &= 0 \\ \rho\tau &= 0 \\ \theta\upsilon &= 0 \end{aligned}\right\} . \qquad (8.8)$$

الآن سوف نتطرق إلى حل مسألة الطمر، من الممكن حل مسألة الطمر بواسطة أي من طرق النقطة الداخلية والتي تتبع المسار الأوسط. ولكن في البداية لابد من إعادة صياغة الرموز للمسار الأوسط لمسألة الطمر، وذلك لأن مسألة الطمر ليست على الصورة القياسية، فنبدأ بإرخاء relax شروط أمثلة التمام (8.8) على الشكل التالي:

$$\begin{aligned} XS &= \mu I \\ \rho\tau &= \mu \\ \theta\upsilon &= \mu \end{aligned}$$

إذا عرّفنا المتغيرات الجديدة

$$\hat{X} = \begin{bmatrix} X & & \\ & \tau & \\ & & \upsilon \end{bmatrix}, \quad \hat{S} = \begin{bmatrix} S & & \\ & \rho & \\ & & \theta \end{bmatrix}$$

فإن المسار الأوسط يعرَّف بشكل وحيد على النحو التالي: $\hat{X}\hat{S} = \mu I, \quad \mu > 0$ تحت القيود (6.8)، ونرمز له بالرمز $(\hat{X}(\mu), \hat{S}(\mu))$ لكل $\mu > 0$.

4.8 الخوارزمية الشكلية ودالة الأوسطية

Frame Algorithm and Centrality Function

نفترض في بقية هذا الباب والباب التاسع أن المسألتين الأساسية والثنائية لهما حلول مسموح بها فعلياً. وكذلك نرمز لـ $(X(\mu), y(\mu), S(\mu))$ بالحل الوحيد لنظام شروط المسار الأوسط

$$\begin{aligned} &\operatorname{tr}(A_i X) = b_i, \qquad i = 1, \ldots, m \\ &\sum_{i=1}^{m} y_i A_i + S = C, \quad S \succ 0 \\ &XS = \mu I \end{aligned}$$

وتذكر أن وجود وحدانية الحل يتحقق من كون $(X(\mu), y(\mu), S(\mu))$ هي النقطة الصغرى الوحيدة للدالة المحدبة الأساسية الثنائية فعلياً

$$f_{\mu}(X,S) = \frac{1}{\mu}\ \mathrm{tr}\ (XS) - \log\ \det(XS) - n$$

والمعرَّفة داخل $\mathcal{P} \times \mathcal{D}$. إن دالة الحاجز الأساسية الثنائية هي عبارة عن الفرق بين دالة الحاجز الأساسية ودالة الحاجز الثنائية سالب الثابت n. والمعرفة كالتالي:

$$p_{\mu}(X) = \frac{1}{\mu}\ \mathrm{tr}\ (CX) - \log\ \det X$$

و

$$d_{\mu}(y,S) = -\frac{1}{\mu} b^{T} y - \log\ \det S$$

إن المسار الأوسط الأساسي يقابل تصغير $X(\mu)$ لـ $p_{\mu}(X)$. ولهذا يقال للوسيط μ أنه وسيط الأوسطية centering parameter، أو وسيط الحاجز barrier parameter.

سوف نشرح الآن خوارزمية الخطوة القصيرة short step algorithm والتي تتبع المسار الأوسط للمسألة الأساسية، واتجاه البحث ΔX هو عبارة عن إسقاط اتجاه نيوتن لحاجز المسألة الأساسية، ومن الممكن استنتاج إسقاط اتجاه نيوتن بتصغير تقريب تايلور التربيعي لدالة $p_{\mu}(X)$ تحت الشرط $\Delta X \in \mathcal{L}^{\perp}$ للاتجاهات المسموح بها للمسألة الأساسية. أو بشكل آخر ΔX هي حل مسألة التصغير التالية:

$$\begin{aligned} \Delta X = \arg\ \ \text{minimize} \quad & \nabla p_{\mu}(X)^{T}\ \Delta X + \frac{1}{2}\Delta X^{T} \nabla^{2} P_{\mu}(X) \Delta X \\ \text{subject to} \quad & tr\ (A_i \Delta X) = 0, \quad i = 1, \ldots, m \end{aligned}$$

حيث arg ترمز للنقطة التي تكون الدالة فيها أصغر ما يمكن.

الآن نستطيع كتابة خوارزمية الخطوة الصغيرة للمسألة الأساسية.

خوارزمية 3.8

مدخلات X° , μ بحيث أن X° حل مسموح بها فعلياً وقريبة من المسار الأوسط بشكل كافٍ.

الوسائط

وسيط الدقة $\varepsilon > 0$

وسيط التحديث $\theta = \dfrac{1}{4\sqrt{n}+2}$

إبدأ $X = X^{\circ}$; $\mu = \mu^{\circ}$

بينما $n\mu > \varepsilon$

$X = X + \Delta X$

$\mu = (1-\theta)\mu$

نهاية

نهاية

لتكن X داخل $\mathcal{P}$ ومعطى الوسيط $\mu > 0$ نعرف

$$(S(X,\mu), y(X,\mu)) = \arg \ \ \text{minimize} \left\| \frac{1}{\mu} X^{\frac{1}{2}} S X^{\frac{1}{2}} - I \right\|$$

$$\text{subject to } \sum_{i=1}^{m} y_i A_i + S = C$$

$$S \in \mathcal{S}_n, \ \ y \in \mathbb{R}^m$$

$S(X, \mu)$ تحقق القيود الثنائية المسموح بها دون أن نشترط أنها موجبة شبه معرفة. نعرِّف الآن دالة المسار الأوسط

$$\delta_p(X,\mu) = \left\| \frac{X^{\frac{1}{2}} S(X,\mu) X^{\frac{1}{2}}}{\mu} - I \right\|$$

لاحظ أن

$$.\delta_p(X,\mu)=0 \Leftrightarrow X=X(\mu)$$

المصفوفة $S(X,\mu)$ تلعب دوراً أساسياً في تحليل الخوارزمية. وبالخصوص اتجاه البحث حيث يمكن أن يكتب بدلالته كما سنوضح ذلك في الفصل التالي.

5.8 إسقاط اتجاه نيوتن لدالة الحاجز الأساسية
Projected Newton Direction for Primal Barrier Function

تذكر أن إسقاط اتجاه نيوتن لدالة الحاجز الأساسية

$$p_\mu(X)=\frac{1}{\mu}\ \text{tr}\,(CX)-\log\ \det X$$

عند الزوج (X,μ) معرف بـ

$$\begin{aligned}\Delta X &= \arg\ \ \text{minimize}\ \ \nabla p_\mu{}^T \Delta X+\frac{1}{2}\Delta X^T\ \nabla^2 p_\mu \Delta X\\ &= \arg\ \ \text{minimize}\ \ \text{tr}\,(\nabla p_\mu \Delta X)+\frac{1}{2}\text{tr}\,(\nabla^2 p_\mu \Delta X^2)\end{aligned} \tag{9.8}$$

تحت شروط وجود الحلول المسموح بها

$$\text{tr}\,(A_i\Delta X)=0\quad,\quad i=1,..,m$$

حيث $\nabla p_\mu=\nabla p_\mu(X,\mu)$ تمثل الاشتقاق gradient و $\nabla^2 p_\mu$ تمثل مصفوفة هاس Hessian . وبمعنى آخر إسقاط اتجاه نيوتن يُصغر تقريب تايلور التربيعي لدالة p_μ تحت شرط وجود اتجاه مسموح به. وسوف نرمز له عند X بـ ΔX .

نتيجة 4.8

لتكن $f: ri\ (S_n^+)\to\mathbb{R}$ بحيث

$$f(x)=\log\ \det\ \text{X}$$

فإن $\nabla f(x)=X^{-1}$

البرهان: انظر [VB].

نتيجة 5.8

لتكن $f : ri\,(S_n^+) \to \mathbb{R}$ بحيث

$$f(x) = \mathrm{tr}\,(CX)$$

حيث $C \in \mathcal{S}_n$ فإن $\nabla f\,(x) = C$.

البرهان: مباشر.

نتيجة 6.8

لتكن $f : ri\,(S_n^+) \to \mathbb{R}$ بحيث

$$f\,(x) = \log\ \det\ \mathrm{X}$$

فإن

$$\nabla^2 f(x) H = -X^{-1} H X^{-1}, \qquad \forall H \in \mathcal{S}_n$$

هو عبارة عن مؤثر خطي linear operator لكل مصفوفة X قابلة للعكس.

البرهان: انظر [VB].

نظرية 7.8

الدالة $f : ri\,(S_n^+) \to \mathbb{R}$ بحيث

$$f\,(x) = -\log\ \det\ (\mathrm{X})$$

هي دالة محدبة فعلياً.

البرهان: انظر [HJ].

لدينا من نتيجة 4.8

$$\nabla p_\mu(X) = \frac{1}{\mu} C - X^{-1}$$ ،

وكذلك $\nabla^2 p_\mu(X,\mu):\mathcal{S}_n \to \mathcal{S}_n$ هي عبارة عن مؤثر خطي يحقق

$$.\nabla^2 p_\mu(X)\Delta X = X^{-1}\Delta X\ X^{-1}, \qquad \forall \Delta X \in \mathcal{S}_n$$

بتعويض المشتقة ومصفوفة هاس في (9.8) تحصل على

$$\begin{aligned} \Delta X = \arg\ \ \text{minimize}\ \ & \frac{1}{\mu}\operatorname{tr}(C\Delta X) - \operatorname{tr}(X^{-1}\Delta X) + \frac{1}{2}(\operatorname{tr}(X^{-1}\Delta X)^2) \\ \text{subject to}\ \ & \operatorname{tr}(A_i\Delta X) = 0, \qquad i = 1,\dots,m \end{aligned}$$

إن شروط الأمثلة KKT لهذه المسألة هي

$$\begin{aligned} \frac{1}{\mu}C - X^{-1} + X^{-1}\Delta X\ X^{-1} + \sum_{i=1}^{m} y_i A_i = 0 \\ \operatorname{tr}(A_i\Delta X) = 0, \quad i = 1,\dots,m \end{aligned}$$

وبالتلاعب المباشر بهذه الشروط نحصل على

$$\operatorname{vec}(X^{-\frac{1}{2}}\Delta X\ X^{-\frac{1}{2}}) = -\left[I - \mathcal{A}_X{}^T\left(\mathcal{A}_X\mathcal{A}_X{}^T\right)^{-1}\mathcal{A}_X \right]\operatorname{vec}(\mathrm{Z}) \qquad (10.8)$$

حيث $Z = \frac{1}{\mu}X^{\frac{1}{2}}CX^{\frac{1}{2}} - I$ و vec تعني تحويل المصفوفة إلى متجه يضع جميع الأعمدة فوق بعض كمتجه واحد، و $\mathcal{A}_X$ هي عبارة عن المصفوفة $m \times n^2$ بحيث أن صفوفها هي:

$$.\operatorname{vec}(X^{\frac{1}{2}}A_j\ X^{\frac{1}{2}})^T, \qquad j = 1,\dots,m$$

إن المعادلة (10.8) هي عبارة عن إسقاط عمودي للمتجه

$$\operatorname{vec}\left(\frac{1}{\mu}X^{\frac{1}{2}}CX^{\frac{1}{2}} - I\right)$$

على الفضاء الصفري null-space لـ $\mathcal{A}_X{}^2$. لاحظ أن فضاء الصف لـ $\mathcal{A}_X$ معطى بـ

$$\text{، } \operatorname{span}\{\operatorname{vec}(X^{\frac{1}{2}}A_1\ X^{\frac{1}{2}}),\dots\dots,\operatorname{vec}(X^{\frac{1}{2}}A_m\ X^{\frac{1}{2}})\}$$

والفضاء الصفري متمم عمودي على هذا الفضاء.

وبالرجوع إلى فضاء المصفوفات المتماثلة $\mathcal{S}_n$ من الواضح استنتاج اتجاه ΔX عن طريق إسقاط المصفوفة $\frac{1}{\mu}X^{\frac{1}{2}}CX^{\frac{1}{2}}-I$ على المتممة العمودية لـ

$$\text{span}\,\{X^{\frac{1}{2}}A_1\,X^{\frac{1}{2}},......,\,X^{\frac{1}{2}}A_m\,X^{\frac{1}{2}}\}.$$

مؤثر إسقاطي وثيق الصلة بالإسقاط السابق هو $P_{\mathcal{A}_X}:\mathcal{S}_n\to\mathcal{S}_n$ المعطى بالشكل التالي:

$$\begin{aligned} P_{\mathcal{A}_X}(M)=\text{arg minimize}\ \ &\|W-M\| \\ \text{subject to}\ \ &\text{tr}\,(\mathrm{X}^{\frac{1}{2}}A_i\ \mathrm{X}^{\frac{1}{2}}\ W)=0,\quad i=1,\dots,m \\ &W\in\mathcal{S}_n \end{aligned} \qquad (11.8)$$

نستطيع الآن كتابة اتجاه البحث $\Delta\mathrm{X}$ بدلالة $\mathrm{S}(\mathrm{X},\mu)$.

نظرية 8.8

اسقاط اتجاه نيوتن عند $\mathrm{X}\in ri\,(\mathcal{P})$ له الشكلين التاليين:

$$\Delta X=-X^{\frac{1}{2}}\left(P_{\mathcal{A}_X}\left(\frac{X^{\frac{1}{2}}\ C\ X^{\frac{1}{2}}}{\mu}-I\right)\right)X^{\frac{1}{2}}=-\left(\frac{XS\left(X,\mu\right)X}{\mu}-X\right)$$

حيث شروط الأمثلة KKT للمسألة

$$\begin{aligned} \text{minimize}\ \ &\left\|\frac{X^{\frac{1}{2}}\ S\ X^{\frac{1}{2}}}{\mu}-I\right\| \\ \text{subject to}\ \ &\sum_{\mathrm{i}=1}^{\mathrm{m}}y_{\mathrm{i}}A_i+S=C \\ &y\in\mathbb{R}^m,\ \ S\in\mathcal{S}_n \end{aligned}$$

هي

$$\left.\begin{array}{l} \dfrac{X\ S\ X}{\mu^2} - Q = \dfrac{X}{\mu} \\ \operatorname{tr}\ (A_i Q) = 0, \qquad i = 1,\ldots,m \\ \sum_{i=1}^{m} y_i A_i + S = C \end{array}\right\} \qquad (12.8)$$

حيث $Q \in \mathcal{S}_n$

البرهان: انظر [Dk].

من الممكن حل شروط الأمثلة (12.8) بإعادة كتابتها على الشكل التالي:

$$\sum_{i=1}^{m} y_i \operatorname{tr}\ (XA_i XA_j) = \operatorname{tr}\ (XA_j XC) - \mu \operatorname{tr}\ (A_j X), \quad j = 1,\ldots,m \qquad (13.8)$$

إن حل هذا النظام الخطي $m \times m$ يعطينا الحل $y(X,\mu)$ ، ومصفوفة المعاملات $\operatorname{tr}\ (XA_i XA_j)$ للنظام الخطي (13.8) متماثلة موجبة شبه معرفة لأن المصفوفات

$$A_1, A_2, \ldots, A_m$$

مستقلة خطياً. وبجعل

$$S(X,\mu) = \sum_{i=1}^{m} y_i(X,\mu) A_i - C$$

نستطيع حساب اتجاه البحث بواسطة

$$\Delta X = -\frac{1}{\mu} X\ \ S(X,\mu) X + X .$$

6.8 اتجاه الموازنة التآلفي Affine – Scaling Direction

نظرية 8.8 تبين لنا أننا نستطيع تقسيم اتجاه البحث ΔX إلى حدين

$$\Delta X = \frac{1}{\mu} \Delta X_a + \Delta X_c$$

حيث

$$\Delta X_a = -X^{\frac{1}{2}}(P_{\mathcal{A}_X}(X^{\frac{1}{2}}CX^{\frac{1}{2}}))X^{\frac{1}{2}}$$

و

$$.\ \Delta X_c = X^{\frac{1}{2}}(P_{\mathcal{A}_X}(I))X^{\frac{1}{2}}$$

يسمى الحد الأول ΔX_a معاملات التوازن التآلفي affine scaling، ويسمى الحد الثاني ΔX_c معاملات اتجاه البحث الأوسطية centering. لاحظ أن معاملات التوازن التآلفي ΔX_a لاتجاه البحث ΔX يصبح هو المسيطر عندما تكون μ صغيرة. تذكر أننا نحاول حساب $X(\mu)$ على المسار الأوسط الأمثل. إن دور الاتجاه ΔX_a هو الحصول على أكبر قدر من التصغير في دالة الهدف في دورة واحدة دون محاولة الاقتراب من المسار الأوسط.

هذا التمثيل الهندسي سيتضح أكثر بواسطة صيغ مختلفة لـ ΔX_a.

بواسطة تعريف الإسقاط $P_{\mathcal{A}_X}$ في (11.8) نستطيع كتابة تعريف ΔX_a على الشكل التالي:

$$\begin{aligned} X^{-\frac{1}{2}}\Delta X_a X^{-\frac{1}{2}} &= \arg\ \text{minimize} \left\| W - X^{\frac{1}{2}}CX^{\frac{1}{2}} \right\| \\ &\text{subject to}\ \ \text{tr}\,(X^{\frac{1}{2}}A_iX^{\frac{1}{2}}W) = 0, \quad i = 1,\dots,m \qquad (14.8) \\ &W \in \mathcal{S}_n \end{aligned}$$

من الممكن تعريف اتجاه التوازن التآلفي بطريقة مختلفة كالتالي:

$$\begin{aligned} \Delta X_a &= \arg\ \text{minimize tr}\ (C\Delta X) \\ &\text{subject to}\ \left\| X^{-\frac{1}{2}}\Delta X X^{-\frac{1}{2}} \right\|^2 \le 1 \qquad (15.8) \\ &\text{tr}\,(A_i\Delta X) = 0, \quad i = 1,\dots,m \end{aligned}$$

من السهل التحقق أن هذين التعريفين متكافأن بمقارنة شروط الأمثلة لمسألتي التصغير (14.8) و(15.8).

وبالتالي نحصل على $X+\Delta X_a \in \mathcal{P}$ ، كما هو واضح من النتيجة التالية:

نتيجة 9.8

ليكن لدينا $X \in \mathcal{P}$. إذا كانت ΔX اتجاهاً مسموحاً به بمعنى

$$\operatorname{tr}(A_i \Delta X)=0, \quad i=1,\dots,m$$

وكذلك كانت ΔX تنتمي إلى المجموعة $\{\Delta X : \| X^{-\frac{1}{2}}\ \Delta X\ X^{-\frac{1}{2}} \|^2 \le 1\}$ فإن $X+\Delta X \in \mathcal{P}$.

البرهان:

الشرط

$$\left\| X^{-\frac{1}{2}} \Delta X X^{-\frac{1}{2}} \right\|^2 = \sum_{i=1}^{n} \lambda_i (X^{-\frac{1}{2}} \Delta X X^{-\frac{1}{2}})^2 \le 1$$

يقتضي أن

$$\left| \lambda_i (X^{-\frac{1}{2}} \Delta X X^{-\frac{1}{2}}) \right| \le 1, \qquad i=1,\dots,n$$

والذي يوضح لنا

$$I + X^{-\frac{1}{2}} \Delta X X^{-\frac{1}{2}} \succeq 0$$

بالضرب من الجهتين بـ $X^{\frac{1}{2}}$ نحصل على المطلوب، أي أن $X+\Delta X \succeq 0$. □

لكي نحصل على خوارزمية متقاربة لابد من إضافة المعامل الأوسطي ΔX_c .

7.8 السلوك قرب المسار الأوسط Behavior Near the Central Path

ليكن لدينا معطى $\mu > 0$ ونعلم أن $X \in ri(\mathcal{P})$ بحيث أن $\delta_p(X,\mu) < 1$. ونريد أن نعرف تأثير خطوة كاملة لإسقاط نيوتن

$$X^+ = X + \Delta X = 2X - \frac{1}{\mu} X S(X,\mu) X .$$

إن الـزوج $\left(X^{+}, S(X,\mu)\right)$ يحقـق قيـود المسـاواة للمسـألة الأساسـية والثنائية ولكن قد لا يحقق شرط الإيجاب شبه المعرف. والنتيجتان التاليتان توضح لنا أن شرط الإيجاب شبه المعرف متحقق إذا كانت X أوسطية بشكل كافٍ بالنسبة لـ μ.

نتيجة 10.8

إذا كانت $X \succ 0$ وكانت $\delta_p(X,\mu) < 1$، فإن $S(X,\mu) \succ 0$

البرهان:

من تعريف $\delta_p(X,\mu)$ نحصل على

$$\delta_p(X,\mu)^2 = \left\| \frac{X^{\frac{1}{2}}\ S(X,\mu) X^{\frac{1}{2}}}{\mu} - I \right\|^2$$

$$= \operatorname{tr}\left(\left(\tfrac{1}{\mu}\ X^{\frac{1}{2}}\ S(X,\mu)\ X^{\frac{1}{2}} - I \right)^2 \right)$$

$$= \sum_{i=1}^{n} \left(\tfrac{1}{\mu} \lambda_i \left(X^{\frac{1}{2}}\ S(X,\mu)\ X^{\frac{1}{2}} \right) - 1 \right)^2$$

وباستخدام $\delta_p(X,\mu) < 1$ نحصل على

$$\sum_{i=1}^{n} \left(\tfrac{1}{\mu} \lambda_i \left(X^{\frac{1}{2}}\ S(X,\mu)\ X^{\frac{1}{2}} \right) - 1 \right)^2 < 1$$

وهـذا يوضـح أن $\lambda_i \left(X^{\frac{1}{2}} S(X,\mu)\ X^{\frac{1}{2}} \right) > 0$ لكـل $i = 1, \ldots\ldots, n$. أي أن $S(X,\mu) > 0$. □

هـذه النتيجـة تـبين لنـا أنـه رغـم اسـتخدامنا للحـل المسـموح بـه الأساسـي $X \in ri(\mathcal{P})$ فإنـه مـن خـلال تنفيـذ الخوارزميـة نحصـل علـى حـل مسـموح بـه للمسـألة الثنائيـة $S(X,\mu) \in ri(\mathcal{D})$ كمكسـب عنـدما تكـون $\delta_p(X,\mu) < 1$. وهذا يعطينا حداً أعلى

$$\operatorname{tr}(CX) - p^{*} \le \operatorname{tr}(XS(X,\mu))$$

للتفريق بين قيمة دالة الهدف عند الدورة الحالية وبين القيمة المثلى بواسطة نظرية الثنائية الضعيف.

الآن سوف نوضح أن $X^{+} = X + \Delta X$ حل مسموح به، إذا كانت X أوسطية بشكل كافٍ.

نتيجة 11.8

لتكن

$$X^{+} = X + \Delta X = 2X - \frac{1}{\mu} X \; S(X, \mu) X$$

إذا كانت $\delta_p(X, \mu) < 1$ و $X \succ 0$ فإن $X^{+} \succ 0$.

البرهان:

لاحظ أننا نستطيع كتابة X^{+} على الشكل التالي: أي أن

$$X^{+} = X^{\frac{1}{2}}(2I - X^{\frac{1}{2}} \frac{S(X, \mu)}{\mu} X^{\frac{1}{2}}) X^{\frac{1}{2}} \tag{16.8}$$

لأن $\delta_p(X, \mu) < 1$ أي أن

$$\left\| \frac{1}{\mu} X^{\frac{1}{2}} S(X, \mu) X^{\frac{1}{2}} - I \right\| < 1$$

ينتج

$$\sum_{i=1}^{n} \lambda_i^2 \left(\frac{X^{\frac{1}{2}} S(X, \mu) X^{\frac{1}{2}}}{\mu} - I \right) < 1$$

وبالتالي لدينا

$$\lambda_i \left(\frac{1}{\mu} X^{\frac{1}{2}} S(X, \mu) X^{\frac{1}{2}} \right) \in (0, 2), \quad i = 1, \dots, n$$

والذي يقتضي

$$\lambda_i \left(2I - X^{\frac{1}{2}} \frac{S(X, \mu)}{\mu} X^{\frac{1}{2}} \right) \in (0, 2), \quad i = 1, \dots, n$$

وكنتيجة لذلك نحصل على $X^{+} \succ 0$ من (16.8). □

أيضاً نستطيع الحصول على تقارب تربيعي على الدورة الأساسية المؤدية إلى المسار الأوسط.

نتيجة 12.8

إذا كانت $X \in ri(\mathcal{P})$ و $\delta_p(X,\mu) < 1$ فإن

$$X^{+} = X + \Delta X = 2X - \frac{1}{\mu} X \; S(X,\mu) X$$

تحقق $\delta_p(X^+,\mu) \le \delta_p^2(X,\mu)$.

البرهان:

بتعريف $\delta_p(X,\mu) < 1$ نحصل على

$$\delta_p(X^+,\mu)^2 = \left\| \frac{X^{+\frac{1}{2}} \, S\left(X^+,\mu\right) X^{+\frac{1}{2}}}{\mu} - I \right\|^2 \le \left\| \frac{X^{+\frac{1}{2}} S\left(X,\mu\right) X^{+\frac{1}{2}}}{\mu} - I \right\|^2$$

$$= \operatorname{tr}\left(\left(\tfrac{1}{\mu} S(X,\mu) X^+ - I \right)^2 \right)$$

وبالتعويض عن $X^+ = 2X - \frac{1}{\mu} XS\left(X,\mu\right) X$ ينتج

$$\delta_p(X^+,\mu)^2 \le \operatorname{tr}\left(\left(\tfrac{1}{\mu} S(X,\mu) \left[2X - \tfrac{1}{\mu} XS(X,\mu) X \right] - I \right)^2 \right)$$

$$= \operatorname{tr}\left(\left(\tfrac{1}{\mu} S(X,\mu) X - I \right)^4 \right)$$

$$= \operatorname{tr}\left(\left(\tfrac{1}{\mu} X^{\frac{1}{2}} S(X,\mu) X^{\frac{1}{2}} - I \right)^4 \right)$$

$$= \left\| \left(\tfrac{1}{\mu} X^{\frac{1}{2}} S(X,\mu) X^{\frac{1}{2}} - I \right)^2 \right\|^2$$

$$\le \left\| \tfrac{1}{\mu} X^{\frac{1}{2}} S(X,\mu) X^{\frac{1}{2}} - I \right\|^4 = \delta_p^4(X,\mu)$$

حيث المتباينة الثانية ناتجة من خواص معيار فروبينيس Frobenius norm . □

8.8 تحديث وسيط المسار الأوسط Updating the Centering Parameter

عندما تكون دورة المسألة الأساسية X أوسطية بشكل كافٍ، أي أن $\delta_p(X,\mu)\le\tau$ لتَحَمُّل ما τ فإننا نستطيع تخفيض الوسيط μ. نُحدث وسيط الحاجز بطريقة ما بحيث يبقى $\delta_p(X,\mu^+)\le\frac{1}{2}$ بعد تحديث μ بـ μ^+. الخطوة التالية $X+\Delta X$ تولد X^+ حلاً مسموحاً به يحقق $\delta_p(X^+,\mu^+)\le\frac{1}{4}$ بواسطة **نتيجة** 12.8.

نتيجة 13.8

نعرف تحديث لـ μ بـ $\mu^+=(1-\theta)\mu$ بحيث $0<\theta<1$ وسيطاً معطى. ومن هذا نستنتج

$$\delta_p(X,\mu^+)\le\frac{1}{1-\theta}\left(\delta_p(X,\mu)+\theta\sqrt{n}\right)$$

البرهان:

$$\begin{aligned}
\delta_p(X,\mu^+) &= \left\|\frac{X^{\frac{1}{2}}S(X,\mu^+)X^{\frac{1}{2}}}{(1-\theta)\mu}-I\right\| \\
&\le \left\|\frac{X^{\frac{1}{2}}S(X,\mu)X^{\frac{1}{2}}}{(1-\theta)\mu}-I\right\| \\
&= \left\|\frac{X^{\frac{1}{2}}S(X,\mu)X^{\frac{1}{2}}}{(1-\theta)\mu}-\frac{1}{1-\theta}I+\frac{\theta}{1-\theta}I\right\| \\
&\le \frac{1}{1-\theta}\left(\left\|\frac{X^{\frac{1}{2}}S(X,\mu)X^{\frac{1}{2}}}{\mu}-I\right\|+\theta\|I\|\right) \\
&= \frac{1}{1-\theta}\left(\delta_p(X,\mu)+\theta\sqrt{n}\right)
\end{aligned}$$

حيث المتباينة الثانية تتحقق من المتباينة المثلثية triangular inequality. □

إن النتيجة السابقة تمكننا أن نختار تحديث للوسيط θ والذي يضمن أن تبقى الدورة الأساسية أوسطية بشكل كافٍ بالنسبة للوسيط الجديد $\mu^+ = (1-\theta)\mu$.

نتيجة 14.8

لتكن $\delta_p(X,\mu) \le \frac{1}{2}$ و $\theta = 1/(4\sqrt{n}+2)$. بعد خطوة $X^+ = X + \Delta X$ والتحديث $\mu^+ = (1-\theta)\mu$ نحصل على $\delta_p(X^+,\mu^+) \le \frac{1}{2}$.

البرهان:

باستخدام نتيجة 13.8 ونتيجة 12.8 نحصل على

$$\begin{aligned}\delta_p(X^+,\mu^+) &\le \frac{1}{1-\theta}\left(\delta_p\left(X^+,\mu\right)+\theta\sqrt{n}\right)\\ &\le \frac{1}{1-\theta}\left(\delta_p^2\left(X,\mu\right)+\theta\sqrt{n}\right).\end{aligned}$$

وبالتعويض عن $\theta = 1\big/(4\sqrt{n}+2)$ نحصل على

$$\begin{aligned}\delta_p\left(X^+,\mu^+\right) &\le \frac{4\sqrt{n}+2}{4\sqrt{n}+1}\left(\frac{1}{4}+\frac{\sqrt{n}}{4\sqrt{n}+2}\right)\\ &= \frac{1}{2}\end{aligned}$$

□

من السهل التحقق من أنه إذا كانت $\delta_p(X,\mu) \le \frac{1}{2}$ فإن التحديث الديناميكي

$$\theta = \frac{\frac{1}{2}-\delta_p(X,\mu)}{\sqrt{n}+\frac{1}{2}} \ge \frac{1}{4\sqrt{n}+2}$$

يضمن أن $\delta_p(X,\mu^+) \le \frac{1}{2}$ إذا كانت $\mu^+ = (1-\theta)\mu$. السؤال الآن هل من الممكن أن نحصل على قيمة صغرى لـ μ^+ بحيث أن الشرط $\delta_p(X,\mu^+) \le \frac{1}{2}$ يبقى متحقق؟ وهذا في الحقيقة ممكن، وذلك بإعادة كتابة $\delta_p(X,\mu)$ على الشكل التالي:

$$\delta_p(X,\mu) = \left\| X^{-\frac{1}{2}}\left(\Delta X_c + \tfrac{1}{\mu}\Delta X_a\right)X^{-\frac{1}{2}}\right\|$$

من تعريف δ_p ونظرية 8.8.

لنرمز $D_a = X^{-\frac{1}{2}}\Delta X_a X^{-\frac{1}{2}}$ و $D_c = X^{-\frac{1}{2}}\Delta X_c X^{-\frac{1}{2}}$ نرى أن أصغر قيمة لـ μ^+ والتي لا زالت تحقق $\delta_p(X,\mu^+) \le \frac{1}{2}$ هي أصغر جذر موجب للمعادلة

$$\delta_p(X,\mu) = \left\| D_c + \frac{1}{\mu} D_a \right\| = \frac{1}{2} .$$

وبتربيع طرفي المعادلة نحصل على المعادلة التربيعية quadratic equation لـ $\frac{1}{\mu}$ التالية:

$$\frac{1}{\mu^2}\|D_a\|^2 + \frac{2}{\mu}\ \mathrm{tr}\ (D_a D_c) + \|D_c\|^2 - \frac{1}{4} = 0$$

والذي من الممكن حلها للحصول على μ^+ المطلوبة.

النظرية التالية تعطينا حداً لأسوإ الأحوال تعقيداً للخوارزمية الشكلية.

نظرية 15.8

لتكن $\varepsilon > 0$ وسيط دقة، $\mu^\circ > 0$ و $\theta = 1/(4\sqrt{n}+2)$. لتكن $X^\circ \succ 0$ نقطة بداية مسموح بها فعلياً بحيث أن $\delta_p(X^\circ,\mu^\circ) \le \frac{1}{2}$. إن الخوارزمية تنتهي بعد $\left\lceil 6\sqrt{n}\ \log\frac{n\mu^\circ}{\varepsilon} \right\rceil$ خطوة على الأكثر، وتكون المصفوفات الأخيرة الناتجة X و $S(X,\mu)$ حلولاً مسموحاً بها فعلياً، والفجوة الثنائية محدودة بـ $\mathrm{tr}\ (XS(X,\mu)) \le \frac{3}{2}\varepsilon$.

البرهان: انظر [Dk].

9.8 خوارزمية الثنائية Dual Algorithm

إن خوارزمية الثنائية هي مشابهه تماماً لخوارزمية المسألة الأساسية، إذا عرفنا

$$X(S,\mu)=\arg\ \text{minimize}\ \left\|\frac{S^{\frac{1}{2}}XS^{\frac{1}{2}}}{\mu}-I\right\|$$

$$\text{subject to}\quad \text{tr}\,(A_iX)=b_i\,,\qquad i=1,....,m$$

$$X\in\mathcal{S}_n$$

وذلك للمتغير المسموح به الفعلي الثنائي $S\succ 0$ ، وبالتالي فإن المرتبة الأولى لشروط الأمثلة والذي يعطينا $X\left(S,\mu\right)$ هي:

$$S\left[\frac{XS}{\mu}-I\right]-\sum_{i=1}^{m}\Delta y_iA_i=0 \qquad (17.8)$$

$$\text{tr}\,(A_iX)=b_i,\quad i=1,....,m \qquad (18.8)$$

وبالضرب من اليمين واليسار للمعادلة (17.8) بـ S^{-1} ، ثم باستخدام المعادلة (18.8). نحصل على:

$$\sum_{i=1}^{m}\Delta y_i\ \text{tr}\left(A_iS^{-1}A_jS^{-1}\right)=\frac{-1}{\mu}b_j+\text{tr}\,(A_jS^{-1}),\quad j=1,....,m \qquad (19.8)$$

وإذا عرفنا

$$\delta_d(S,\mu)=\left\|\frac{S^{\frac{1}{2}}X(S,\mu)S^{\frac{1}{2}}}{\mu}-I\right\|\text{،}$$

فإننا نستطيع إعادة التحليل الذي قمنا به بالنسبة للخوارزمية الأساسية، ولكن بتبديل دور S مع X. إن اتجاه إسقاط نيوتن لدالة الحاجز الثنائية d_μ (اتجاه البحث للخوارزمية) يكون

$$\Delta S=S^{\frac{1}{2}}\left(I-\frac{1}{\mu}S^{\frac{1}{2}}X\left(S,\mu\right)S^{\frac{1}{2}}\right)S^{\frac{1}{2}}=\sum_{i=1}^{m}\Delta y_iA_i\text{،} \qquad (20.8)$$

حيث نحصل على Δy بحل (19.8)، وبالتالي نحصل على ΔS من $\Delta S=-\sum_{i=1}^{m}\Delta y_iA_i$.

الآن نستطيع كتابة خوارزمية الخطوة الصغيرة للمسألة الثنائية.

خوارزمية 16.8

مدخلات زوج من الحلول المسموح بها الفعلية الثنائية (S°, y°)

الوسائط الوسيط μ° بحيث أن $S_d(S^{\circ}, \mu^{\circ}) \le \frac{1}{2}$

وسيط الدقة $\varepsilon > 0$

وسيط التحديث $\theta = 1/(4\sqrt{n}+2)$

ابدأ $S = S^{\circ}, \mu = \mu^{\circ}$

بينما $n\mu > \varepsilon$

$S = 2S - \frac{1}{\mu} SX(S,\mu)S$

$\mu = (1-\theta)\mu$

نهاية

نهاية

نظراً للتشابه في تحليل خوارزمية المسألة الأساسية وخوارزمية المسألة الثنائية فإن المسألتين لهما نفس حد التعقيد المذكور في نظرية 15.8.

من المهم ملاحظة أنه ليس من الضروري صياغة $X(S,\mu)$ بشكل صريح لمعرفة ما إذا كانت موجبة شبه معرفة أو لا، وبالتالي حساب الفجوة الثنائية والتأكد من أنها موجبة، ومن (17.8) نعلم أن

$$X(S,\mu) \ge 0 \Leftrightarrow S + \sum_{i=1}^{m} \Delta y_i A_i \succeq 0 .$$

أيضاً لاحظ أنه إذا كانت $X(S,\mu) \succeq 0$ فإن الفجوة الثنائية عند $\left(X(S,\mu), S\right)$ معطاة بالمعادلة

$$\text{tr}\,(X(S,\mu)S) = \mu\ \text{tr}\,(S - \Delta S)$$

من (20.8). هذه الملاحظات مهمة عند الاستفادة من تناثر المعلومات.

10.8 طرق التحديث الكبيرة Large Update Methods

إن استراتيجية تحديث μ التي شرحناها في الفصول السابقة متحفظة جداً عند استخدامها في التطبيقات العملية. وسوف نشرح في هذا الفصل استراتيجيتين لتحديث μ، والتي تسمح لنا بتخفيض كبير لها. في الخوارزمية التالية سوف نستخدم التحديث الديناميكي لـ μ

$$\mu = \frac{\operatorname{tr}(XS)}{n+\upsilon\sqrt{n}}$$

حيث $S \in ri(\mathcal{D})$ في الدورة الثنائية الحالية، و $X \in \mathcal{P}$ أفضل حل معلوم أساسي، و $\upsilon \geq 1$ وسيط معطى. تسمى هذه الخوارزمية بخوارزمية الموازنة الثنائية:

خوارزمية 17.8

مدخلات زوج $\left(X^{\circ}, S^{\circ}\right) \in ri(\mathcal{P} \times \mathcal{D})$

الوسائط الوسيط $\mu_0 > 0$ بحيث أن $\delta_d(S^{\circ}, \mu_0) \leq \frac{1}{2}$

وسيط الدقة $\varepsilon > 0$

الوسيط $\upsilon \geq 1$

ابدأ $X = X^{\circ}, S = S^{\circ}, \mu = \mu_0$

بينما $\operatorname{tr}(XS) > \varepsilon$

$S = 2S - \frac{1}{\mu} SX(S,\mu) S$

إذا $X(S,\mu) \succ 0$ **فإن** $X = X(S,\mu)$

$$\mu = \frac{\operatorname{tr}(XS)}{n+\upsilon\sqrt{n}}$$

نهاية

نهاية

إن دور المصفوفة X في الخوارزمية هو للترميز، وليس لها دور في الحسابات ولا تحتاج إلى تخزينها.

خوارزمية الخطوة الصغيرة مددت بواسطة [AF] لتستخدم تحديثاً كبيراً لـ μ. الخوارزمية تعمل بخطوات إسقاط نيوتن بالنسبة لقيمة μ، حتى يتحقق $\delta_d(S,\mu) \le \frac{1}{2}$ للدورة الحالية. بعد ذلك μ تخفض بكسر ثابت، بحيث أن $\mu = (1-\theta)\mu$. من الممكن الآن كتابة خوارزمية الخطوة الطويلة للمسألة الثنائية على الشكل التالي:

خوارزمية 18.8

مدخلات زوج $\left(X^{\circ}, y^{\circ}\right) \in ri(\mathcal{D})$

الوسائط وسيط أوسطي $\tau = 1/\sqrt{2},\ \tau > 0$

وسيط μ° بحيث $\delta_d(S^{\circ}, \mu_{\circ}) \le \tau$

وسيط دقة $\varepsilon > 0$ ووسيط تحديث $0 < \theta < 1$

ابدأ $S = S^{\circ}\ , y = y^{\circ}\ , \mu = \mu_{\circ}$

بينما $\operatorname{tr}\left(XS\right) > \varepsilon$

إذاً $\delta_d(S, \mu) \le \tau$

$\mu = \left(1-\theta\right)\mu$

وإلا $\delta_d(S, \mu) > \tau$

احسب $\left(\Delta S, \Delta y\right)$ وأوجد α

$S = S + \alpha \Delta S$

$y = y + \alpha \Delta y$

نهاية

نهاية

نهاية

حيث $\alpha = \arg \ \text{minimize}\ \ d_{\mu}\left(y + \alpha\Delta y, S + \alpha \Delta S\right)$.

هذه الخوارزمية لها حد في أسوإ الحالات، بحيث إذا كانت $\theta = O(1)$ فإن حد الدورات هو $O(nL)$ ، وإذا كانت $\theta = O(1/\sqrt{n})$ فإن حد الدورات هو $O\left(\sqrt{n}L\right)$ انظر [RTV].

لدينا الآن حد التعقيد للخوارزمية 17.8.

نظرية 19.8

إن طريقة الموازنة الثنائية تتوقف بعد

$$O\left(\upsilon\sqrt{n}\ \log\left(\frac{\mathrm{tr}\left(X^{\circ}S^{\circ}\right)}{\varepsilon}\right)\right)$$

دورة. والناتج هو زوج $\left(X\left(S,\mu\right),S\right)\in\mathcal{P}\times\mathcal{D}$ ، بحيث $\mathrm{tr}\left(X\left(S,\mu\right)S\right)\leq\varepsilon$

البرهان: انظر [Ye].

تمارين الباب الثامن

1.8 لدينا البرنامج التالي

$$\text{maximize} \quad y_1 + y_2 \,, \qquad y \in \mathbb{R}^2$$
$$\text{subject to} \quad y_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + y_2 \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \preceq \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

أوجد البرنامج الثنائي؟ ثم أوجد مسألة الطمر المتجانس للمسألتين الأساسية والثنائية؟

2.8 لتكن $n = m = 3$، $b = [1 \quad 0 \quad 0]$ وأن:

$$C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$
$$A_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \; A_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

أوجد مسألة لاجرانج المرافقة؟

3.8 أوجد $p_\mu(X)$ و $f_\mu(X,S)$ و $d_\mu(y,S)$ لتمرين 2.8؟

4.8 حل البرنامج التالي:

$n = 2$ ، $m = 3$ ،

$b = [1 \quad 2 \quad 1]^T$ ، $C = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ ،

$$A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix},$$

باستخدام خوارزمية 3.8؟

5.8 أوجد $\delta_p(X,\mu)$ لتمرين 1.8؟

6.8 حل البرنامج في تمرين 4.8 باستخدام خوارزمية 19.8؟صلي