1 Dualidade e relaxação

- Q: Qual é o contrário do "espaço dual"?
- R: O espaço "donothing".

Eduardo Tengan, em uma aula qualquer

A noção de dualidade não é uma noção perene só na álgebra linear e suas áreas adjacentes (conhecidas em conjunto como "matemática"), também ocupa um lugar especial na teoria de otimização. Em otimização, dualidade vem em diversas formas: dualidade de programação linear, de programação linear inteira (lagrangeana), entre outras. Frequentemente, os métodos de otimização mais confiavelmente eficicientes são métodos primal-dual.

A dualidade é um tema unificador porque podemos expressar todos os duais como dual de inferência ou de relaxação (frequentemente vezes, como ambos). Um dual de inferência pode ser visto como o problema de inferir das restrições nas variáveis um maior limitante inferior na imagem da função objetivo (assumimos daqui para frente, sem perda de generalidade, que queremos minimizar a função). Assim, a busca ocorre sobre provas, no lugar de sobre valores das variáveis. Mais precisamente, se tivermos um problema de otimização da forma

$$\begin{aligned} &min\ f(x),\\ &R(x),\\ &x \in X, \end{aligned}$$

em que $R(x)$ denota as restrições sobre x, sendo X o seu domínio, seu respectivo **dual de inferência** nos coloca o problema de maximizar o limitante inferior em $f(x)$ que é inferível das restrições. Isso ocorre na forma de uma busca por uma prova do limitante ótimo (ou "bom o suficiente", se não pudermos pedir por um "ótimo"):

dual de inferência

$$\begin{aligned} &max\ v,\\ &C(x) \vdash^P (f(x) \geq v),\\ &v \in \mathbb{R}, P \in \mathcal{P}, \end{aligned}$$

em que $C(x) \vdash^P (f(x) \geq v)$ indica que P é uma prova de $(f(x) \geq v)$, usando o que é assumido em $C(x)$ (no caso, $C(x)$ indica as restrições). O domínio de P é uma família de provas $\mathcal{P}$, sendo (v, P) um par de solução dual. Quando o problema inicial (dito "**primal**") não tem soluções, o conjunto de soluções para

primal

o dual é ilimitado.

Quando o primal for um problema de programação linear, por exemplo, a família de provas pode consistir de combinações lineares não-negativas de restrições de desigualdade, sendo a solução dual identificada com os multiplicadores na combinação linear que deriva o maior limitante.

Direto da definição, podemos inferir o **princípio de dualidade fraca**: um valor factível v do dual de inferência nunca pode ser maior do que qualquer valor factível do primal, o que nos possibilita limitar o valor para o ótimo do primal. Ademais, a diferença (a "lacuna") entre os valores ótimos do primal e do dual é chamada **lacuna de dualidade**. A lacuna de dualidade pode ser igual a zero, mas não é garantido que esse possa ser o caso (isto é, não é garantido que as soluções ótimas primal e dual possam ter o mesmo custo). Em particular, pode ser que exista um maior limitante inferior a $f(x)$, mas esse limitante inferior não ser inferível em $\mathcal{P}$. Quando esse é o caso, $\mathcal{P}$ é dita incompleta. Mais em geral, se $C(x)$ implica C mas não existe $P \in \mathcal{P}$ que infira C, $\mathcal{P}$ é dita incompleta. Caso contrário, ela é dita completa. Quando $\mathcal{P}$ for completa, a lacuna de dualidade pode, ao menos em princípio, ser reduzida a zero pelo dual, o que constitui o princípio de **dualidade forte**[1].

princípio de dualidade fraca

lacuna de dualidade

dualidade forte

A ideia de dualidade por inferência provê um plano unificador para várias ideias centrais em otimização no geral e em programação por restrições em particular. Entre outras coisas, dualidade por inferência nos provê métodos para *aprendizado de cláusulas*, frequentemente muito úteis. Buscas com *aprendizado de cláusulas* cresceram de pesquisa em *aprendizado baseado em explicações*, um ramo de inteligência artificial que busca melhorar a eficiência de métodos de busca por *backtracking* por meio de "explicações de falha", na forma de novas restrições ao problema original. Para CSPs em geral, tais "explicações" são chamadas *nogood* (ou, às vezes, "conflitos"). Foi mostrado que uma grande quantidade de problemas que não podem (até o presente momento) ser resolvidos por outras técnicas podem ser resolvidos de maneira eficiente por *aprendizado de cláusulas*. Em particlar, problemas (enteriormente) abertos em teoria de grupos foram resolvidos com essa técnica (entre outras). Para mais informações, veja [1]

1.0.1 Dual Substituto

Existem na literatura de programação por restrições, assim como na de programação matemática, diversas noções específicas de dual, para diversos tipos de problemas diferentes. Entre eles, destacam-se como especialmente úteis os duais *substituto* e *lagrangeano*, aos quais devotaremos alguma atenção nesta seção e na próxima, colocando ênfase especial na relação entre eles.

[1] Os "princípios" de dualidade fraca e forte são às vezes referidos como teoremas (porque precisam ser provados). Suas provas, no entanto, são tão simples, que foram omitidas a fim de não aborrecer a leitora.

Seja um problema do tipo

$$
\begin{aligned}
& min f(x), \\
& g(x) \geq 0, \\
& x \in X,
\end{aligned} \tag{1}
$$

em que $g(x)$ não necessariamente é linear e $x \in X$ indica as demais restrições em x. Note que problemas de programação inteira mista são um caso particular em que $g(x)$ é uma função linear e $X = \mathbb{R}^n + \mathbb{Z}^m$. Obtemos o dual de inferência desse problema tomando combinações lineares não negativas e implicação como método de inferência, fazendo

$$
\begin{aligned}
& max\ v, \\
& (g(x) \geq 0) \vdash^P (f(x) \geq v), \\
& P \in \mathcal{P}, v \in \mathbb{R},
\end{aligned} \tag{2}
$$

em que cada prova em $\mathcal{P}$ corresponde a um vetor u de multiplicadores e a prova P deduz $f(x) \geq v$ de $g(x) \geq 0$ quando a $ug(x) \geq 0 \Rightarrow f(x) \geq v$. Isto é o mesmo que dizer que o mínimo de $f(x)$ com $ug(x) \geq 0$ e $x \in X$ é no mínimo v. Daí segue que essa formulação é equivalente a

$$
\begin{aligned}
& max\ \sigma(u), \\
& \sigma(u) = min_{x \in X}\{f(x) | ug(x) \geq 0\}, \\
& u \geq 0,
\end{aligned}
$$

Pergunta: o *dual substituto* é completo?

1.0.2 Dual Lagrangeano

O dual lagrangeano ao problema 1 é semelhante ao dual substituto, exceto que, no dual lagrangeano, as provas fazem uso de combinações lineares a *dominação*[2]. O dual é o problema 2, em que $\mathcal{P}$ consiste de provas que mesclam combinação linear com dominação: $f(x) \geq v$ é inferido de $g(x) \geq 0$ se $\exists\, l \geq 0 : lg(x)$ domina $f(x) \geq v$.

Como dominação é um requisito mais fraco do que implicação, o dual lagrangeano resulta em um dual de inferência mais fraco do que o *substituto* e, assim, possibilita limitantes mais fracos em relação ao *substituto*. Apesar disso,

[2] Dizemos que uma desigualdade $g(x) \geq 0$ domina uma $h(x) \geq 0$ se $g(x) \leq h(x)\ \forall x$, ou se não existe x tal que $g(x) \geq 0$

para outros fatores o lagrangeano goza de propriedades mais interessantes do que o *substituto*, o que o torna atraente para uma grande variedade de problemas. Por exemplo, o conjunto de soluções factíveis para o *dual lagrangeano* é côncavo (o que não vale, no geral, para o *dual substituto*). Isso pode ser visto facilmente, se, supondo que $lg(x) \geq 0$ é factível para $l \geq 0$, observarmos que o lagrangeano maximiza v com $l \geq 0$ e $lg(x) \geq f(x) - v \; \forall x \in X$. Reescrevendo $lg(x) \leq f(x) - v$ como $v \leq f(x) - lg(x)$ e fazendo $\lambda(l, x) = f(x) - lg(x)$, nosso lagrangeano é equivalente a

$$
\begin{aligned}
& max \; \lambda(l), && (3)\\
& \lambda(l) := inf_{x \in X}\{f(x) - lg(x)\}, && (4)\\
& l \geq 0, && (5)
\end{aligned}
$$

e o conjunto de ínfimos de funções afim (como em 4) é um conjunto côncavo (λ é afim para x fixo). O vetor l é popularmente conhecido como vetor de multiplicadores de lagrange.

1.1 Complexidade

Um certificado de factibilidade (infactilibidade/otimalidade) é um pedaço de informação que permite a verificação da factibilidade (infactibilidade/otimalidade) da instância de um problema. Dizemos que um problema de otimização pertence ao NP (de *nondeterministic polynomial*) se existe um certificado de factibilidade polinomial para qualquer instância factível do problema: a computação requerida para testar factibilidade, pelo certificado, é limitada por uma função polinomial do tamanho da instância do problema. Um problema de otimização é co-NP se existe um certificado polinomial de otimalidade para qualquer instância.

Se um problema de otimização primal for co-NP, seu dual de inferência é NP para alguma família de provas $\mathcal{P}$: se o problema for co-NP, seja $\mathcal{P}$ a família de provas P_i (que assumimos polinomiais) de otimalidade da instância i, de valor ótimo v_i, provê um certificado polinomial de factibilidade do dual (que é, portanto, NP).

Programação linear pertence a ambos NP e co-NP. Problemas desse tipo são tidos ter "boa caracterização", no sentido de que soluções factíveis e provas de otimalidade são facilmente escritas.

Leituras adicionais

[1] Beame Paul and Kautz Henry and Sabharwal Ashish, Towards Understanding and Harnessing the Potential of Clause Learning, Journal of Artificial Intelligence Research 22 (2004) 319-351.

[2] Hooker N. John, Integrated Methods for Optimization, Second Edition, Springer Verlag, 2012.