1. Qual a qualidade do limitante dual obtido ao resolver o dual lagrangiano ?
2. Como se pode resolver o dual lagrangiano ?

- Qualidade do limitante:
  - supor que $X = \{x^1, x^2, \dots, x^T\}$, com T finito;

- Resolvendo o dual Lagrangiano:
  - alternativa 1: usar um algoritmo de otimização por subgradientes;
  - alternativa 2: usar PL e um algoritmo de planos de corte.

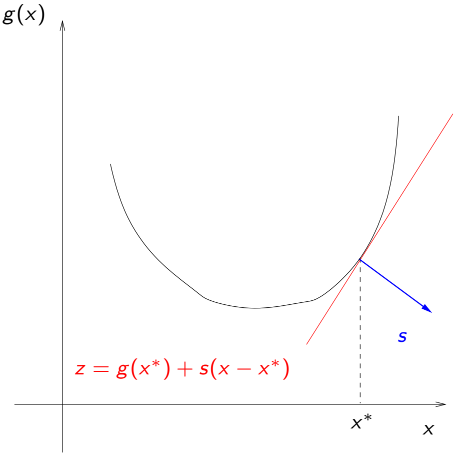
- **Definição 5**: $g : \mathbb{R}^n \to \mathbb{R}$ é uma **função convexa** se, para todos pontos x e y de $\mathbb{R}^n$ e $\lambda \in [0, 1]$,
$$g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda)g(y)$$

- **Proposição 6**: $z(u) = \max\{cx + u(d - Dx)\}$ é convexa.

Prova: $\begin{cases} w = \lambda u + (1-\lambda)v \\ z(w) = c\overline{x} + w(d - D\overline{x}) \qquad (\overline{x} \in X) \end{cases}$

$$\begin{array}{rcll} c\overline{x} + u(d - D\overline{x}) & \leq & z(u) & [\lambda] \\ c\overline{x} + v(d - D\overline{x}) & \leq & z(v) & [1-\lambda] \\ & & & \\ \lambda z(u) + (1-\lambda)z(v) & \geq & z(w) & \square. \end{array}$$

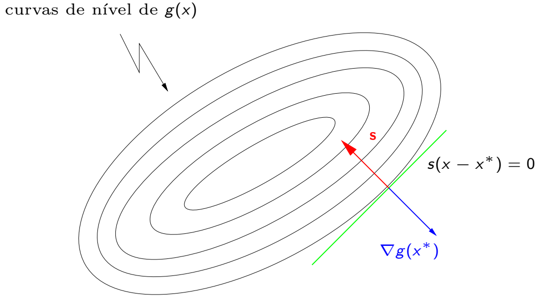
- **Proposição 7**: uma função $g : \mathbb{R}^n \to \mathbb{R}$ é convexa *se e somente se*, para todo $x^* \in \mathbb{R}^n$ existe $s \in \mathbb{R}^n$ tal que $g(x^*) + s(x - x^*) \leq g(x)$ para todo $x \in \mathbb{R}^n$.



- Estando em x^*, que direção deve ser seguida para minimizar $g(x)$?

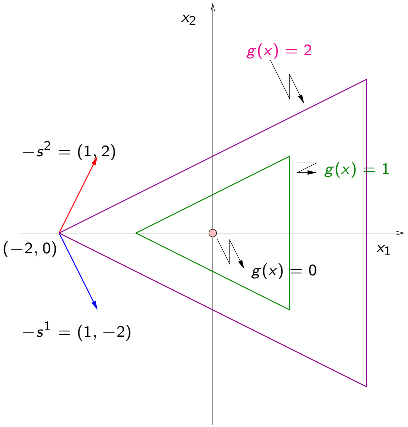
  Usando a Proposição 7, se x é tal que $g(x) < g(x^*)$, então necessariamente $s(x - x^*) < 0$. Ou seja, a partir de x^* movendo-se (de uma quantidade adequada) na direção oposta àquela de s, $g(.)$ irá diminuir de valor !

- Como achar s ? Se g é diferenciável em x^*, fazer $s = -\nabla g(x^*)$ (gradiente de g em x^*).



- E se $g(.)$ é convexa mas não for diferenciável ?

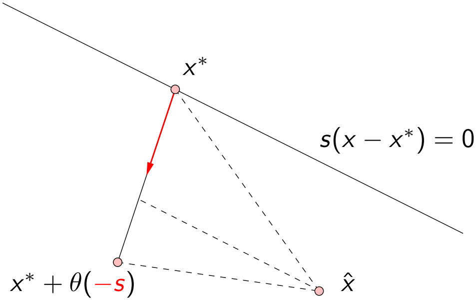
$$g(x_1, x_2) = \max\{x_1, -x_1 + 2x_2, -x_1 - 2x_2\}$$



- $s^1 = (-1, 2)$
- $s^2 = (-1, -2)$
- o hiperplano suporte não é único em $(-2, 0)$!
- as direções $-s^1$ e $-s^2$ não são de descida !

- Pelo menos sabemos que a solução ótima satisfaz
  $s(x - x^*) < 0$!

- Portanto, é possível sair de x^* e dar um passo *pequeno* na direção $-s$ de modo a me aproximar mais de um ponto ótimo $\hat{x}$.



- **Definição 8**: se $g : \mathbb{R}^n \to \mathbb{R}$ é uma função convexa, então s é um **subgradiente** de g em x^* *se e somente se*
$$s(x - x^*) \leq g(x) - g(x^*), \forall x \in \mathbb{R}^n.$$
O **subdiferencial** ($\delta g(x^*)$) de g em x^* é o conjunto de todos os subgradientes de g neste ponto.

- **Proposição 9**: Se g é convexa e $0 \in \delta g(x^*)$ então $g(x^*) = \min\{g(x) : x \in \mathbb{R}^n\}$, ou seja, x^* é **ótimo**.

- Em teoria: se queremos minimizar uma função convexa g, basta começar de um ponto qualquer e iterativamente ir se deslocando a pequenos passos na direção oposta àquela de um subgradiente de g naquele ponto até que 0 pertença ao subdiferencial do ponto corrente.

- Como calcular um subgradiente de $z(u) = \max\{cx + u(d - Dx) : x \in X\}$ no ponto u ?

- **Proposição 10**: Seja $\overline{x} \in X$ tal que $z(u) = c\overline{x} + u(d - D\overline{x})$. Então, $(d - D\overline{x})$ é um subgradiente de $z(u)$ em u.

Prova:

$z(v) \geq cx + v(d - Dx), \ \forall x \in X,$ logo
$z(v) \geq c\overline{x} + v(d - D\overline{x}).$

Mas, $c\overline{x} = z(u) - u(d - D\overline{x}),$ portanto,
$z(v) \geq z(u) + (v - u)(d - D\overline{x}).$ $\square$

- O **algoritmo de subgradiente** para resolver o problema dual lagrangiano:

> Inicialização: $k = 0$, escolher u^0.
> Iteração k:
> - $u = u^k$;
> - Resolver $IP(u^k)$. Seja $x(u^k)$ uma solução ótima.
> - Passo na direção oposta ao subgradiente:
> $u^{k+1} \leftarrow \max\{u^k - \mu_{\mathbf{k}}(d - Dx(u^k)), 0\}$
> - $k \leftarrow k + 1$;
> - Se o **critério de parada** não está satisfeito, **execute** mais uma iteração.

- μ é chamado de **tamanho do passo**.
- Qual o tamanho do passo que deve ser dado ?
  (problemas de convergência)

- **Teorema 11**:
  (a) Se $\sum_k \mu_k \to \infty$ e $\mu_k \to 0$ então $z(u^k) \to w_{LD}$.
  (b) Se $\mu_k = \mu_0 \rho^k$, com $\rho < 1$ e μ_0 e ρ são suficientemente grandes, então $z(u^k) \to w_{LD}$.
  (c) Se $\overline{w} \geq w_{LD}$ e $\mu_k = \epsilon \frac{z(u^k) - \overline{w}}{\|d - Dx(u^k)\|^2}$ com $0 < \epsilon < 2$, então $z(u^k) \to \overline{w}$, ou o algoritmo encontra u^k com $\overline{w} \geq z(u^k) \geq w_{LD}$ para algum k finito.

- **Comentários**:
  (a) converge lentamente !
  (b) dificuldade em escolher μ_0 e ρ "altos".
  Na prática só faz decréscimo geométrico do passo após um número constante de iterações.
  (c) dificuldade em encontrar um limitante superior $\overline{w}$.
  Na prática, usa-se uma estimativa para o limitante superior, mas isto não garante convergência !
  Esta estimativa muitas vezes é substituída pelo melhor limitante primal conhecido ... (ver discussão no livro texto)
  É o método mais usado !

Método do Subgradiente aplicado ao TSP simétrico

- restrições dualizadas: grau dos nós $(= 2)$.
- Direção do passo:

$$u_i^{k+1} = u_i^k + \mu_k \left(2 - \sum_{e \in \delta(i)} x_e(u^k) \right).$$

- tamanho do passo (regra (c)):

$$\mu_k = \epsilon_k(\underline{w} - z(u^k)) / \sum_{i \in V} (2 - \sum_{e \in \delta(i)} x_e(u^k))^2.$$

- **Nota**: o problema dual neste caso é de maximização e os multiplicadores são irrestritos em sinal.

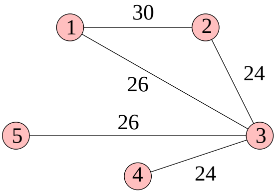
- Instância: grafo completo com 5 nós e custos dados por

$$c = \begin{bmatrix} - & 30 & 26 & 50 & 40 \\ & - & 24 & 40 & 50 \\ & & - & 24 & 26 \\ & & & - & 30 \\ & & & & - \end{bmatrix}$$

- Uma solução heurística: *tour* $\{1, 2, 3, 4, 5, 1\}$ de custo 148.
- Método do subgradiente com a regra (c) e $\epsilon_k = 1$ ...

- Iteração 1: supondo $u^1 = \{0, 0, 0, 0, 0\}$, a matriz de custos lagrangianos coincide com a matriz original.

$$\tilde{c} = \begin{bmatrix} - & 30 & 26 & 50 & 40 \\ & - & 24 & 40 & 50 \\ & & - & 24 & 26 \\ & & & - & 30 \\ & & & & - \end{bmatrix}$$



- 1-árvore ótima com custo $z(u^1) = 130$;
- subgradiente em u^1:

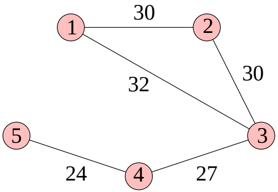
$$(2 - \sum_{e \in \delta(i)} x_e(u^1))_{i=1,\dots,5} = (0, 0, -2, 1, 1).$$

- atualização dos multiplicadores:

$$u^2 = u^1 + 1.[(148 - 130)/6](0, 0, -2, 1, 1)$$
$$u^2 = (0, 0, -6, 3, 3)$$

- Iteração 2: para $u^2 = (0, 0, -6, 3, 3)$, a matriz de custos lagrangianos é dada por:

$$\tilde{c} = \begin{bmatrix} - & 30 & 32 & 47 & 37 \\ & - & 30 & 37 & 47 \\ & & - & 27 & 29 \\ & & & - & 24 \\ & & & & - \end{bmatrix}$$



- 1-árvore ótima com custo $z(u^2) = 143 + 2\sum_i u_i^2 = 143$;
- subgradiente em u^2:

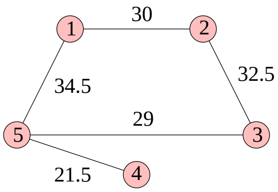
$$(2 - \sum_{e \in \delta(i)} x_e(u^2))_{i=1,\dots,5} = (0, 0, -1, 0, 1).$$

- atualização dos multiplicadores:

$$u^3 = u^2 + 1.[(148 - 143)/2](0, 0, -1, 0, 1)$$
$$u^3 = (0, 0, -8.5, 3, 5.5)$$

- Iteração 3: para $u^3 = (0, 0, -8.5, 3, 5.5)$, a matriz de custos lagrangianos é dada por:

$$\tilde{c} = \begin{bmatrix} - & 30 & 34.5 & 47 & 34.5 \\ & - & 32.5 & 37 & 44.5 \\ & & - & 29.5 & 29 \\ & & & - & 21.5 \\ & & & & - \end{bmatrix}$$



- 1-árvore ótima com custo $z(u^2) = 147.5$;
- **solução ótima**: os custos inteiros e já se conhece um limitante superior de 148 !
- A relaxação lagrangiana pode ser usada dentro de um **algoritmo exato** do tipo *branch-and-bound* para produzir limitantes duais.