الباب الرابع

نظام كاريش كون تكر
Karush-Kuhn-Tucker System

- مقدمة • نظام KKT المخفض
- المعادلات الناظمية

1.4 مقدمة Introduction

إن الجانب الأكثر استهلاكا للوقت عند كل دورة في طريقة تابع المسار هو حل نظام المعادلات التي تُعرِّف متجهات اتجاه الخطوة $\Delta x, \Delta w, \Delta y, \Delta z$:

$$A\Delta x + \Delta w = \rho \tag{1.4}$$

$$A^T\Delta y - \Delta z = \sigma \tag{2.4}$$

$$Z\Delta x + X\Delta z = \mu e - XZe \tag{3.4}$$

$$W\Delta y + Y\Delta w = \mu e - YWe \tag{4.4}$$

بعد تعديلات بسيطة، هذه المعادلات يمكن أن تكتب على شكل مصفوفة من عدة قوالب كالتالي:

$$\begin{bmatrix} -XZ^{-1} & 0 & -I & 0 \\ 0 & 0 & A & I \\ -I & A^T & 0 & 0 \\ 0 & I & 0 & YW^{-1} \end{bmatrix} \begin{bmatrix} \Delta z \\ \Delta y \\ \Delta z \\ \Delta w \end{bmatrix} = \begin{bmatrix} -\mu Z^{-1}e + x \\ \rho \\ \sigma \\ \mu W^{-1}e - y \end{bmatrix} \tag{5.4}$$

هـذا النظـام يسـمى نظـام كـاريش- كـون- تكـر Karush-Kuhn-Tucker ، أو بطريقة أبسط يسمى نظام KKT. وهو نظام خطي متماثل مكون من $2n+2m$ معادلة في $2n+2m$ مجهول. ويمكن حلها عن طريق تحليل هذا النظام الكبير وبعد ذلك نقوم بالتعويض الأمامي والخلفي. وعلى أية حال إن عمل بعض هذه الحسابات يدوياً ومن ثم حل أنظمة صغيرة يجعل حل النظام أكثر فعالية. ويوجد مرحلتان لعمل هذه الأنظمة الصغيرة، المرحلة الأولى هي عبارة عن تحويل النظام الكبير إلى أنظمة صغيرة. ويسمى النظام المتبقي بنظام KKT المخفض، وبعد المرحلة الثانية يسمى نظام المعادلات الناظمي. وسوف نناقش هذين النظامين في الفصلين القادمين.

2.4 نظام KKT المخفض The Reduced KKT System

تـذكّر أن X و Y هـن مصـفوفات قطريـة وبالتـالي المعـادلات (3.4) و(4.4) سهلة الحل، ولذلك يمكن إزالتها من البداية. ولإبقاء التماثل الذي رأيناه في (5.4) يجب حل (3.4) و(4.4) بالنسبة لـ Δz و Δw على التوالي:

$$\Delta z = X^{-1}(\mu e - XZe - Z\Delta x)$$

$$\Delta w = Y^{-1}(\mu e - YWe - W\Delta y)$$

بتعويض هـذه الصيغ في (1.4) و(2.4)، نحصل على ما أسميناه نظام KKT المخفض:

$$A\Delta x - Y^{-1}W\Delta y = \rho - \mu Y^{-1}e + w \qquad (6.4)$$

$$A^T\Delta y + X^{-1}Z\Delta x = \sigma + \mu X^{-1}e - z \qquad (7.4)$$

بالتعويض في تعريف كل من ρ و σ ومن ثم كتابة النظام في صيغة مصفوفة نحصل على:

$$\begin{bmatrix} -Y^{-1}W & A \\ A^T & X^{-1}Z \end{bmatrix}\begin{bmatrix} \Delta y \\ \Delta x \end{bmatrix} = \begin{bmatrix} b - Ax - \mu Y^{-1}e \\ c - A^T y + \mu X^{-1}e \end{bmatrix}.$$

لاحظ أن مصفوفة KKT المخفضة هي أيضاً مصفوفة متماثلة. إن الطرف الأيمن يُظهر تماثلاً بين المسألة الأساسية والثنائية. ولزيادة تخفيض النظام نحتاج أولاً لكسر هذا التماثل الذي أبقيناه إلى الآن.

3.4 المعادلات الناظمية The Normal Equations

للمرحلة الثانية، لدينا خياران: أما الأول فهو أن نقوم بحل (6.4) بالنسبة لـ Δy ونحذفها من (7.4)، وأما الثاني فهو أن نقوم بحل (7.4) بالنسبة لـ Δx ونحذفها من (6.4). الآن لنفترض أننا نتبع المنهج الأخير. في هذه الحالة نحصل من (7.4) على:

$$\Delta x = XZ^{-1}(c - A^T y + \mu X^{-1}e - A^T \Delta y)$$ ،

التي نستخدمها لحذف Δx من (6.4) لنحصل على

$$-(Y^{-1}W + AXZ^{-1}A^T)\Delta y = b - Ax - \mu Y^{-1}e - AXZ^{-1}(c - A^T y + \mu X^{-1}e)$$

النظام الأخير هو نظام مكون من m معادلة في m مجهول، ويسمى نظام معادلات ناظمي للمسألة الأصلية. وهو نظام المعادلات التي تتضمن المصفوفة $Y^{-1}W + AXZ^{-1}A^T$. والحد $Y^{-1}W$ هو مصفوفة قطرية، ولذا فإن الجزء المهم في هذه المصفوفة هو الذي يحتوي على الحد $AXZ^{-1}A^T$.

لدينا A متناثرة sparse (وهذا هو الواقع في كثير من المسائل التي تأتي من التطبيقات الحقيقية العملية)، ويمكن أن نتوقع أن المصفوفة $AXZ^{-1}A^T$ متناثرة على نفس النمط. وعلى أية حال نحتاج أن نحقق في تناثر $AXZ^{-1}A^T$ أو عدمه بشكل أكثر دقة. لاحظ أن العناصر $(i,j)th$ في هذه المصفوفة معطاة كالتالي:

$$(AXZ^{-1}A^T)_{ij} = \sum_{k=1}^{n} a_{ik}\frac{x_k}{z_k}a_{jk} \quad (8.4)$$

إن العنصر $(i,j)th$ هو عبارة عن ضرب داخلي موزون بين الصف i والصف j للمصفوفة A. إذا كانت هذه الصفوف لها نمط غير صفري متباعد فإن الضرب الداخلي لابد أن يكون صفراً، وفيما عدا ذلك يجب التعامل معه على أنه غير صفر. وهذا غير جيد لأنه إذا كانت A متناثرة ولها عمود واحد كثيف تصبح المصفوفة (8.4) كلها كثيفة.

ويجب أن لا ننس أنه لا يلزم أن نحل المسألة عن طريق معادلات الناظمي الأساسي، ولكن بدلاً من ذلك بإمكاننا أن نختار البديل الآخر للحل (6.4) بالنسبة لـ Δy

$$\Delta y = -YW^{-1}(b - Ax - \mu Y^{-1}e - A\Delta x)$$

ونقوم بحذفها من (7.4):

$$\begin{aligned}(A^T YW^{-1}A + X^{-1}Z)\Delta x = c - A^T y + \mu X^{-1}e \\ + A^T YW^{-1}(b - Ax - \mu Y^{-1}e)\end{aligned}$$

إن هذا النظام هو نظام مكون من n معادلة في n مجهول، يسمى نظام معادلات ناظمي للمسألة الثنائية. لاحظ أن الأعمدة الكثيفة لا تمثل مشكلة في هذه المعادلات. بينما هذا النظام أكبر من النظام السابق، وهو أيضاً متناثر، وهذا التناثر دائماً ما يكون له أكثر أهمية من بعد المصفوفة.

وبشكل عام لا نستطيع القول أن حل نظام معادلات ناظمي للمسألة الأساسية أو حل نظام معادلات ناظمي للمسألة الثنائية هو الأفضل، فلكل حالة خصائصها وإيجابياتها و سلبياتها. وإنما نستطيع القول أنه إذا كانت

المصفوفة A فيها عمود كثيف يصبح حل نظام ناظمي للمسألة الثنائية أفضل، وإذا كان فيها صف كثيف يصبح حل نظام ناظمي للمسألة الأساسية أفضل، وتكون المشكلة أكبر إذا كانت المصفوفة A فيها صف وعمود كثيفان حيث يصبح حل النظامين مكلفاً جداً.

تمارين الباب الرابع

1.4 أثبت أن المعادلة التالية صحيحة:

$$(E^{-1}+ADA^{T})^{-1}=E-EA(A^{T}EA+D^{-1})^{-1}A^{T}E$$

حيث أن المصفوفات أعلاه قابلة للعكس؟

2.4 استخدم المعادلة في التمرين 1.4 للتحقق من التكافؤ بين المعادلة

$$\Delta x=(D^{2}-D^{2}A^{T}(E^{-2}+AD^{2}A^{T})^{-1}AD^{2})(c-A^{T}y+\mu X^{-1}e)$$
$$+D^{2}A^{T}(E+AD^{2}A^{T})(b-Ax-Ye)$$

والمعادلة

$$\Delta x=(A^{T}E^{2}A+D^{-2})^{-1}(c-A^{T}y+\mu X^{-1}e)$$
$$+(A^{T}E^{2}A+D^{-2})^{-1}A^{T}E^{2}(b-Ax-\mu Y^{-1}e)$$

وذلك لـ Δx ، حيث $D^{2}=XZ^{-1}$ و $E^{2}=YW^{-1}$ ؟