

Math 420/620 Final Question 1

1. Weather forecasting software has to produce a forecast within 4 hours to meet the quality of service agreed upon by a government contractor. Currently the program takes 5 hours to run on an Origin 2000 parallel computer consisting of 16 processors. The contractor is looking for ways to speed the computation.

(i) About 60 percent of the time all 16 processors work in parallel, but since not all the work can run in parallel, 40 percent of the time only 1 processor is active. By Amdahl's law

$$M(P/16 + 1 - P) = 5 \text{ hours}$$

where M is how long the program would run using only one processor and P represents the percent of the total computation that is parallelizable. Thus,

$$MP/16 = 3 \text{ hours} \quad \text{and} \quad M(1 - P) = 2 \text{ hours.}$$

Solve for M and P .

Since $MP/16 = 3$ then $MP = 48$. It follows that

$$M(1 - P) = M - MP = M - 48 = 2$$

Consequently, $M = 50$.

Now, substitute back into the first equation to solve for P . Thus,

$$P = 48/M = 48/50 = 24/25 = 0.96.$$

```
In [1]: M=50; P=24/25
```

```
Out[1]: 0.96
```

(ii) How many more processors are needed to meet the 4 hour deadline?

Consider a parallel computer with n total processors. We must solve for the least n such that

$$M(P/n + 1 - P) \leq 4.$$

Thus,

$$MP/n \leq 4 + MP - M$$

or

$$n \geq \frac{MP}{(4 + MP - M)} = \frac{48}{4 + 48 - 50} = 24.$$

Since a total of $n = 24$ processors would be sufficient, then 8 more processors are needed to upgrade the current 16 processor computer to 24.

(iii) There was an accident and the old computer destroyed. A new computer will be purchased with n processors that each run at a speed s . Here $s = 200$ MHz corresponds to the processors in the old computer. Explain why

$$(200/s)M(P/n + 1 - P) = 4 \text{ hours}$$

ensures the new computer will produce a forecast in 4 hours.

Increasing the speed of the processors proportionately increases the speed at which the weather forecast is finished while holding the number of processors fixed. For example, doubling the speed means the time taken is only half. Since the time taken by the original computer was given by Amdahl's law as

$$T_0 = M(P/n + 1 - P),$$

faster processors will complete the work in time

$$T(s) = (200/s)T_0 = (200/s)M(P/n + 1 - P).$$

The condition above comes by taking $T(s) = 4$.

(iv) Processors cost $c(s) = As^3$ where A is a constant and s is processor speed. Assuming the processors in the old computer cost \$24000, solve for A .

We want $c(200) = 24000$. Thus $A200^3 = 24000$, which implies

$$A = \frac{24000}{200^3} = 0.003.$$

We remark that A is a dimensional constant with units $\$/\text{MHz}^3$. The power dissipation in integrated circuits scales as $sV(s)^2$ where experience suggests voltage V must be scaled linearly with frequency. Thus, the expected power dissipation of a faster processor scales at least as s^3 .

The actual pricing of processors will be related to supply and demand constraints as well as engineering constraints, so while $c(s) = As^3$ may hold for s near 200, many assumptions have been made.

```
In [2]: A=24000/200^3
```

```
Out[2]: 0.003
```

(v) Suppose the processors are the main expense. Then the total cost to build a computer with n processors is $C(n, s) = nAs^3$. What values of n , s and $C(n, s)$ correspond to the cheapest computer that can produce a forecast in 4 hours?

Our goal is to minimize $C(n, s)$ subject to the constraint

$$(200/s)M(P/n + 1 - P) = 4.$$

This can be done using Lagrange multipliers; however, since the constraint is easily solved for s , we instead substitute

$$s = 50M(P/n + 1 - P)$$

into the objective function to obtain

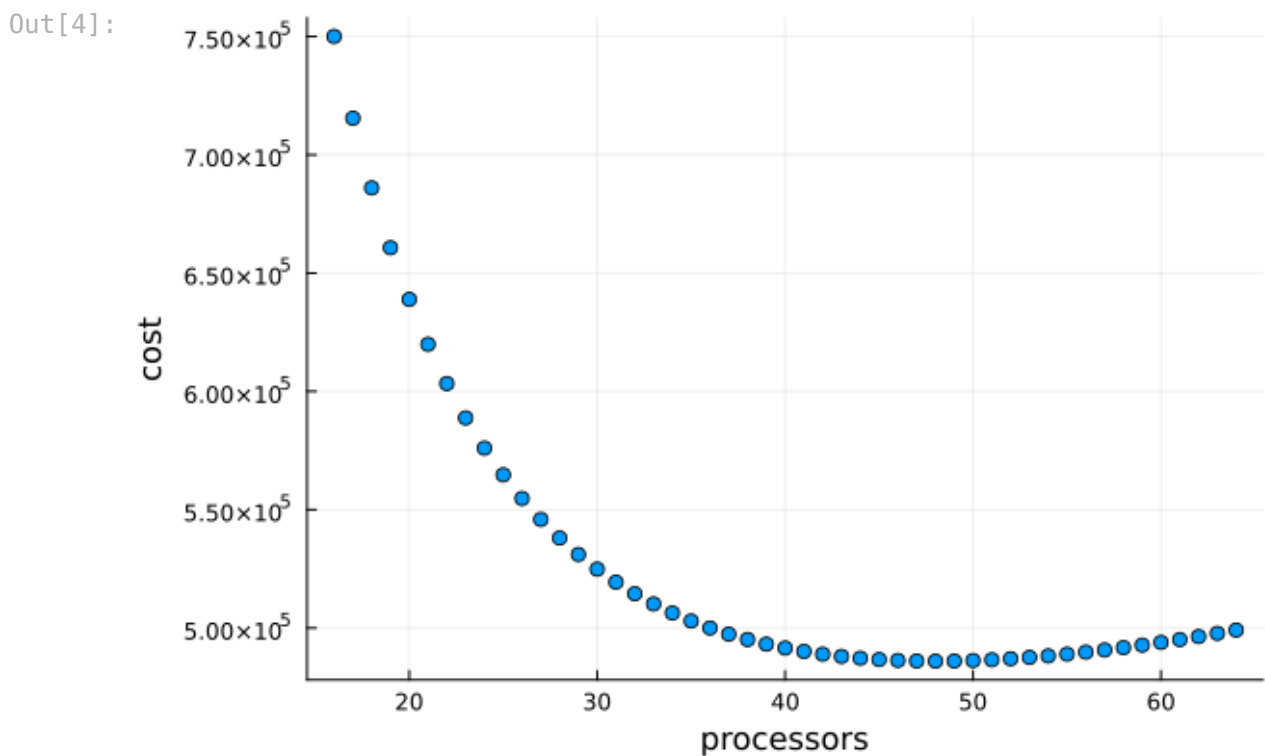
$$C(n) = nA50^3M^3(P/n + 1 - P)^3.$$

Now search for the n which minimizes $C(n)$.

Now solve $C'(n) = 0$ and if n is not an integer check $\lfloor n \rfloor$ and $\lceil n \rceil$ to determine which is the least. Begin by plotting C as a function of n

In [3]: `using Plots`

```
In [4]: s(n)=50*M*(P/n+1-P)
C(n)=n*A*s(n)^3
scatter(C, 16:64, legend=:false,
        xlabel="processors", ylabel="cost")
```



From the graph it appears the optimal number of processors is somewhere between 40 and 60.

```
In [5]: using Symbolics
        D(f,x)=expand_derivatives(Differential(x)(f))
```

```
Out[5]: D (generic function with 1 method)
```

```
In [6]: @variables x
```

```
Out[6]: 1-element Vector{Num}:
         x
```

```
In [7]: fn=D(C(x),x)
        fs="f(x)="*string(fn)
        eval(Meta.parse(fs))
```

```
Out[7]: f (generic function with 1 method)
```

Now set $f(x) = C'(x)$ and use Newton's method

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

to solve for the x where $C'(x) = 0$.

```
In [8]: dfn=D(f(x),x)
        dfs="df(x)="*string(dfn)
        eval(Meta.parse(dfs))
```

```
Out[8]: df (generic function with 1 method)
```

```
In [9]: xk=50.0
        for k=1:10
            xk=xk-f(xk)/df(xk)
        end
        xopt=xk
```

```
Out[9]: 47.99999999999996
```

```
In [10]: Copt=C(48)
```

```
Out[10]: 486000.00000000146
```

```
In [11]: C(47)
```

```
Out[11]: 486071.9782707112
```

Since $C(48) < C(47)$ we conclude that the number of processors that minimize cost is $n = 48$.

```
In [12]: nopt=48  
s(48)
```

```
Out[12]: 150.00000000000014
```

Moreover, the speed of the processors that minimize cost is $s = 150$ MHz and the minimal cost of the replacement computer is $C(n, s) = 486000$ dollars.

(vi) Find the relative sensitivity $S(C, P)$ of the optimal cost C depending on the percentage P of parallelizable work found in part (i). Interpret your result.

```
In [13]: # define Cp which is a function of the parallelizable work  
sp(p)=50*M*(p/nopt+1-p)  
Cp(p)=nopt*A*sp(p)^3
```

```
Out[13]: Cp (generic function with 1 method)
```

```
In [14]: @variables p
```

```
Out[14]: 1-element Vector{Num}:  
p
```

```
In [15]: Cp(p)
```

```
Out[15]: 2.2500000000000005e9((1 - p + p / 48)^3)
```

```
In [16]: dCdp=D(Cp(p), p)
```

```
Out[16]: -6.609375000000001e9((1 - p + p / 48)^2)
```

```
In [17]: substitute(p/Copt*dCdp, p=>P)
```

```
Out[17]: -46.99999999999993
```

The sensitivity is negative, which reflects the fact that if the amount of parallel work in the weather forecasting computation is more than estimated the cost will go down. If the amount of parallel work is less than estimated the cost will go up. Note also the fact that $S(C, P) \approx -47$ is large in magnitude reflects that the optimal cost depends strongly on the estimate for the amount of parallel work.

```
In [ ]:
```