



Conception et Planification de Réseaux

Routage Optimal

Olivier BRUN et Jean-Marie GARCIA

Plan

- **Fonction coût**
- **Problème du routage optimal**
- **Caractérisation du routage optimal**
- **Méthodes de directions admissibles**
- **Méthode de Frank-Wolfe (flow deviation)**
- **Méthode du gradient projeté**



Fonction Coût

Fonction Coût

● Interêt

- Evaluer les performances d'un algorithme de routage,
- Quantifier la notion de congestion
 - Modèle d'évaluation de performances basé sur le trafic des liens du réseau.

● Evaluation de la congestion

- On utilise le trafic moyen sur chaque lien
- F_{ij} : taux d'arrivée des paquets sur chaque lien (i, j) .
 - F_{ij} est appelé le flot sur le lien (i, j) .
 - F_{ij} est mesuré en unités de données/seconde.

Fonction Coût

● Hypothèses

- Les processus d'arrivées des paquets sur les liens de communications sont stationnaires (leurs statistiques n'évoluent pas au cours du temps).
- Raisonnable quand ces statistiques changent très lentement par rapport au temps nécessaire pour vider une file d'attente
- Typiquement le cas pour des réseaux dans lesquels le nombre de flots pour chaque couple origine-destination est très grand, le trafic de chaque flot n'étant qu'une petite fraction du trafic total.
- Approprié pour des réseaux d'opérateurs ou d'ISP

Fonction Coût

● Fonction coût additive :

- Expression de la forme :

$$\sum_{(i,j)} D_{ij}(F_{ij}) \quad (1)$$

- D_{ij} caractérise la congestion sur le lien (i, j)
- Une formule souvent utilisée est,

$$D_{ij}(F_{ij}) = \frac{F_{ij}}{C_{ij} - F_{ij}} + d_{ij} F_{ij} \quad (2)$$

où C_{ij} est la capacité de transmission du lien (i, j) , mesurée en paquets/seconde, tandis que d_{ij} représente le délai de propagation.

Le coût global représente alors le nombre de paquets dans le réseau.

Fonction Coût

● Hypothèses :

- **Chaque lien se comporte comme une file M/M/1**
 - Arrivées aléatoires suivant un processus de Poisson,
 - Temps de traitement exponentiellement distribués,
 - Indépendance des processus d'arrivées et de service,
 - Réseaux ouverts de Jackson.
- **Hypothèses généralement non vérifiées dans un réseau réel.**
- **Formule intéressante néanmoins car elle exprime le fait que la congestion apparaît lorsque $F_{i,j} \rightarrow C_{i,j}$**

Fonction Coût

● Fonction coût non additive :

- Expression de la forme :

$$\max_{(i,j)} \frac{F_{i,j}}{C_{i,j}} \quad (3)$$

- Correspond au taux maximal d'utilisation des liens du réseau,
- Propriétés qualitatives similaires à la précédente fonction coût,
- Résultats assez similaires d'un point de vue optimisation du routage,
- Dans la suite, nous utiliserons des fonctions coûts additives de la forme $\sum_{(i,j)} D_{ij}(F_{ij})$ qui sont souvent plus faciles à optimiser.



Problème de Routage Optimal

Problème de Routage Optimal

● Introduction & Notations

- Un couple $w = (o, d)$ de noeuds distincts o et d sera appelé dans la suite un couple origine-destination (OD).
- Le trafic exogène du noeud o vers le noeud d est supposé stationnaire et sera noté r_w .
- r_w représente le taux d'arrivée au noeud o des paquets qui ont pour destination d .
- L'objectif du routage est de répartir chaque demande r_w sur les différents chemins reliant l'origine o à la destination d de telle sorte que le flot total F_{ij} résultant sur chaque lien (i, j) minimise la fonction coût.

Problème de Routage Optimal

● Notations

- W est l'ensemble des couples origine-destination,
- P_w est l'ensemble des chemins reliant les extrêmités du couple OD w ,
 - En général on ne considère qu'une sélection des chemins entre o et d
- x_p le flot en paquets/seconde transmis sur le chemin $p \in P_w$ du couple OD w .
- Les flots sur les chemins $\{x_p \mid w \in W, p \in P_w\}$ sont sujets à des contraintes de conservation des demandes r_w :

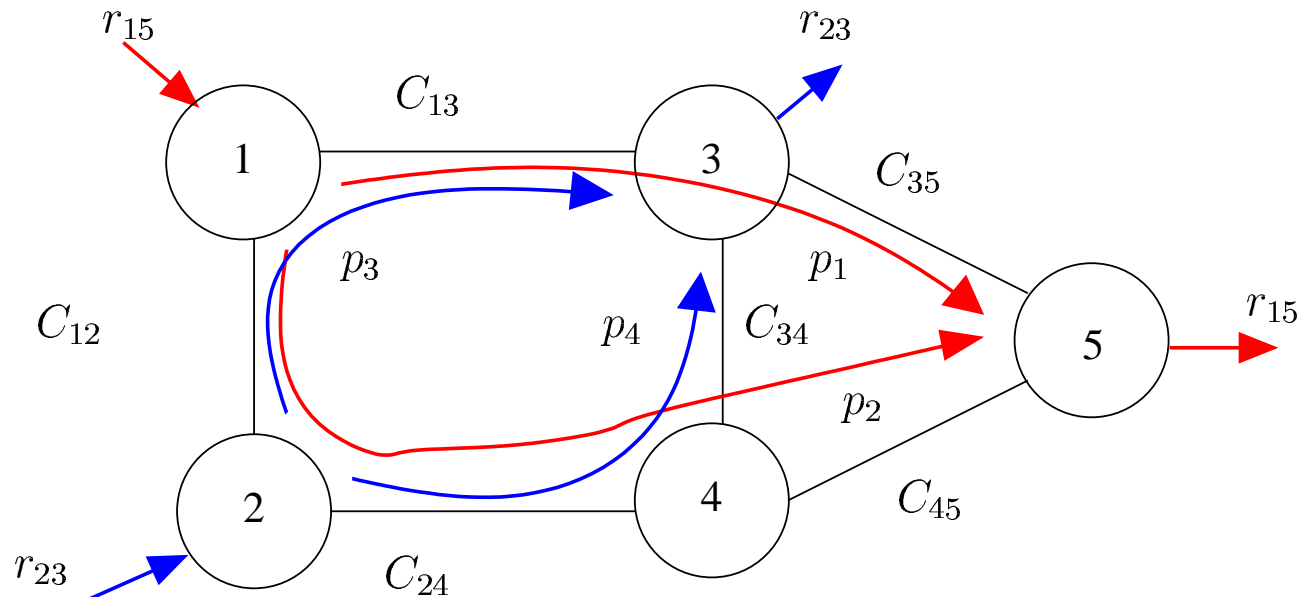
$$\sum_{p \in P_w} x_p = r_w \quad \forall w \in W$$

$$x_p \geq 0 \quad \forall p \in P_w, \forall w \in W$$

Problème de Routage Optimal

● Exemple

- **Demandes r_{15} et r_{23} :** $W = \{(1, 5), (2, 3)\}$
- **Chemins de r_{15} :** $P_{15} = \{p_1, p_2\}$ où $p_1 = \{1, 3, 5\}$ et $p_2 = \{1, 2, 4, 5\}$
- **Chemins de r_{23} :** $P_{23} = \{p_3, p_4\}$ où $p_3 = \{2, 1, 3\}$ et $p_4 = \{2, 4, 3\}$
- **Contraintes :** $x_{p_1} + x_{p_2} = r_{15}$ et $x_{p_3} + x_{p_4} = r_{23}$ avec $x_{p_i} \geq 0$



Problème de Routage Optimal

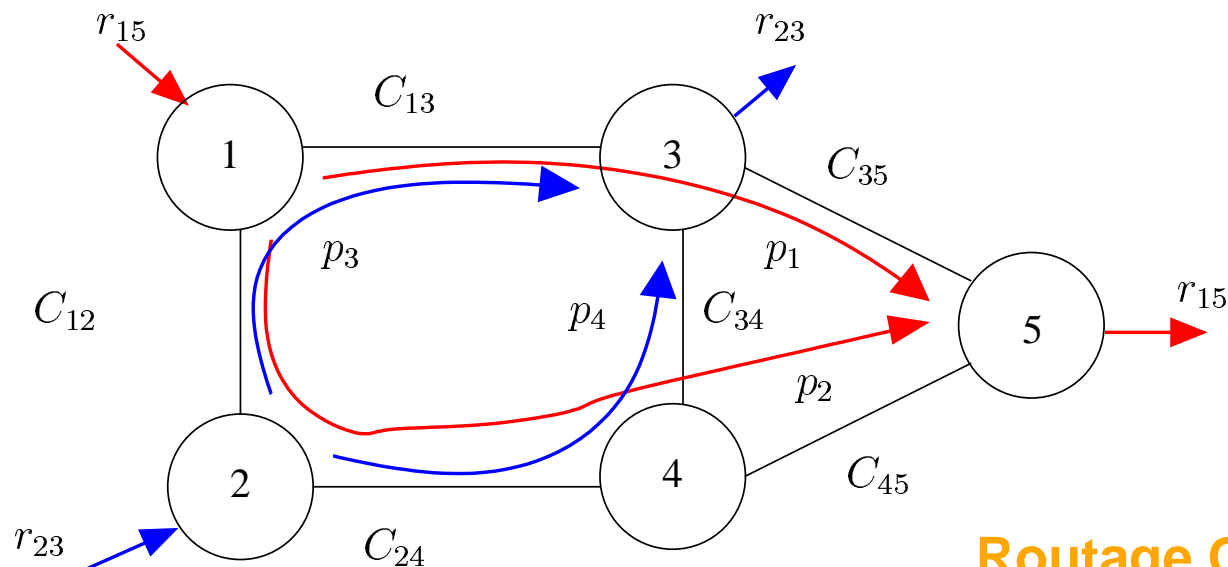
● Flot sur les liens

- Le trafic total sur le lien (i, j) est la somme des flots qui passent par ce lien,

$$F_{ij} = \sum_{p/(i,j) \in p} x_p \quad (4)$$

où la somme est sur tous les chemins p contenant le lien (i, j) .

- Exemple : $F_{13} = x_{p_1} + x_{p_3}$



Problème de Routage Optimal

● Problème de routage optimal

- Trouver l'ensemble des flots $\{x_p\}$ pour chaque demande r_w qui minimise la fonction coût,

$$\sum_{(i,j)} D_{ij} (F_{ij}) \quad (5)$$

sous les contraintes de conservation des demandes,

$$\sum_{p \in P_w} x_p = r_w \quad \forall w \in W$$
$$x_p \geq 0 \quad \forall p \in P_w, \forall w \in W$$

Problème de Routage Optimal

● Problème de routage optimal

● Sachant que,

$$F_{ij} = \sum_{p/(i,j) \in p} x_p \quad (6)$$

le problème revient à trouver le vecteur $x = (x_p)_{p \in P_w, w \in W}$ solution du programme mathématique non linéaire suivant :

$$\left\{ \begin{array}{l} \text{Minimiser :} \\ \text{sous les contraintes :} \\ \sum_{p \in P_w} x_p = r_w \\ x_p \geq 0 \end{array} \right. \quad \begin{array}{l} D(x) = \sum_{(i,j)} D_{ij} \left(\sum_{p/(i,j) \in p} x_p \right) \\ \forall w \in W \\ \forall p \in P_w, \forall w \in W \end{array}$$

Problème de Routage Optimal

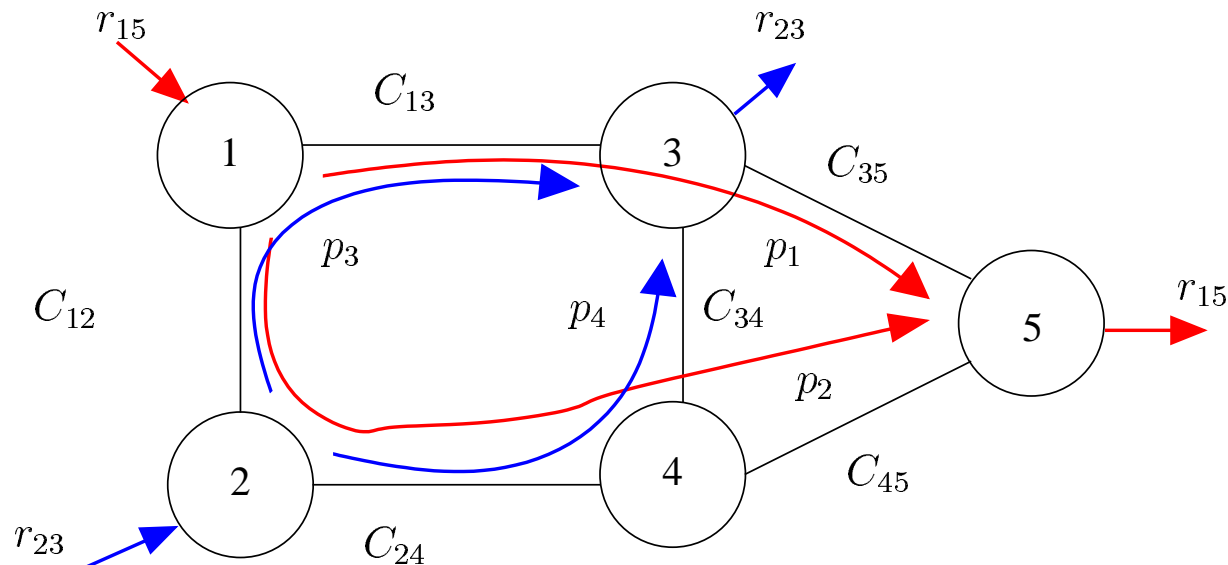
● Exemple - Minimiser :

$$\frac{F_{13}}{C_{13} - F_{13}} + \frac{F_{12}}{C_{12} - F_{12}} + \frac{F_{24}}{C_{24} - F_{24}} + \frac{F_{34}}{C_{34} - F_{34}} + \frac{F_{35}}{C_{33} - F_{33}} + \frac{F_{45}}{C_{45} - F_{45}}$$

où

$$\begin{aligned} F_{13} &= x_{p_1} + x_{p_3} & F_{12} &= x_{p_2} + x_{p_3} & F_{24} &= x_{p_2} + x_{p_4} \\ F_{34} &= x_{p_4} & F_{35} &= x_{p_1} & F_{45} &= x_{p_2} \end{aligned}$$

sous les contraintes : $x_{p_1} + x_{p_2} = r_{15}$ et $x_{p_3} + x_{p_4} = r_{23}$ avec $x_{p_i} \geq 0$





Caractérisation du Routage Optimal

Caractérisation du Routage Optimal

● Pourquoi ?

- Pour montrer que le routage optimal propage les flots sur des plus courts chemins au sens de certaines métriques,
- Ces métriques dépendent du trafic sur les liens.
- Cette caractéristique intéressante du routage optimal est à la base des algorithmes que nous verrons dans la suite.

Caractérisation du Routage Optimal

● Notations :

- Coût associé à une solution $x = (x_p)_{p \in P_w, w \in W}$:

$$D(x) = \sum_{(i,j)} D_{ij} \left[\sum_{p/(i,j) \in p} x_p \right]$$

- On note D'_{ij} la dérivée du coût D_{ij} (par rapport au trafic F_{ij} du lien (i, j))

Caractérisation du Routage Optimal

● Analyse des variations du coût

- Dérivons le coût $D(x)$ par rapport à x_p ,

$$\begin{aligned}\frac{\partial D(x)}{\partial x_p} &= \sum_{(i,j)} \frac{\partial D_{ij}}{\partial x_p} \left[\sum_{q/(i,j) \in q} x_q \right] \\ &= \sum_{(i,j)} \frac{\partial D_{ij}}{\partial F_{ij}} \frac{\partial F_{ij}}{\partial x_p} \left[\sum_{q/(i,j) \in q} x_q \right]\end{aligned}$$

- Or $\partial F_{ij} / \partial x_p = 1$ si $(i, j) \in p$ et 0 sinon. Par conséquent,

$$\frac{\partial D(x)}{\partial x_p} = \sum_{(i,j) \in p} D'_{ij}$$

où les dérivées D'_{ij} sont évaluées pour les flots F_{ij} obtenus avec la solution x .

Caractérisation du Routage Optimal

- **Interprétation des dérivées partielles du coût**

$$\frac{\partial D(x)}{\partial x_p} = \sum_{(i,j) \in p} D'_{ij}$$

- $\partial D(x)/\partial x_p$ apparaît comme la longueur du chemin p quand les métriques des liens du chemin sont prises égales aux dérivées premières D'_{ij} évaluées en x .
- $\partial D(x)/\partial x_p$ est appelé la longueur du chemin p au sens des dérivées premières.

Caractérisation du Routage Optimal

● Condition d'optimalité

- Soit $x^* = \{x_p^*\}$ une solution optimale. Considérons une demande OD w et un chemin $p \in P_w$ tel que $x_p^* > 0$.
- En déviant une partie δ du flot du chemin p sur un autre chemin p' , le coût ne peut diminuer car la solution est optimale. Au premier ordre, la variation du coût est

$$\delta \frac{\partial D(x^*)}{\partial x_{p'}} - \delta \frac{\partial D(x^*)}{\partial x_p} \geq 0$$

- Condition d'optimalité :

$$x_p^* > 0 \quad \Rightarrow \quad \frac{\partial D(x^*)}{\partial x_{p'}} \geq \frac{\partial D(x^*)}{\partial x_p} \quad \forall p' \in P_w \quad (7)$$

Caractérisation du Routage Optimal

● Interprétation de la condition d'optimalité

$$x_p^* > 0 \quad \Rightarrow \quad \frac{\partial D(x^*)}{\partial x_{p'}} \geq \frac{\partial D(x^*)}{\partial x_p} \quad \forall p' \in P_w$$

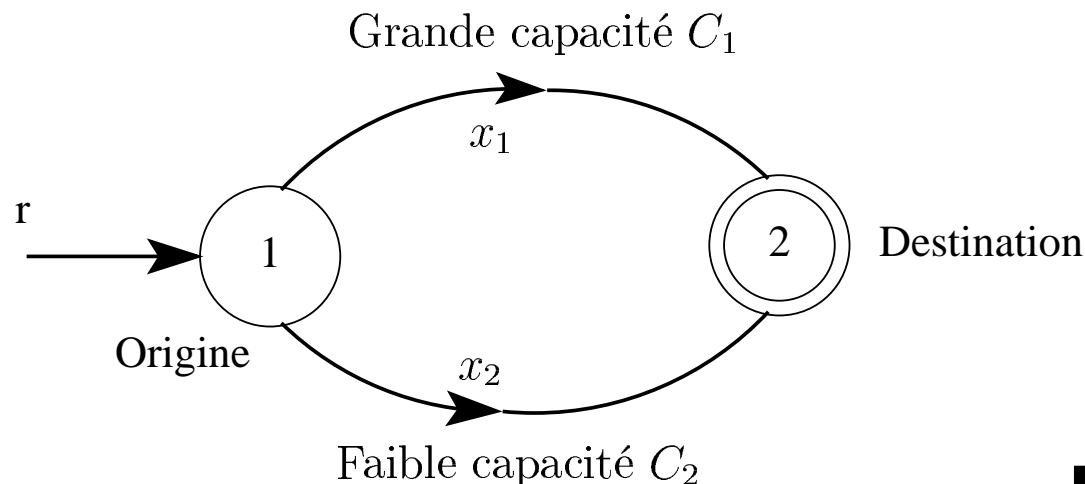
- La solution optimale ne propage du flot sur le chemin p que si c'est un chemin de longueur minimale au sens des dérivées premières.
- La solution optimale ne répartit une demande r_w sur plusieurs chemins que si ces chemins sont de longueurs égales (et minimales) au sens des dérivées premières.
- C'est une condition nécessaire. Elle est suffisante si les fonctions coûts D_{ij} sont convexes.

Caractérisation du Routage Optimal

● Exemple

- Réseau à deux liens de capacités C_1 et C_2 , $C_1 \geq C_2$.
- Le noeud 1 est la seule origine et le noeud 2 la seule destination.
- Le trafic r entre 1 et 2 doit être réparti entre deux flots x_1 et x_2 pour minimiser la fonction coût,

$$D(x) = D_1(x) + D_2(x) \quad \text{où} \quad D_i(x) = \frac{x_i}{C_i - x_i} \quad i = 1, 2$$



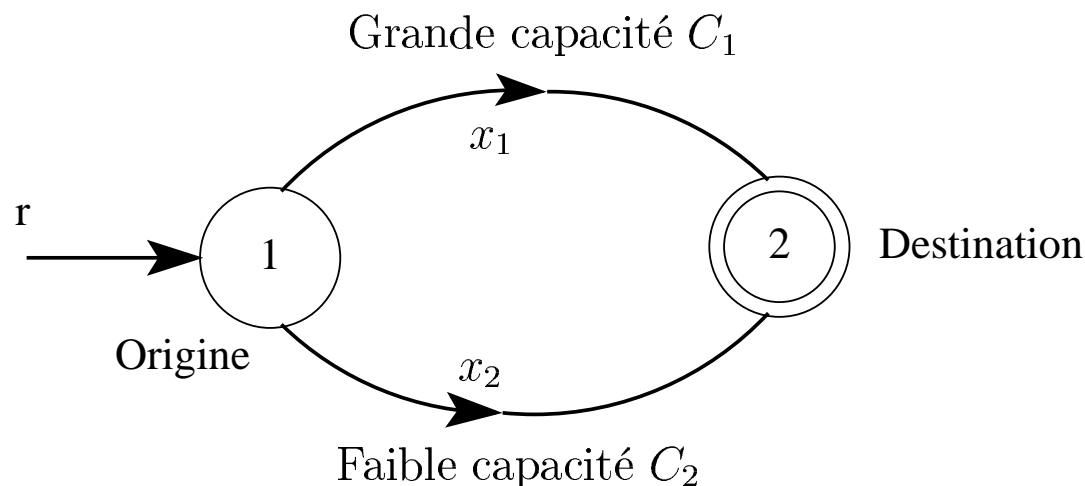
Caractérisation du Routage Optimal

● Exemple

- A l'optimum, la condition d'optimalité doit être vérifiée de même que les contraintes,

$$x_1^* + x_2^* = r, \quad x_1^* \geq 0, \quad x_2^* \geq 0$$

- Le flot x_1 ne peut être inférieur à x_2 ($C_1 \geq C_2$). On a donc 2 cas :
 - Cas où $x_1^* = r$ et $x_2^* = 0$,
 - Cas où $x_1^* > 0$ et $x_2^* > 0$



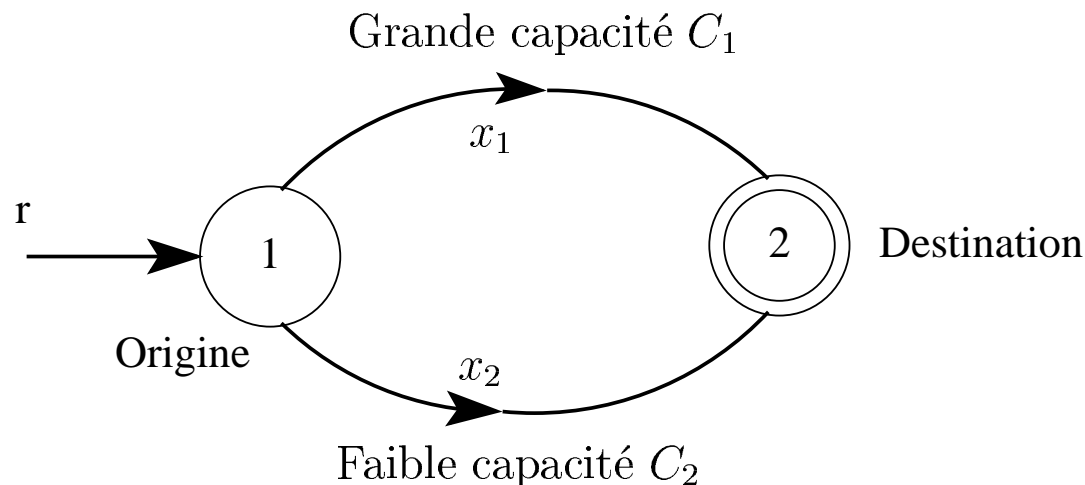
Caractérisation du Routage Optimal

● Exemple : cas où $x_1^* = r$ et $x_2^* = 0$

- D'après la condition d'optimalité, on a $dD_1(r)/dx_1 \leq dD_2(0)/dx_2$
- Sachant que $D'_i(x_i) = C_i/(C_i - x_i)^2$, on a donc :

$$\frac{C_1}{(C_1 - r)^2} \leq \frac{1}{C_2}$$

- Ce qui implique que $r \leq C_1 - \sqrt{C_1 C_2}$.



Caractérisation du Routage Optimal

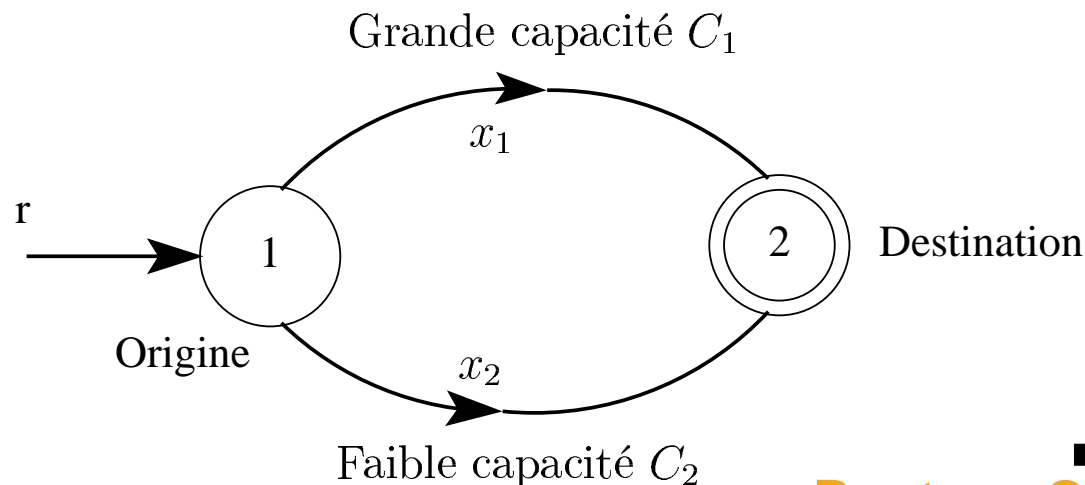
● Exemple : cas où $x_1^* > 0$ et $x_2^* > 0$

- Condition d'optimalité : $dD_1(x_1^*)/dx_1 = dD_2(x_2^*)/dx_2$

$$\frac{C_1}{(C_1 - x_1^*)^2} \leq \frac{C_2}{(C_2 - x_2^*)^2}$$

- Cette équation et la contrainte $x_1^* + x_2^* = r$ donnent :

$$x_1^* = \frac{\sqrt{C_1} [r - (C_2 - \sqrt{C_1 C_2})]}{\sqrt{C_1} + \sqrt{C_2}} \quad \text{et} \quad x_2^* = \frac{\sqrt{C_2} [r - (C_1 - \sqrt{C_1 C_2})]}{\sqrt{C_1} + \sqrt{C_2}}$$



Caractérisation du Routage Optimal

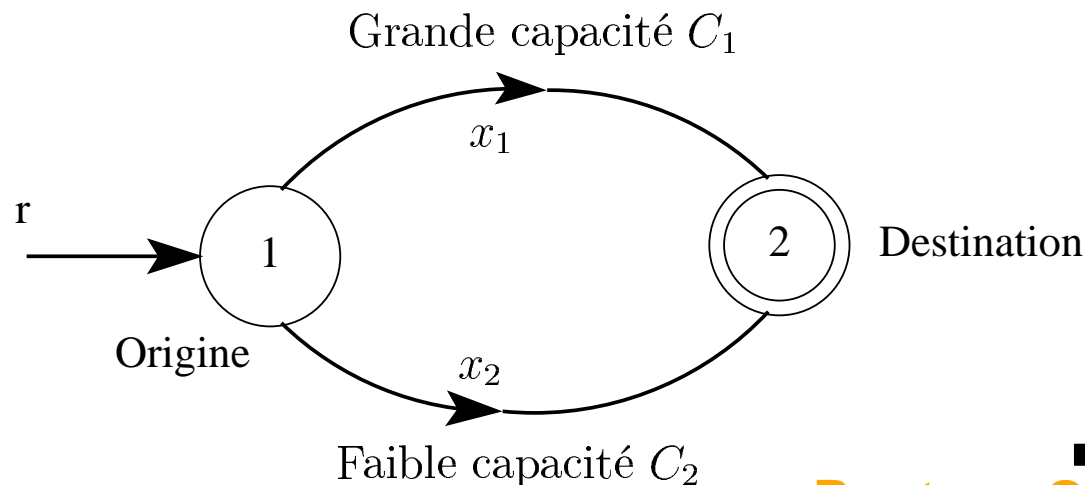
● Exemple : cas où $x_1^* > 0$ et $x_2^* > 0$


- Condition d'optimalité : $dD_1(x_1^*)/dx_1 = dD_2(x_2^*)/dx_2$

$$\frac{C_1}{(C_1 - x_1^*)^2} \leq \frac{C_2}{(C_2 - x_2^*)^2}$$

- Cette équation et la contrainte $x_1^* + x_2^* = r$ donnent :

$$x_1^* = \frac{\sqrt{C_1} [r - (C_2 - \sqrt{C_1 C_2})]}{\sqrt{C_1} + \sqrt{C_2}} \quad \text{et} \quad x_2^* = \frac{\sqrt{C_2} [r - (C_1 - \sqrt{C_1 C_2})]}{\sqrt{C_1} + \sqrt{C_2}}$$





Méthodes de directions admissibles pour le routage optimal

Méthodes de directions admissibles

● Que nous dit la condition d'optimalité ?

- La solution optimale ne partage chaque demande que sur les chemins de longueur minimale au sens des dérivées premières.
- Une solution ne peut pas être optimale si une partie de la demande est routée sur un chemin de longueur non minimale.
- Cela suggère qu'une solution non optimale peut être améliorée en déviant une partie du flot des chemins de longueurs non minimales vers les chemins de longueur minimale.
- Les méthodes de directions admissibles se basent sur cette idée
 - Elles calculent la solution optimale de routage itérativement.
 - A chaque itération, elles font décroître le coût de la solution courante en déviant du flot des chemins de longueurs non minimales vers les chemins de longueur minimale.

Méthodes de directions admissibles

● Principe

- **Considérons un vecteur solution $x = \{x_p\}$ admissible, c'est à dire vérifiant les contraintes.**
- **Soit la solution $x' = x + \beta \Delta x$ où $\Delta x = \{\Delta x_p\}$ est une direction et β le pas que l'on fait dans cette direction pour modifier x .**
- **Cette solution x' nous intéresse si elle remplit deux conditions :**
 - **La solution x' est admissible.**
 - **Cette solution fait décroître le coût, i.e. $D(x') \leq D(x)$.**

Méthodes de directions admissibles

● Condition d'admissibilité

- Δx est une direction admissible si x' est admissible,

$$\begin{aligned}\sum_{p \in P_w} x'_p &= \sum_{p \in P_w} x_p + \beta \Delta x_p \\ &= \sum_{p \in P_w} x_p + \beta \sum_{p \in P_w} \Delta x_p \\ &= r_w\end{aligned}$$

- Puisque $\sum_{p \in P_w} x_p = r_w$, la condition pour que la direction Δx soit admissible est que :

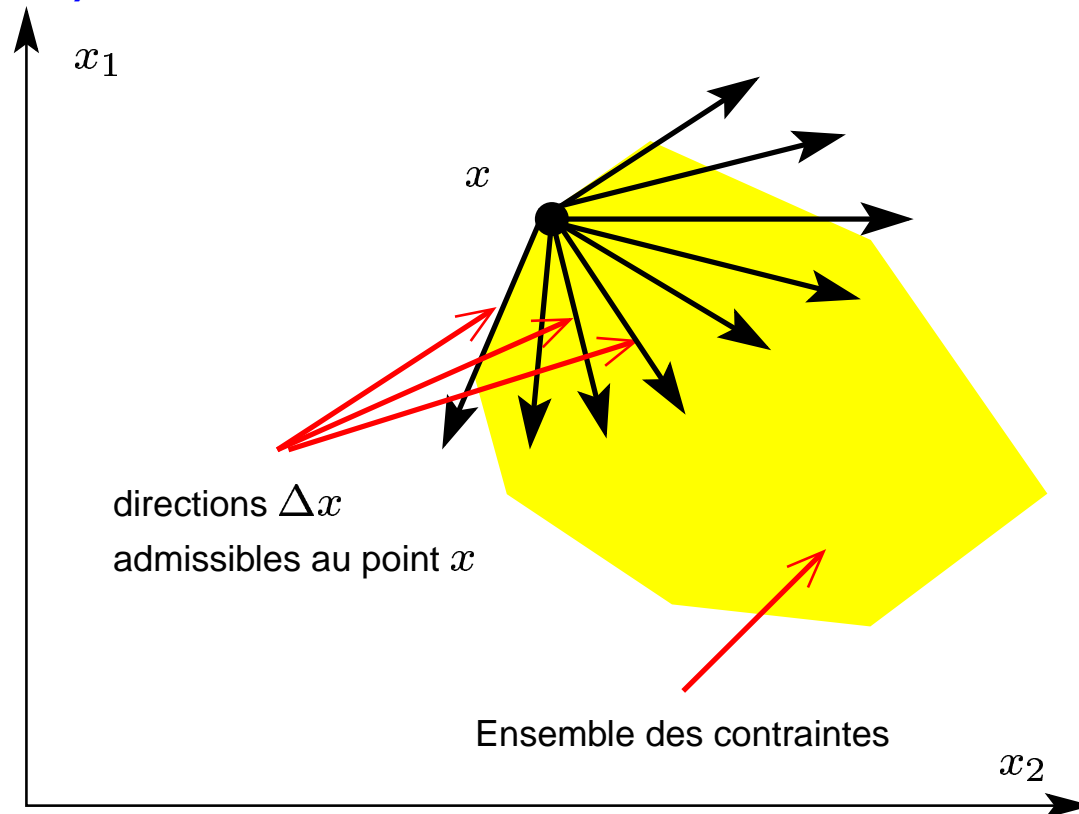
$$\sum_{p \in P_w} \Delta x_p = 0 \quad \forall w \in W \quad (8)$$

- Cette condition exprime tout simplement que pour maintenir l'admissibilité à partir de la solution x , toute augmentation du flot sur certains chemins doit être compensée par des diminutions du flot sur d'autres chemins.

Méthodes de directions admissibles

● Illustration des directions admissibles

- Δx est une direction admissible si $x' = x + \beta \Delta x$ est admissible (pour β petit),



Méthodes de directions admissibles

● Condition de descente

- Δx est une direction de descente si $D(x') \leq D(x)$,
- Considérons le gradient $\nabla D(x)$ de $D(x)$ au point x

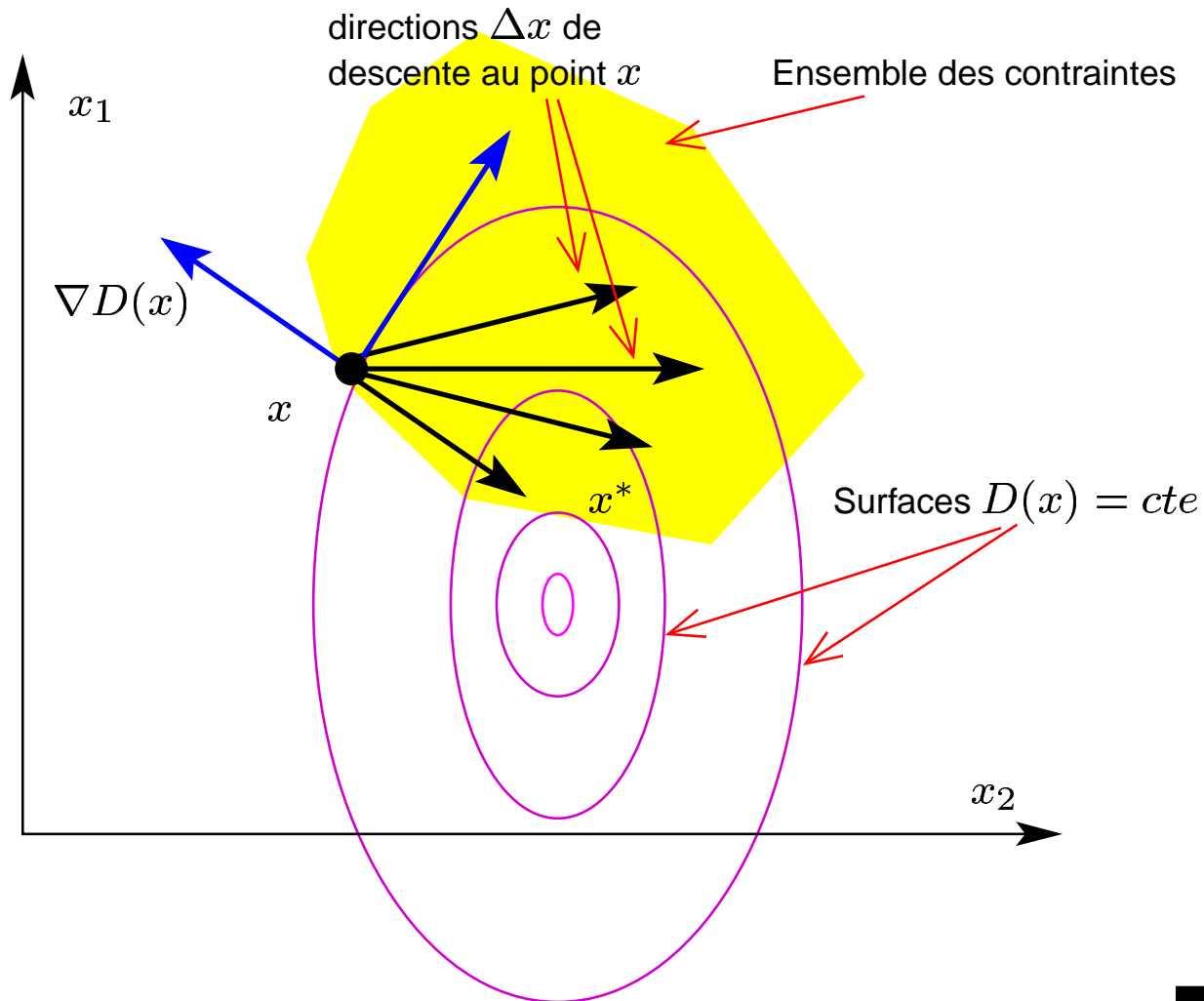
$$\nabla D(x) = \left[\frac{\partial D(x)}{\partial x_1}, \dots, \frac{\partial D(x)}{\partial x_n} \right]^T$$

- Le gradient $\nabla D(x)$ est normal aux surfaces isocoût $D(x) = cte$,
- Le gradient $\nabla D(x)$ représente la direction d'augmentation du coût
- Les directions de descente sont “en sens inverse” du gradient
- La condition de descente s'exprime par le fait que le produit scalaire de $\nabla D(x)$ et de Δx doit être négatif, i.e.

$$\sum_{w \in W} \sum_{p \in P_w} \frac{\partial D(x)}{\partial x_p} \Delta x_p < 0$$

Méthodes de directions admissibles

● Illustration des directions de descente admissibles



Méthodes de directions admissibles

● Génération de directions de descente admissibles

● Si Δx est une direction admissible, i.e. $\sum_{p \in P_w} \Delta x_p = 0$, il suffit de prendre,

(a) $\Delta x_p \leq 0$ si p n'est pas un plus court chemin au sens des dérivées premières,

(b) $\Delta x_p < 0$ pour au moins un chemin non optimal p .

Méthodes de directions admissibles

- **Algorithme des méthodes de directions admissibles**

- L'itération de base est de la forme :

$$x := x + \beta \Delta x$$

- où Δx est une direction de descente admissible et β est un pas positif dans cette direction tel que,

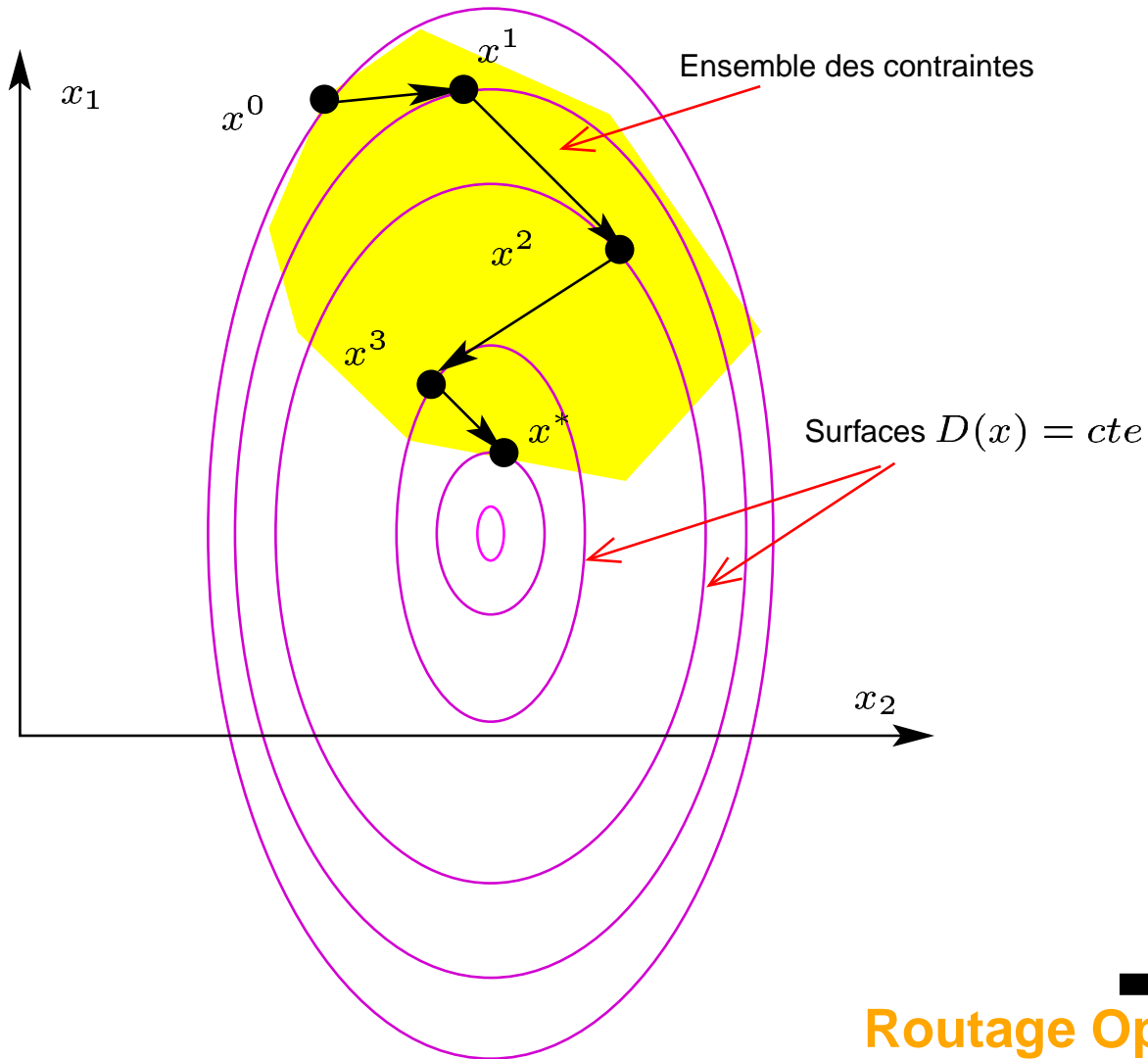
$$D(x + \beta \Delta x) < D(x)$$

et que la solution $x + \beta \Delta x$ est admissible.

- Le pas β peut être différent à chaque itération.

Méthodes de directions admissibles

● Illustration du processus d'optimisation





Méthode de Frank-Wolfe (Flow Deviation)

Flow Deviation

● Algorithme de Flow Deviation

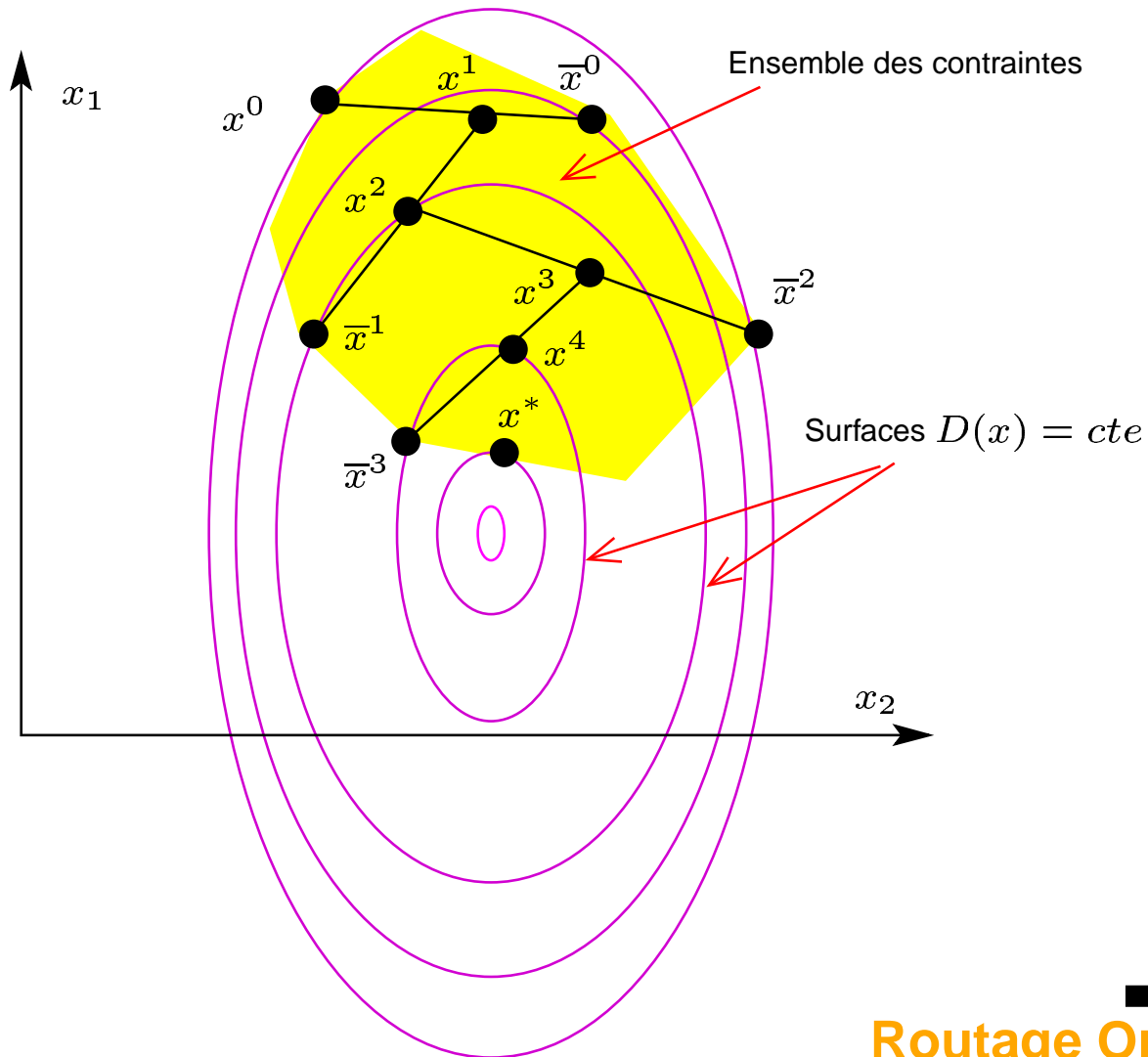
- Soit $x = \{x_p\}$ une solution admissible.
- Trouver pour chaque couple OD w un chemin de longueur minimale au sens des dérivées premières
- Soit $\bar{x} = \{\bar{x}_p\}$ la solution obtenue en routant toutes les demandes r_w sur ces plus courts chemins
- Mettre à jour la solution x avec

$$x_p := x_p + \beta (\bar{x}_p - x_p)$$

- Le pas β peut être choisi de manière à minimiser le coût de la nouvelle solution,
- On peut utiliser un pas adaptatif, e.g. $\beta := 0.8 \times \beta$ si le coût augmente et $\beta := 1.2 \times \beta$ s'il diminue
- On peut aussi utiliser un pas fixe (à éviter)

Flow Deviation

● Illustration de Flow Deviation



Flow Deviation

● Caractéristiques de Flow Deviation

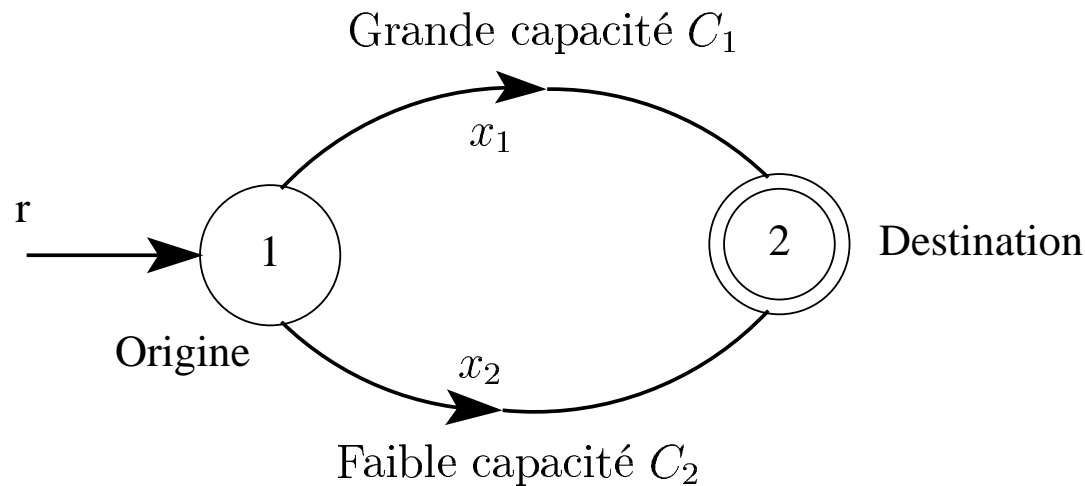
- **L'algorithme converge vers la solution optimale de partage de charge**
 - Mais la convergence est de plus en plus lente au fur et à mesure qu'on se rapproche de l'optimum
 - Les directions de descente générées ont tendance à devenir orthogonales à la direction vers l'optimum, d'où une progression en zig-zag
- **Les flots sont déviés vers les chemins les plus courts au sens des dérivées premières dans des proportions égales**
 - C'est la différence majeure avec les méthodes de projection (gradient projeté, gradients conjugués)
 - Avec ces méthodes, la quantité de flot déviée sur un plus court chemin est différente pour chaque couple OD w

Flow Deviation

● Exemple

- Capacités $C_1 = 20$ et $C_2 = 10$.
- Le trafic $r = 8$
- Solution initiale : $x^0 = (4, 4)$ de coût

$$D(x) = \frac{x_1}{C_1 - x_1} + \frac{x_2}{C_2 - x_2} = \frac{4}{20 - 4} + \frac{4}{10 - 4} = 0.916$$



Flow Deviation

● Exemple : itération 1

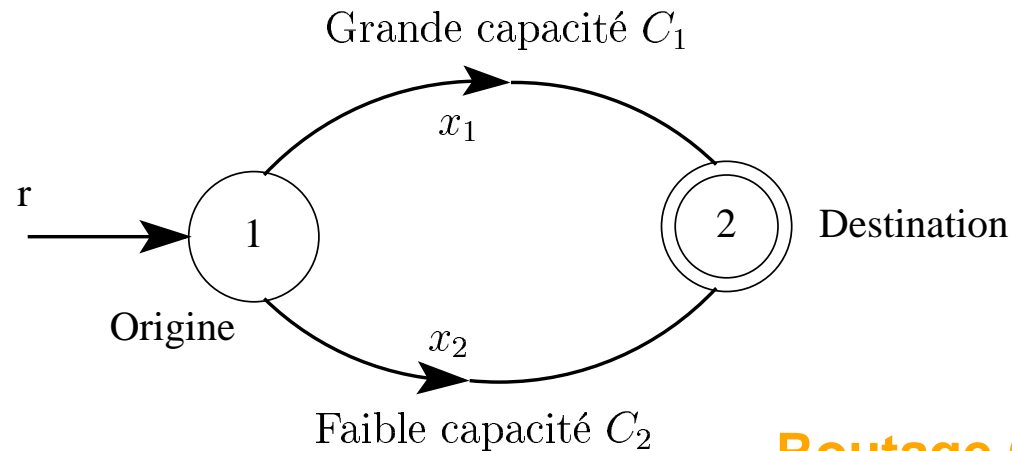
- Calcul du point extrême \bar{x}^0 .

$$\frac{\partial D(x^0)}{\partial x_1} = 0.0781 \text{ et } \frac{\partial D(x^0)}{\partial x_2} = 0.2777 \Rightarrow \bar{x}^0 = (8, 0)$$

- Avec un pas $\beta = 0.5$, la nouvelle solution est:

$$x^1 = x^0 + \beta (\bar{x}^0 - x^0) = (4, 4) + 0.5 \times (4, -4) = (6, 2)$$

- Coût : $D(x^1) = 0.679$



Flow Deviation

● Exemple : itération 2

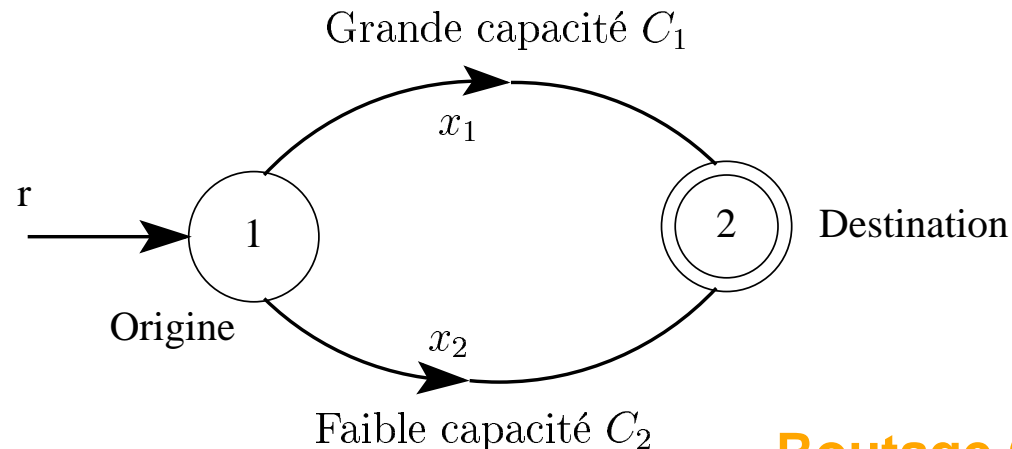
- Calcul du point extrême \bar{x}^1 .

$$\frac{\partial D(x^1)}{\partial x_1} = 0.102 \text{ et } \frac{\partial D(x^1)}{\partial x_2} = 0.156 \Rightarrow \bar{x}^1 = (8, 0)$$

- Avec un pas $\beta = 0.6$, la nouvelle solution est:

$$x^2 = x^1 + \beta (\bar{x}^1 - x^1) = (6, 2) + 0.6 \times (2, -2) = (7.2, 0.8)$$

- Coût : $D(x^2) = 0.649$



Flow Deviation

● Exemple : itération 3

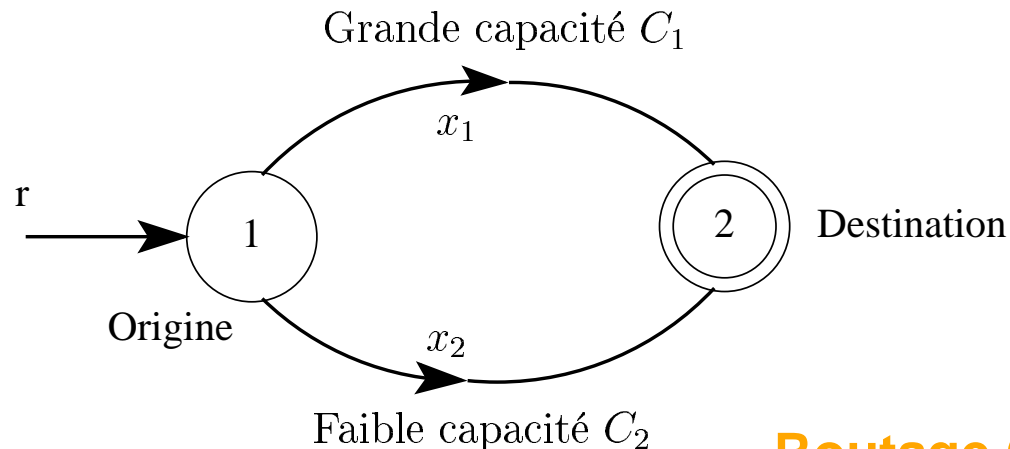
- Calcul du point extrême \bar{x}^2 .

$$\frac{\partial D(x^2)}{\partial x_1} = 0.122 \text{ et } \frac{\partial D(x^2)}{\partial x_2} = 0.118 \Rightarrow \bar{x}^2 = (0, 8)$$

- Avec un pas $\beta = 0.01$, la nouvelle solution est:

$$x^3 = x^2 + \beta (\bar{x}^2 - x^2) = (7.2, 0.8) + 0.01 \times (-7.2, 7.2) = (7.128, 0.872)$$

- La solution optimale est $x^* = (7.112, 0.888)$





Méthode du Gradient Projeté

Méthode du Gradient Projeté

● Algorithme du Gradient Projeté

- Soit $x = \{x_p\}$ une solution admissible.
- Faire :

$$x := x + \beta \Delta x \quad \text{où} \quad \Delta x_p = \frac{\partial D(x)}{\partial x_p} - \frac{1}{|P_w| - 1} \sum_{q \neq p} \frac{\partial D(x)}{\partial x_q}$$

- Le pas β peut être choisi de manière à minimiser le coût de la nouvelle solution,
- On peut utiliser un pas adaptatif, e.g. $\beta := 0.8 \times \beta$ si le coût augmente et $\beta := 1.2 \times \beta$ s'il diminue
- On peut aussi utiliser un pas fixe (à éviter)

Méthode du Gradient Projeté

● Caractéristiques du gradient projeté

- L'algorithme converge vers la solution optimale de partage de charge
 - Mais la convergence est de plus en plus lente au fur et à mesure qu'on se rapproche de l'optimum
 - Le gradient projeté reste quand même beaucoup plus efficace que Flow Deviation