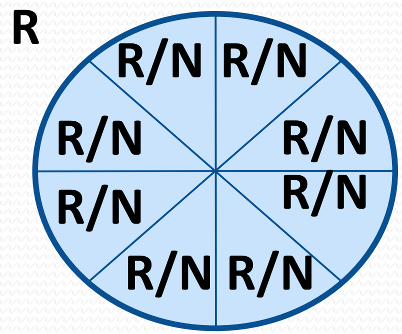
Master-MIRI
Topics on Optimization and Machine Learning (TOML)

José M. Barceló Ordinas
Departament d'Arquitectura de Computadors
(UPC)

- **COP applications to networking**
  - Many networking problems can be solve using optimization tools.
    - MAC protocols in WSN (Project)
    - Resource Allocation: Network Utility problems
    - Network Resource Allocation: Power control
    - Queueing Theory examples
    - **Textbook:** papers

- **Resource allocation**
  - Let us assume that an authority has a resource of size R that wants to divide between N users. What kind of resource allocation can we apply ?
    - E.g., we can divide the resource R in N equally parts, it is to say, give R/N amount to each user



    - E.g., What if the users valuate their allocations in a different way ?

Resource allocation

- Let us define the **utility** of user x as **U(x)** and we define it as the measure of value that a user gives to an allocation.

Satisfaction from increasing from 10$→20$ is not the same than passing from 100$→110$ (10% of increment) or passing from 1000$→1010$ (1% of increment) or passing from 10000$→10010$ (0.1% of increment).

Utility functions are strictly concave functions that represent the degree of satisfaction of a user.

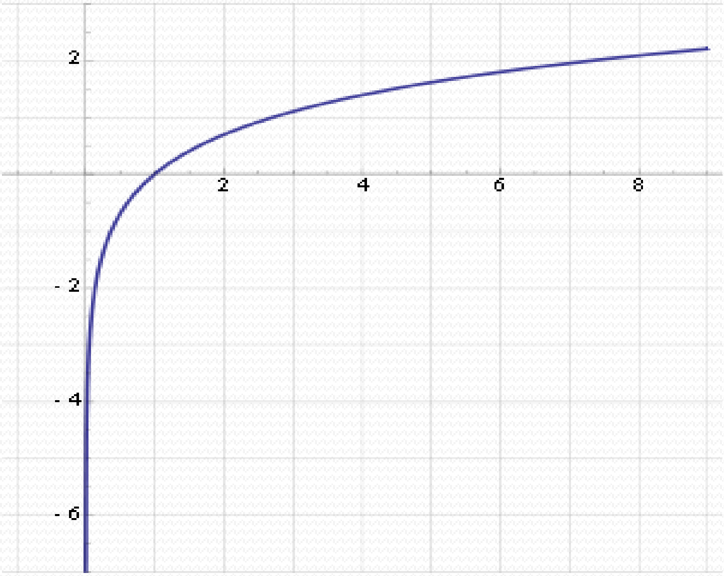
- How can we divide the resource R in such a way that we maximize the **utility function of a user** ?
- How can we divide the resource R in such a way that we maximize the **social welfare** (sum of users utilities) ?

Utility Maximization

- Let us define $U_i(x_i)$ as the utility that user i=1,…, N obtains after being assigned an allocation x_i of resource R → concave function
  - $U_i(x_i)= \log x_i$ for user i, would be an example of a **concave utility function**.
  - $W(x_1, …, x_N) = \Sigma_{i=1,..N} U_i(x_i) = \Sigma_{i=1,..N} \log x_i$ would be a **welfare function**
- Let us consider a resource R, e.g. a link with capacity C.
  - a user feels more utility of a rate increase from 1Kbps to 100 Kbps than from 1 Mbps to 1.1 Mbps

U_i(1 Kbs) = 3.0 units
U_i(100 Kbs) = 4.0 units

U_i(1 Mbs) = 6.0 units
U_i(1.1 Mbs) = 6.041 units



- Remember that allocations aim to maximize utilities and therefore social welfare (the sum of utilities). **How to measure whether we have a good allocation ?**

- ## Resource utility:
  - **Definition:** a measure to determine whether all the resource is used.
  - It is measured as the ratio of sum of allocation by the resource capacity. It is to say $\mathbf{0 \leq \rho = (\Sigma_{i=1,..N}\, x_i\,)/R \leq 1}$. We aim to $\rho = 1$,
  - For example, a link with capacity C=16 Mbps, and 4 users. The allocations are 2.5 Mbps for each user, resulting in a total of 10 Mbps ≤ 16 Mbps. It's network utilization is low since it is of $\boldsymbol{\rho} = 10/16 = 0.625 < 1$.

- ## Pareto efficiency:
  - **Definition:** a measure to determine that the allocation is optimal. That means that there is no another allocation that produces better utilities.
  - For example, a link with capacity C=16 Mbps, and 4 users. The allocations are 2.5 Mbps for each user, resulting in a total of 10 Mbps ≤ 16 Mbps. It is not efficient, since there are other allocations more efficient.

- **Fairness in networks:**
  - **Definition:** a measure to determine whether users or applications are receiving a fair share of system resources.

Jain index: measures how fair is the allocation. **$0 \le JI \le 1$**, the more near 1, the more fair is the allocation.

$$JI = F(x_1,\ldots,x_n) = (\Sigma_i\, x_i)^2/(n\, \Sigma_i\, x_i^2) \le 1$$

- **For example:** 4 users share a link with capacity C=16 Mbps and they have demands of c_1= 2 Mbps, c_2= 4 Mbps, c_3= 10 Mbps, c_4=15 Mbps.
  - Uniform allocation C/N = 4 Mbps → JI = 1, but it is fair, however it is not efficient since user 1 needs 2 Mbps and receives 4 Mbps, thus 2 Mbps are unused, and sure that there are other allocations more efficient,
  - allocate beginning from the least demanding user so allocation is x =(2,4,10,0) → and JI = $(2+4+10+0)^2/(4\cdot(4+16+100+0))$= 256/480 = 0.533, not too efficient

Convex Optimization Problems. Applications

Maximum Throughput allocation (scheduling): schedule packets to the less demanding flows until the resource is consumed (e.g. CPU, capacity, etc).

The most demanding ones will suffer “starvation” since they will have to wait until the less demanding end (e.g. have no data to transfer).

Customer satisfaction: low, since most of them suffer starvation (low Jain index)

Network operator satisfaction: low, since the most demanding ones typically pay more, thus, profit is not maximized

Network (resource) utilization: in general is optimal since the resource is fully used.

Convex Optimization Problems. Applications

- In general, we can express the demands of each user as a linear system of equations, it is to say as an affine linear system. For example, a link with capacity C=16 Mbps and they have demands of c_1= 2 Mbps, c_2= 4 Mbps, c_3= 10 Mbps, c_4=15 Mbps:

$$x_i \leq c_i \qquad i=1,\ldots, N$$
$$\Sigma_i x_i \leq C$$

Max-min fair allocation: a vector **x*** (x_i, $i \in 1,\ldots, N$) of allocations is max-min fair if:

(i) it is a feasible set of solutions (i.e. $x \geq 0$ and $Ax \leq b$), and

(ii) if for each i=1,…, N, x^*_i can not be increased, while maintaining the feasibility, without decreasing x_i for some i for which $x_i \leq x^*_i$

- The max-min fair allocation gives more priority to smaller flows (see Kelly paper), in the sense that if $x_i < x^*_i$ then no increase in x^*_i, no matter how large, can compensate for any decrease in x_i, no matter how small.
- In general, max-min provides better fairness than maximum throughput schedulers since shares the resource among all users (avoids starvation)

Frank Kelly et al "Rate control for communication networks: shadow prices, proportional sharing and stability

- **Example of max-min fair allocation**: 4 users share a link with capacity C=16 Mbps and they have demands of c_1= 2 Mbps, c_2= 4 Mbps, c_3= 10 Mbps, c_4=15 Mbps.
  - Label users with demands from lowest to highest. Assign $x_i = C/N$ (uniform allocation) bit/s to each user.
    - If $x_1 = C/N \leq c_1$, stop
    - If $x_1 = C/N > c_1$, add (x_1-c_1) to the rest of users 2…N, so, they receive

      $x_i = C/N + (x_1-c_1)/(N-1)$, with i= 2, …, N and $x_1 = c_1$,
  - If with this assignment, user 2 receives more than its demand c_2, give the excess to users 3…N, and iterate
  - For example:
    1. $x_1 = x_2 = x_3 = x_4 = C/N = 4$. Since $x_1 > c_1$, then
    2. $x_1 = c_1 = 2$; $x_2 = x_3 = x_4 = 4 + (4-2)/3 = 4 + 2/3 = 4.66 > c_2$,
    3. $x_1 = 2$; $x_2 = c_2 = 4$; $x_3 = x_4 = 4.66 + 0.66/2 = 5 < c_3$,

Final ***max-min fairness*** assignment: $x_1 = 2$; $x_2 = 4$; $x_3 = 5$; $x_4 = 5$.

$$JI=F(x_1,\ldots,x_n)= (2+4+5+5)^2/(4\cdot(4+16+25+25))= 256/280 = 0.914$$

Convex Optimization Problems. Applications

Let be the following optimization problem (COP):

$$\begin{array}{lll} \text{Max} & \sum_i U(x_i) = \sum_i m_i \log(x_i) & \\ \text{s.t.} & Ax \leq b & \\ & x \geq 0 & \end{array} \left.\vphantom{\begin{array}{l}a\\a\\a\end{array}}\right\} \quad x^*_i = m_i / (\sum_{r \in S} \lambda_r)$$

Proportionally fair allocation: a vector of rates **x*** (x^*_i, i=1,..,N) is proportionally fair if

(i) it is a feasible set of solutions (i.e. $x \geq 0$ and $Ax \leq b$), and

(ii) if for any other feasible vector **x** (x_i, i=1,..N) the aggregate of proportional changes is zero or negative:

$\nabla^T U(x^*_i)\,(x-x^*) = \sum_{i=1,\dots,N} (x_i - x_i^*)/x_i^* \leq 0$ (optimality condition for concave functions)

- The proportionally fair allocation also gives more priority to smaller flows, but less emphatically. It says that if user i increases its allocation there will be at least one other user whose rate will decrease, and further, the proportion by which it decreases will be larger than the proportion by which the rate increases for user i.
- It also avoids starvation as the max-min fair allocation,
- If $U(x_i)=\log(x_i)$ then, $x^*_i = m_i / (\sum_{r \in S} \lambda_r)$ **satisfies** the proportionally fair allocation

Solve the optimization problem of the example:

$$
\begin{array}{ll}
\text{minimize} & \sum_i U(x_i) = -\log(x_1) - \log(x_2) - \log(x_3) - \log(x_4) \\
\text{s.t.} & x_1 \leq 2 \\
 & x_2 \leq 4 \\
 & x_3 \leq 10 \\
 & x_4 \leq 15 \\
 & x_1 + x_2 + x_3 + x_4 \leq 16 \\
 & x_1, x_2, x_3, x_4 \geq 0 \\
\text{var} & x_1, x_2, x_3, x_4
\end{array}
$$

Solving the optimization problem (try with CVXPY) you will get: x_1=2 Mbps, x_2=4 Mbps, x_3= 5 Mbps, and x_4=5 Mbps,

In this case, the result is the same than the max-min fair allocation, thus in this specific case, the max-min also gives the optimal allocation

Schedulers used in routers, switches, servers, etc to achieve these fairness policies:

- **Maximum Throughput:** requires the evaluation of cost functions
- **Max-min:** uses Round Robin and Fair Queueing (FQ). This last assigns a queue to each flow and round-robins over the queues with packets.
- **Proportional sharing:** uses Weighted Fair Queueing (WFQ) that assigns a weight e,g,. $(\phi_i/\sum\phi_i)R$ to each flow (here R is the capacity of the link). If all weights are equal, then WFQ = FQ.

Network Utility Maximization

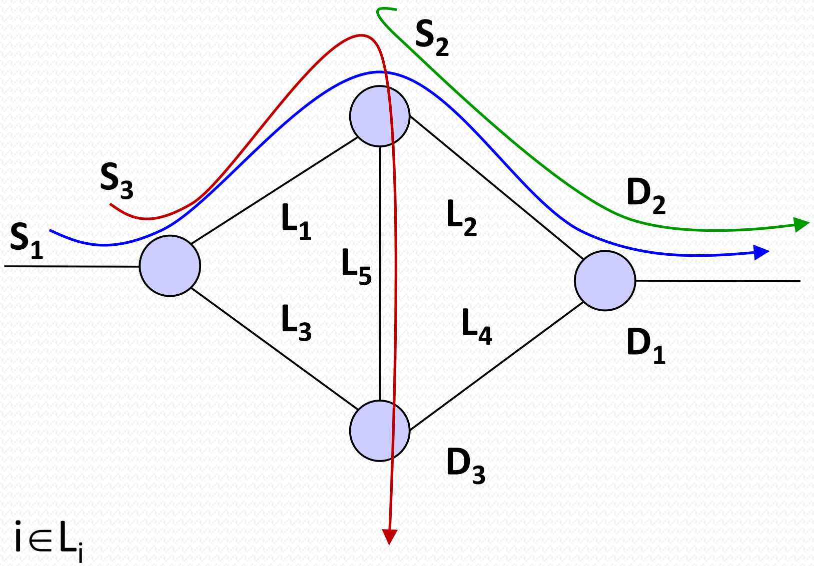
- Let be a network of routers with a set of traffic sources $S= \{S_1,\ldots,S_n\}$ and a set of links $L=\{L_1,\ldots,L_k\}$. Each link has a capacity c_l and each source S_r has a route R, such as the source $s_i \subset R_s$ transmits at some rate x_s (bps).
- For example, if source s_1 traverses links L_1 and L_2, then its route $R_1= L_1 \cup L_2$.
- Define $\mathbf{x_s}$ as the amount of rate (bps) that a source receives on route R_1,
- Let us define now the utility $\mathbf{U_s(x_s) = \log x_s}$ that a source obtains from transmitting on route s at rate x_s → concave function

The Network Utility Optimization Problem is:

Maximize:	$\sum_{s\in S} U_s(x_s)$	
subject to:	$\sum_{s:l\in R_s} x_s \leq c_l$	for all $l\in L$
	$x_s \geq 0$	for all $s\in S$

Network Utility Maximization

- **Example:** lthere are 5 resources: links L_1, L_2, …, L_5 with capacities C_1, C_2, …, C_5 and 3 sources and then 3 routes $\mathbf{R_1}$**:** $L_1 \cup L_2$, $\mathbf{R_2}$**:** L_2, $\mathbf{R_3}$**:** $L_1 \cup L_5$.
- Let us represent the following network utility maximization convex optimization problem:



Max $\quad \sum_{s\in S} U_s(x_s) = \sum_{s\in S} \log (x_s)$

s.t. $\quad Ax \le C$

$\quad x \ge 0$

With A a matrix with coefficients:

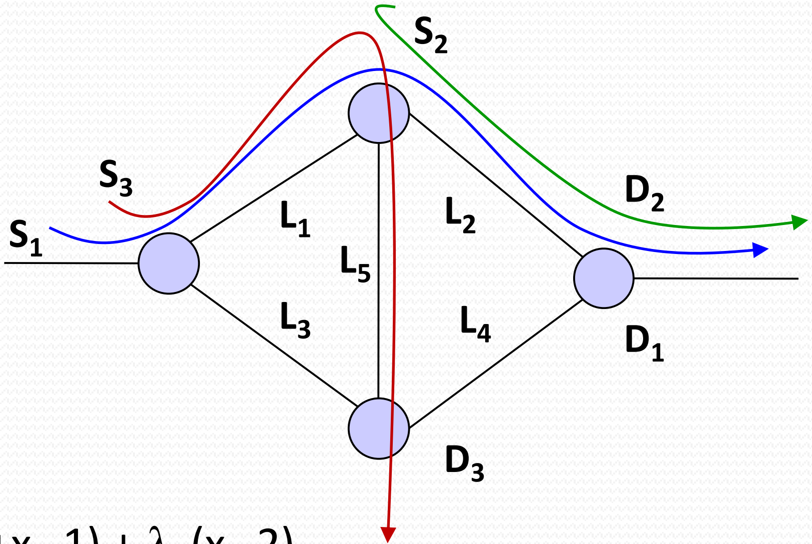
$a_{ij}=1$ if source j lies on link $i\in L_i$

$a_{ij}=0$ if source j does not lie on link $i\in L_i$

Columns j of A represent routes R (which links) source j use. **Rows i of A** represent which sources j occupy link i.

- **Network Utility Maximization**
  - **Example:** links with C_1=C_3=C_5=1 and C_2=C_4=2 capacity units

$$
\begin{array}{llll}
\text{maximize} & \log x_1 + \log x_2 + \log x_3 & & \\
\text{s.t.} & x_1 + \quad\quad x_3 & \leq 1 & \rightarrow L_1 \\
 & x_1 + \; x_2 & \leq 2 & \rightarrow L_2 \\
 & \quad\quad\quad x_3 & \leq 1 & \rightarrow L_5 \\
 & x_1, x_2, x_3 & \geq 0 & \\
\text{var} & x_1, x_2, x_3 & &
\end{array}
$$



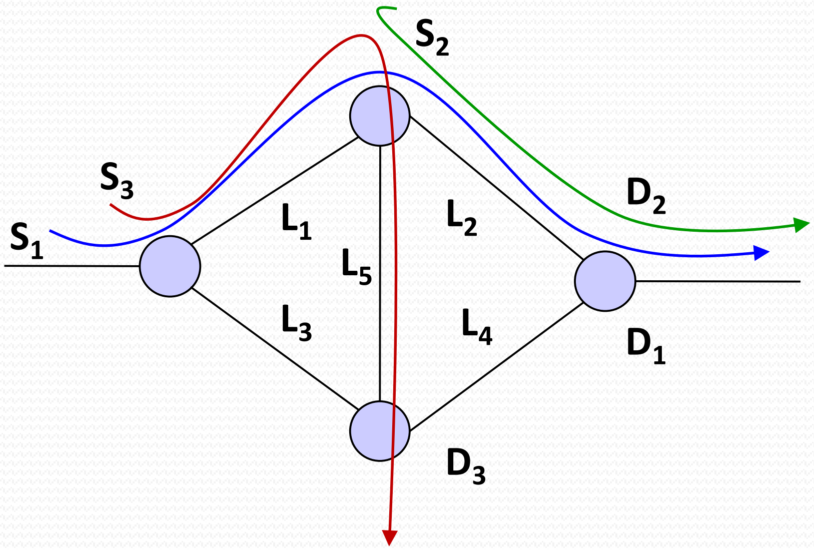
- $L(x,\lambda) = -\log x_1 - \log x_2 - \log x_3 + \lambda_1(x_1+x_3-2) + \lambda_2(x_1+x_2-1) + \lambda_5(x_3-2)$

  Setting $\partial L(x,\lambda)/\partial x_s = 0$:

$$x_1 = 1/(\lambda_1 + \lambda_2) \qquad x_2 = 1/\lambda_2 \qquad x_3 = 1/(\lambda_1 + \lambda_5) = 1/\lambda_1$$

setting $x_1+x_3 = 2$ and $x_1+x_2 = 1 \rightarrow \lambda_1 = 1.732$, $\lambda_2 = 0.634$, $\lambda_3 = \lambda_4 = \lambda_5 = 0$, and the optimal rates are: $x_1 = 0.422$, $x_2 = 1.577$, $x_3 = 0.577$

- **Example:**



In general: the optimal transmission rate for source s is only determined by the Lagrange multipliers on its route. Proof:

$$L(x,\lambda)= - \sum_{s\in S} U_s(x_s) + \lambda^T(Ax-C)$$

Being $U_s(x_s)=m_s \log x_s$

Then applying KKT, $\nabla_x L(x,\lambda)=0$, $\lambda^T(Ax-C) =0$

$$x_s= m_s/(\sum_{s\in Rs} \lambda_s)$$

The Lagrange multipliers associated with a link can be thought as the price for using that link and the price of a route can be thought as the sum of the prices of its links. **The transmission rate, x_s, of a source is inversely proportional to the prices of the links it uses.**

If the price of the route increases → the transmission rate of a source using that route decreases

- **Resource Allocation in Wireless Networks**
  - Let us model the network like a graph and let be $T=[T_{ij}]$ the fraction of time that node i transmits to node j (i.e., **capacity of a link** is 1 packet per unit), and let x_i be the traffic allocated to each user
    - For example, in the picture, T_{12} can not transmit simultaneously with T_{23} since there would be a collision.
    - Links are bi-directional. Then, $T_{12} + T_{23} + T_{32} \leq 1$ in order not to have collisions.

Maximize $\Sigma_{s=1\ldots3} U_s(x_s)$

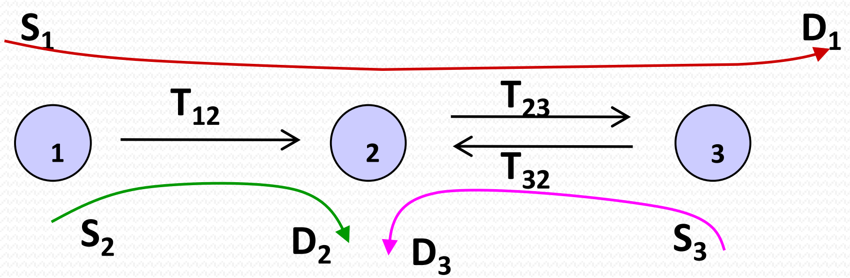
s.t. $x_1+x_2 \leq T_{12}$

$x_1 \leq T_{23}$

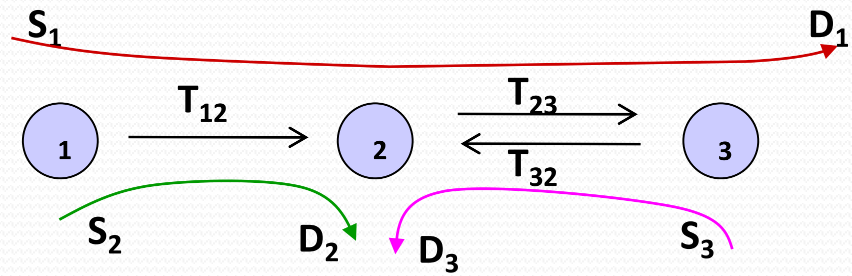
$x_3 \leq T_{32}$

$T_{12} + T_{23} + T_{32} \leq 1$

$x, T \geq 0$



Resource Allocation in Wireless Networks



Let us write the Lagrangian as:

$$L(x,\lambda,\nu) = \text{Max } \Sigma_s U_s(x_s) - \lambda_1(x_1+x_2-T_{12}) - \lambda_2(x_1-T_{23}) - \lambda_3(x_3-T_{32}) - \upsilon(T_{12}+T_{12}+T_{23}-1)$$

This problem can be decomposed as two optimization problems:

Utility Maximization Problem (same as wired networks)

Max $\quad \Sigma_s U_s(x_s) - \lambda_1(x_1+x_1) - \lambda_2 x_1 - \lambda_3 x_3$

s.t. $\quad x_1, x_2, x_3 \geq 0$

Scheduling Problem (λ_1 are weights to determine fractions of time a particular hop is scheduled) → "Maximum Weight Algorithm"

Max $\quad \lambda_1 T_{12} + \lambda_2 T_{23} + \lambda_3 T_{32}$

s.t. $\quad T_{12}+T_{12}+T_{23} \leq 1$ and $T_1, T_2, T_3 \geq 0$

Geometric Programming: application to Power Control

- **Objective:**
  - **Power control** is a mechanism to control the transmission power of devices in order to achieve a good performance in terms of several aspects such as reduction of interference from other devices or reduction of owns energy consumption (battery).
  - Some QoS metrics are non-linear functions of SIR (Signal to Interference Ratio) which is in turn not non-linear (nor concave, nor convex) function of P_t (Transmission power) of devices
  - When SIR is much larger than 0 dB, Geometric Programming may solve the problem. When SIR is comparable or less than 0 dB, the problem is non-linear and difficult to solve

Geometric Programming: application to Power Control

- **Basic Model**: Consider cellular or multi-hop network with n Tx/Rx with transmit powers $P_1, \ldots P_n$
  - The power received by receiver i and transmitted by node j is given by $G_{ij}P_j$, where $G_{ij} \geq 0$ is the path gain proportional to $G_{ij} \cong G_{ij}/d_{ij}^{\gamma}$ and where γ is the power fall-off factor and usually is $2 \leq \gamma \leq 4$. G_{ii} is the gain of the antenna of node i.
  - The SIR (Signal to Interference and Noise Ratio) of the i^{th} receiver is:

$$\mathrm{SIR}_i = \frac{G_{ii}P_i}{\sigma_i + \sum_{k \neq i} G_{ik}P_k} \geq \mathrm{SIR}^{\min} \qquad i = 1, \ldots, n$$

where σ_i is the noise at receiver i, and there are the following restrictions:

$$P_i^{\min} \leq P_i \leq P_i^{\max} \qquad i = 1, \ldots, n$$

Convex Optimization Problems. Applications

• Let us define the following Optimization Problem (OP): Maximize the SIR_i of particular user i*.

Maximize:	$\mathbf{SIR_{i*}\ (P)}$		
subject to:	$\mathbf{SIR_i\ (P) \geq SIR_{i,min}}$	**for all i**	**(1)**
	$\mathbf{P_{i,min} \leq P_i \leq P_{i,max}}$	**for all i**	**(2)**
variables	**P**		

• constraint (1) sets a floor on the SIR of other users and protects these users from user i* increasing her transmit power excessively.

• constraint (2) is system limitation on transmit powers.

• For CDMA systems, another constraint is added: $\mathbf{P_{i1}G_{i1}=P_{i2}G_{i2}}$, that solves the near-far problem of CDMA systems: the expected receiver power from one transmitter i1 must be equal that from another i2,

• In case of having other schemes such OFDM (4G), then other constraints can appear.

Although "Maximize the SIR_i of particular user i*" is not a GP problem, it can easily be converted to one:

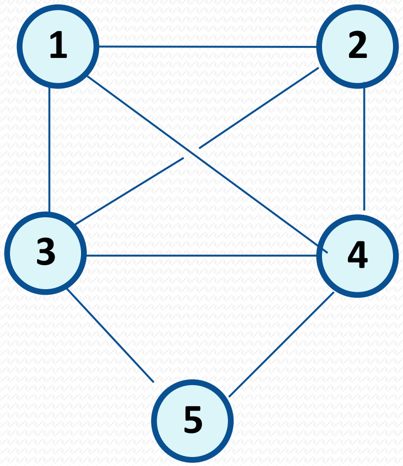
- Objective function: change maximize by minimize (1/SIR), then 1/SIR = $\sigma_i^{-1} G_{ii}^{-1} P_i^{-1} + \sum_{j \neq i} G_{ii}^{-1} G_{ij} P_i^{-1} P_j$, that is a posynomial,
- Constraints:

$$SIR_i = \frac{G_{ii} P_i}{\sigma_i + \sum_{j \neq i} G_{ij} P_j} \geq SIR_{i,\min} \Rightarrow \sigma_i + \sum_{j \neq i} G_{ij} P_j \leq \frac{G_{ii} P_i}{SIR_{i,\min}}$$

$$\Rightarrow \qquad \frac{\sigma_i + \sum_{j \neq i} G_{ij} P_j}{G_{ii} P_i} \leq \frac{1}{SIR_{i,\min}}$$

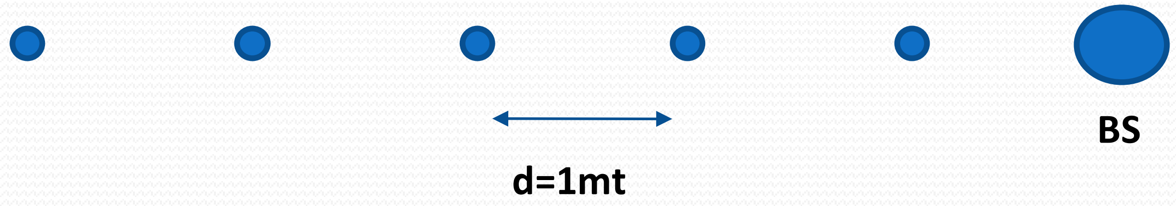
That is a posynomial, and then allow us to write the "Maximize the SIR_i of particular user i*" as a GP problem.

- **Example 1 (Mung Chian Tutorial, section 3.3, pp 78):** N=5 users that have the Gain matrix as expressed below. Obtain the "Maximize the SIR_i of particular user i*" problem.
  - P_{min}= 0.1 W, P_{max}= 1 W,
  - σ_i (i=1,…5) = 0.5 (1.12 dB)
  - SIR_{min} variable (test some, e.g. -5dB, -2.5 dB, 0 dB, 2.5 dB).



$$G = \begin{pmatrix} 1 & 0.1 & 0.2 & 0.1 & 0 \\ 0.1 & 1 & 0.1 & 0.1 & 0 \\ 0.2 & 0.1 & 1 & 0.2 & 0.2 \\ 0.1 & 0.1 & 0.2 & 1 & 0.1 \\ 0 & 0 & 0.2 & 0.1 & 1 \end{pmatrix}$$

- **Example 2 (Mung Chian Tutorial, section 3.3, pp 78):** N=5 users that transmit to a Base Station (BS) at distances of 1 meter of each other. Obtain the “Maximize the SIR_i of particular user i*” problem.
    - Fall-off factor γ=4 → G_{ij}= 0.5 $d^{-\gamma}$, G_{ii}= 1.0,
    - P_{min}= 0.1 W, P_{max}= 1 W,
    - σ_i (i=1,…5) = 0.5 (1.12 dB)
    - SIR_{min} variable (test some, e.g. -5dB, -2.5 dB, 0 dB, 2.5 dB).



- We define **channel outage** as the effect of not being able to recover (decode) a transmission due to fading or interference. Thus, Let us have a **channel outage** (and then a packet lost) when the received SIR falls below a given threshold SIR_{th}.

$$P_{o,i} = \Pr\{SIR_i \leq SIR_{th}\} = \Pr\{G_{ii}P_i \leq SIR_{th} \sum_{j\neq i} G_{ij}P_j\} \leq P_{o,i,max}$$

assuming low thermal noise σ_i with respect interference, and exponential distributions, then, the outage probability can be expressed as:

$$P_{o,i} = 1 - \prod_{j\neq i} \frac{1}{1 + \frac{SIR_{th} G_{ij} P_j}{G_{ii} P_i}} \leq P_{o,i,max}$$

and then can be written as a posynomial in variable **P**:

$$\prod_{j\neq i} \left(1 + \frac{SIR_{th} G_{ij} P_j}{G_{ii} P_i}\right) \leq \frac{1}{1 - P_{o,i,max}}$$

• Constellation size M (MQAM modulation) used by a link can be approximated as:

$$M = 1 + \frac{-\varphi_1}{\ln(\varphi_2 \mathrm{BER})} \mathrm{SIR} = 1 + K \cdot \mathrm{SIR}$$

and the **Data Rate** on link *i* will be: $R_i = Bw \log_2(1+K\, SIR_i)$

That can be approximated as:

$$R_i = Bw \log_2(K\, SIR_i)$$

when K·SIR is much larger than 1 (0 dB)

• For the following assume that constant K is absorbed by SIR_i formula and the aggregated data rate will be:

$$R_{system} = \Sigma_i R_i = Bw\, \Sigma_i \log_2(K\, SIR_i) = Bw \log_2(\Pi_i SIR_i)$$

Aggregate data maximization is equivalent to maximizing product of SIR

Note that: $\boxed{\text{Maximize } \Sigma_i \log(h_i(x)) \cong \text{Minimize } \Pi_i h^{-1}_i(x)}$

Convex Optimization Problems. Applications

- Let us now, maximize the total system throughput under user data rate constraints and outage probability constraints:

Maximize:	$R_{system}(P)$		
subject to:	$R_i(P) \geq R_{i,min}$	**for all i**	**(1)**
	$P_{o,i}(P) \leq P_{o,i,max}$	**for all i**	**(2)**
	$P_{i,min} \leq P_i \leq P_{i,max}$	**for all i**	**(3)**
variables	P		

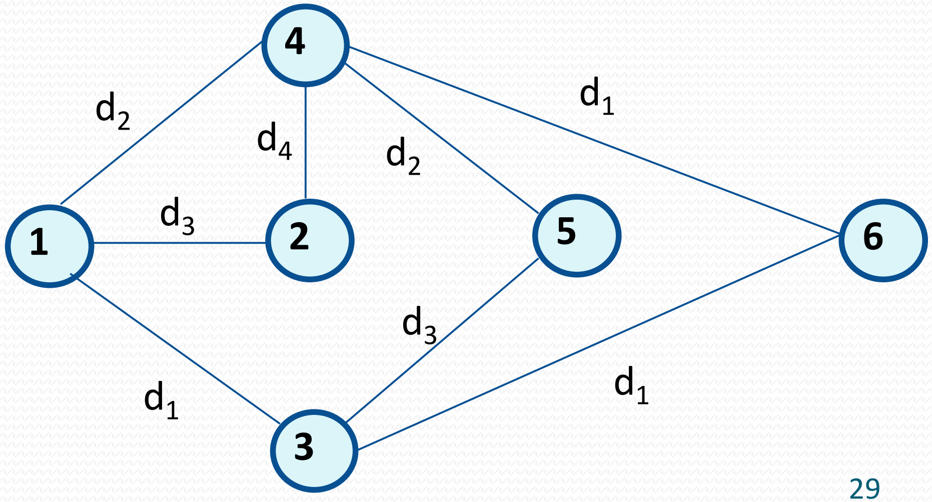
  - constraint (1) sets the minimum data rate demanded by user i.
  - constraint (2) fixes the outage probability upper bounds $P_{o,i,max}$
  - constraint (3) is system limitation on transmit powers

- Other variations of this problem are to lower bound R_{system} as a constraint and maximize R_{i*} for a particular user i*. We also can use the total power $\Sigma_i P_i$ as constraint or objective function.

- **Example 3:** Let's have a network like the figure.
  - Fall-off factor $\gamma=4 \rightarrow G_{ij}= G\, d^{-\gamma}$, $G_{ii}= 1.0$, with $G=10^5$. Distances: $(d_1, d_2, d_3, d_4) = (70, 60, 50, 30)$ mt.
  - $P_{min}= 10$ mW, $P_{max}= 1$ W,
  - σ_i (i=1,…n) = 10^{-6} (-60 dB) and $SIR_{min}= 10^{-5}$ (-50 dB)
  - R_{min}=1 Kb/s (for a K=4), Bw= 1 KHz.
  - Outage probability $P_{o,i,max} = 10^{-4}$,

Obtain

- Maximize "the SIR_i of particular user i*" problem
- Maximize "the R_{system}" problem.



- **Geometric Programming: application to Queueing Theory**
  - Let us assume simple Queueing models and see that we can optimize several intrinsic parameters
  - M/M/1 queue system: a system whose arrival rate is Poisson distributed of parameter λ and the service is exponentially distributed of parameter μ.
    - The load of the system is $\rho= \lambda/\mu$
    - The average number of customers in the system: $N= \rho/(1-\rho)=\lambda/(\mu-\lambda)$
    - The average delay per customer in the system: $D= N/\lambda = 1/(\mu-\lambda)$
    - The average delay per customer in the queue: $W = 1/(\mu-\lambda)-(1/\mu) = \rho/(\mu-\lambda)$
    - The average number of customers in the queue: $Q = \lambda W = \rho^2/(1-\rho)$

Geometric Programming: application to Queueing Theory

- The following optimization problem is GP:

Maximize: μ/λ

subject to: $W \le W_{max}$

$D \le D_{max}$

$Q \le Q_{max}$

$\lambda \ge \lambda_{min}$

$\mu \le \mu_{max}$

variables λ, μ

- Where W_{max}, D_{max}, Q_{max} are bounds on the maximum waiting time in the queue, maximum total delay and maximum number in the system.
- λ_{min} Bounds the incoming traffic and μ_{max} bounds the service rate

- **Geometric Programming: application to Queueing Theory:**
  - A telephone call center: maximize the probability that a particular number of telephone lines are in use at a given time
    - M/M/m/m queue: Poisson arrivals of parameter λ and service is exponentially distributed of parameter μ, m servers, the last m indicates the limit of users in the system (m-server loss system).
    - The probability to stay in state k (k servers are occupied) is:

$$p_k = \frac{\left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!}}{\sum_{i=1}^{m} \left(\frac{\lambda}{\mu}\right)^i \frac{1}{i!}}$$

that is called the Erlang-B formula

and give us the probability that k servers are occupied.

- **Geometric Programming: application to Queueing Theory:**
  - A telephone call center: maximize the probability that a particular number of telephone lines are in use at a given time. The problem is GP.

Maximize:	$p_k(\lambda,\mu)$
subject to:	$p_j(\lambda,\mu) \geq C_j \quad j=1,\ldots,m$
	$\lambda \geq \lambda_{min}$
	$\mu \leq \mu_{max}$
variables	λ, μ

  - There are other optimization problems related to queueing theory and Markovian chains in which buffer overflows or stationary probabilities are optimized and that are used in many telecommunication systems.